

# **Disconnection of drug-response and placebo-response in acute-phase antipsychotic drug trials on schizophrenia? Meta-regression analysis**

**NPP-19-0113**

Stefan Leucht, MD,<sup>1\*</sup> Anna Chaimani, PhD,<sup>2\*</sup> Dimitris Mavridis, PhD,<sup>3</sup> Claudia Leucht, MD,<sup>1</sup>

Maximilian Huhn, MD,<sup>1</sup> Bartosz Helfer, PhD,<sup>4</sup> Myrto Samara, MD,<sup>1</sup> Andrea Cipriani, MD, PhD,<sup>5,6</sup> John

R Geddes, MD,<sup>5,6</sup> John M. Davis, MD<sup>7</sup>

## **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](mailto:Stefan.Leucht@tum.de)

\*These authors have contributed equally to the manuscript.

1. Department of Psychiatry and Psychotherapy, Technische Universität München

Klinikum rechts der Isar

2. Paris Descartes University; INSERM, UMR1153 Epidemiology and Statistics, Sorbonne Paris

Cité Research Center (CRESS), METHODS Team; Cochrane France, Paris, France

3. Department of Primary Education, University of Ioannina School of Education, University Campus

Ioannina 45110, Ioannina, Greece

4. Department of Medicine, Imperial College London, London W2 1PG, UK

5. Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK

6. Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford OX3 7JX, UK

7. Psychiatric Institute, University of Illinois at Chicago (mc 912), 1601 W. Taylor St., Chicago, IL

60612, and Maryland Psychiatric Research Center, Baltimore, MD, USA

Word count: 3457

## ABSTRACT

Differences in efficacy between antipsychotics and placebo in schizophrenia trials have decreased over the past decades. Previous studies have tried to identify potential explanatory factors focusing on response to placebo; however, it is still not clear which factors, if any, specifically moderate *drug-response*, as they may be different from those moderating *placebo-response*.

Therefore, in this meta-regression analysis we explore whether there is an interaction between drug-response and placebo-response in terms of effect size. We systematically searched multiple electronic databases, ClinicalTrials.gov and the US Food and Drug Administration website for randomized, placebo-controlled trials investigating the efficacy of antipsychotics in patients with acute schizophrenia (last update: October 2016). The main outcome was the change on the Brief Psychiatric Rating Scale or the Positive and Negative Syndrome Scale score from baseline to endpoint after at least 3 weeks of treatment. Multiple patient-, design-, and drug-related potential predictors of response were analyzed by meta-regressions, as predefined in the study protocol.

Overall, 167 trials with 28,102 participants were included. Publication year, the number of participants and sites, mean dose, minimum severity threshold as an inclusion criterion, chronicity, industry sponsorship, type of rating scale, diagnostic criteria and number of medications had a different impact on drug and placebo response. By contrast, baseline severity, duration of wash-out, study duration, and study region affected drug and placebo response in a similar way without a net effect on effect sizes. No other factors had a significant effect on either drug-response or placebo-response.

In conclusion, as individual moderators may have different impact on placebo-response and drug-response, it is important to consider also the specific factors influencing drug-response in order to fully understand the difference between antipsychotics and placebo. These results shed further light on the phenomenon of decreasing effect size of antipsychotics for schizophrenia over time and should help design future randomized controlled trials in the field (Prospero registration number CRD42013003342).

## INTRODUCTION

Differences in efficacy between antipsychotic drugs and placebo in clinical trials of acute schizophrenia have decreased over the decades and we have recently demonstrated that this is to large degree due to an increasing placebo response [1]. This phenomenon might help to account for some unexpected trial results where even standard drugs such as haloperidol did not outperform placebo [2]. Implications of these results pose a challenge for antipsychotic drug development as well as for assessment of their efficacy in everyday clinical practice.

We previously have demonstrated that drug-placebo differences can be explained mainly by the increasing placebo response and to a lesser extent industry sponsorship (~~which is probably a composite of various factors such as large sample size, multiple centers and countries, recruitment in professional centers~~). When we investigated factors explaining *placebo*-response in a subsequent paper, lower chronicity and higher sample size were correlated with this phenomenon [3], replicating some of the findings of previous publications which were based on approximately 2 times smaller datasets [4-6].

However, as effect sizes are defined as a difference between drug effects and placebo effects, the remaining question is which factors moderate *drug-response* and whether these are different from those moderating placebo-response. Intuitively, one would assume that a factor leading to an increase in placebo response would lead to a parallel increase of drug response (this is sometimes referred to as the “additivity assumption” [6,7]). In our original paper on drug-placebo differences, we found that the additivity assumption is not always fulfilled. While placebo-response increased over the years, drug-response remained stable, leading to decreasing effect sizes over the decades [1].

Therefore, here we try to identify modifiers of drug-response and explore how they interplay with modifiers of placebo response in the forming of effect sizes. Our results will provide further in-depth

clarification for the decrease in drug-placebo differences over the years and may have an impact on the design of future placebo-controlled clinical trials of antipsychotic drugs.

## **MATERIAL AND METHODS**

We followed the PRISMA guidelines [8] (see checklist in the data supplement S1) and initially published a protocol in PROSPERO (CRD42013003342, see data supplement S2). We used the same database and largely the same text of the method section as in our previous publications for consistency [1] [3].

### **Inclusion/exclusion criteria**

#### 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 [9]. We excluded relapse prevention studies in stable patients receiving maintenance medication [10], studies in patients with predominant negative symptoms, and studies in patients with major concomitant somatic or psychiatric illness. There is no indication that effect sizes have decreased in relapse prevention studies in stable patients, differences between drug and placebo remain large [10]. The outcome in patients with predominant negative symptoms would be different (negative symptoms rather than overall symptoms). Studies in patients with major concomitant or psychiatric illness are extremely rare and would have increased heterogeneity even further.

#### Interventions

We included all antipsychotics licensed in at least one country, except clozapine, a more efficacious drug [11] 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 [12] 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 drug arms of the

included studies were used for the current analysis. We examined all antipsychotics as a group under the assumption that efficacy differences between drugs are small [11,13-15], except for clozapine, which was excluded for this reason.

### Types of studies

Published and unpublished, double-blind, placebo-controlled randomized controlled trials of at least 3 weeks duration [16] were included. Studies with a high risk in sequence generation or allocation concealment were excluded [17]. We a priori excluded Chinese studies due to quality concerns [18,19] 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 [17].

### **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 [20] 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 [5,11,21-29]. week

### **Outcomes**

The outcome was the mean change from baseline to endpoint of the Positive and Negative Syndrome Scale (PANSS [30]) total score. If the PANSS was not available we used the change from baseline to endpoint of the Brief Psychiatric Rating Scale (BPRS [31]) and converted it to the PANSS using a validated method [32].

### **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). It is nowadays a standard that analyses should be based on intention-to-treat datasets, and it has been shown that MMRM is a more appropriate statistical approach to account for attrition in psychopharmacological trials than LOCF [33]. Missing standard deviations were estimated from test statistics or by using the mean standard deviation of the remaining studies [34,35].

### **Statistical analysis**

We conducted meta-regressions in a frequentist framework with drug-response as dependent variables and compared them with our previously published results on placebo-response [3] and drug-placebo differences [1]. All the potential predictors that we had used in our previous papers [1,3] were analyzed as independent variables. Drug and placebo-response were continuous variables defined as the difference in PANSS/BPRS scale before and after treatment. Drug-placebo differences were calculated as standardized mean differences. The initial choice of predictors had been based on previous evidence [4-6,36,37] which suggested that these factors might be relevant. We categorized the moderators into patient-, study design-, and drug-related factors, although there were expected overlaps. We first ran univariable meta-regressions exploring separately the effect of each potential moderator. For the multivariable meta-regression models we only used factors that were significant in univariable analyses and we followed a formal variable selection procedure using the backward-stepwise algorithm with removal criterion  $p=0.15$ . We monitored how much heterogeneity in drug response, placebo response and drug placebo differences each predictor explains by comparing the heterogeneity of each meta-regression model with the heterogeneity of the model without any covariates.

### Patient-related factors

The patient-related factors were: chronicity measured by the patients' mean age, duration of illness, duration of the current episode and first episode status [5,37], percentage men [37], US American populations versus not/mixed countries [38], severity (PANSS total score) at baseline [36], in- versus outpatient [5], operationalized criteria (e.g. ICD-10 or DSM-III to IV-R) versus unspecific 'clinical diagnoses', and the total dropout rates of both groups combined, because the joint process of response with dropout is rarely accounted for appropriately [39,40].

### Study design-related factors

We analyzed publication year, the impact of risk of bias (appropriate versus unclear randomization [41] and allocation concealment methods [42], blinding [42], and missing outcome data [17,43]), study duration [5], duration of wash-out [5], requirement of a scale-derived minimum of symptoms at baseline [36], PANSS versus BPRS as a scale, sample size [44], number of sites [5], percentage of academic sites [5], number of medications and arms [5], percentage of participants randomized to placebo [4], and drug company sponsorship of at least one study arm (medication donation alone was not considered company sponsorship [45]).

### Drug-related factors

We classified the antipsychotics by their mechanisms according to the "Neuroscience-based Nomenclature" [46], antipsychotic doses in chlorpromazine equivalents according to the International-Consensus-Study-of Antipsychotic-Dosing [6,47], and fixed versus flexible dose studies [4].

We also analyzed whether the degree of placebo response in the studies was associated with the degree of drug response.

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

## **RESULTS**

### **Description of included studies**

The PRISMA flow diagram [8] is presented in the online supplement Figure S1 and a description of the included studies in supplement Table S4. A summary of study characteristics is presented in Table 1.

Overall, 167 studies with 28102 participants met the inclusion criteria, of which 104 studies with 23567 participants (8023 allocated to the placebo groups, and 15544 to the drug groups), which were published between 1969 and 2016 provided data to calculate drug-placebo differences. In the studies with such data, the patients' mean duration of illness was 13.8 (SD 4.0) years, the mean age 38.6 (SD 4.8) years and the median duration of studies with useable outcomes was 6 weeks (range 3-26 weeks, for the outcome examined here all studies except one (26 weeks) lasted  $\leq 12$  weeks). There were no studies in first-episode patients or in treatment-resistant patients. Risk of bias is presented in the online supplement S5. We only included randomized, 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 reflected the high dropout rates in current schizophrenia studies (overall mean 39.3%, SD 17.1). Older studies were poorly reported, making it often impossible to extract data (52 studies (50%) of the had a high risk of selective reporting). 65 studies (62.5%) were sponsored by the manufacturers of one antipsychotic included, 31 (29.8%) were not primarily industry sponsored and in 8 (7.7%) studies the sponsor was unclear.

### **Analysis of potential moderators – univariable analysis**

The mean drug response in PANSS units was 17.45 (95% CI 15.89,19.01; 100 studies (N) with 14933 participants (n)), the placebo response was 6.25 (95% CI 4.64,7.85; 99 studies (N) with 7623 participants (n)) and the mean SMD was 0.47 (95% CI 0.42,0.52,  $p < 0.001$ ). As expected, drug-response and placebo-response were strongly correlated (Figure 1).

In the following text, in Table 2 and in Figure 2 we sorted the moderators in the following way:

- 1) We first present the moderators that had a significant effect on the drug-placebo difference (effect size, SMD), because it is the drug-placebo difference that ultimately counts.
- 2) We then present the results of the moderators that had a significant effect on either drug-response or placebo-response, but without an important effect on the resulting drug-placebo difference.
- 3) The moderators that had no important effect on any result are summarized at the end.



1. Moderators with a significant effect on drug-placebo differences (SMDs), Figure 2a – Figure 2l

Placebo-response was significantly higher in more recent studies, in studies with a larger number of participants and sites, in studies with a minimum baseline severity as an inclusion criterion, use of PANSS instead of the BPRS and in studies with operationalized diagnostic criteria, whereas drug-response was not significantly associated with these factors. The net effect were smaller drug-placebo differences (Figures 2a, 2b, 2c, 2e, 2i, 2j; Table 1).

Studies using higher mean doses in chlorpromazine equivalents had significantly smaller placebo-response and tended to have more drug response (not significant), resulting in larger drug-placebo differences (Figure 2d, Table 1).

The chronicity measures (mean age and mean duration of illness) were negatively correlated with both drug-response and placebo-response. However, the effect was more pronounced in the drug-groups, leading to smaller drug-placebo differences in more chronic patients, although this effect was significant only for mean age (Figures 2f, 2g).

Industry sponsorship, the number of medications and drug mechanism according to the “Neuroscience-based Nomenclature (NbN)” [46] had no significant impact on neither drug-response nor placebo-response, but the slopes were different enough that drug-placebo differences were significantly smaller in industry-sponsored studies, in studies on drugs with another primary mechanism than D<sub>2</sub>-antagonism by NbN and in studies with only one medication (Figures 2h, 2k, 2l; Table 1).

A post-hoc sensitivity analysis following a reviewer request showed that the pattern of increasing placebo response, relatively stable drug response and decreasing drug placebo differences was present before and after the year 2000. The mean coefficients were virtually the same, although confidence intervals show that not all result that were statistically significant in the main analysis were significant in the sensitivity analysis. The smaller number of trials is a likely explanation (supplement Figure S2).

2. Moderators with a significant effect on drug-response and/or placebo-response but without a significant impact on drug-placebo differences (SMDs), Figure 2m – Figure 2p

Studies with higher baseline severity, shorter duration of the wash-out phases, shorter study duration, and studies conducted in countries outside the US or mixed had higher drug-response than their counterparts. But placebo response was affected in the same direction with similar slopes resulting in no significant effect on drug-placebo differences (Figures 2m, 2n, 2o, 2p; Table 2).

3. Moderators without a significant effect on neither drug-response, nor placebo-response nor drug-placebo differences (SMDs), Figure 2q – Figure 2z

The total drop-out rate, the percentage of patients randomized to placebo, the number of arms, the percentage of academic sites, fixed or flexible dosing, percentage of men, and risk of bias in terms of randomization, allocation concealment, blinding and missing outcome data had no important impact on the results.

**Moderators of drug response, placebo response and drug-placebo differences – multivariable analysis**

As some significant predictors are naturally related to each other, we made the following choices for the multivariable models:

- a) We chose mean participant age rather than mean duration ill as a measure of chronicity, because more studies reported this outcome;
- b) We did not include placebo response in the multivariable model of drug response and vice versa because both are strongly correlated (see Figure 1).

Significant factors in the model for drug response included: average age and baseline severity; in the model of placebo response: average age and total number of participants; and for drug-placebo differences: degree of placebo response and industry sponsorship (industry-sponsored trials had smaller effect sizes than non-industry-sponsored ones, see Table 3).

## DISCUSSION

The current analysis provides important insights in the interactions of drug-response and placebo-response in the formation of effect sizes. In our analysis, we found that drug-response differences, placebo-response differences and drug-placebo differences in effect sizes are not affected by the same factors in the same way. Certain factors influenced only either drug-response or on placebo-response so that they had a net effect on drug-placebo differences. Other factors affected drug-response and placebo-response in the same direction and to a similar degree so that they had no net effect on the effect size. Finally, some factors had no influence for any of these relationships.

~~Starting out with~~

### *Factors that had an impact on drug-placebo differences (effect size)*

We have previously demonstrated that placebo-response has increased over the years while drug-response remained stable resulting in decreasing drug-placebo differences (Figure 2a [1]). Here we found a similar pattern of results for factors “sample size” and the related factor “number of sites” (Figures 2b and 2c). Specifically, the more participants and sites, the more placebo-response and the smaller the effect sizes, while drug-response was unaffected by sample size and number of sites. The reason for this is difficult to interpret. We speculate that drug-response may plateau, while in studies with large sample size more patients who benefit from placebo are recruited compared to smaller trials where patients can be more carefully selected.

The effects on drug response and placebo response differed even more with regard to the doses used. In studies with high antipsychotic doses drug-response increased while placebo-response decreased and consequently effect size increased (Figure 2d). It could be that the patients in studies with high doses are more severely ill so that they benefit more from drug and less from placebo. An alternative interpretation

is that in studies with high doses there is more un-blinding by side-effects so that raters may guess the group patients are assigned to.

In studies that employed a threshold for minimum severity as an inclusion criterion we found that placebo-response was higher than in studies without such a criterion, while drug-response is not affected (Figure 2e). One reason could be artificial baseline inflation in such studies. If the baseline ratings were inflated to meet an inclusion criterion, the next rating of patients in the placebo-group may be automatically lower, while in the drug group we may see a true reduction of symptoms. More chronic patients (the moderator mean age) respond less well to drug and to placebo but the effect was more pronounced in the drug groups so that it could be helpful to recruit younger patients than it is currently the case (mean age around 40 years, Figure 2f). If, as in some previous reports [3,4], we would have only analyzed placebo-response, the conclusion might have been to use more chronic patients to avoid placebo-response, demonstrating the importance of examining drug-response simultaneously. For the related factor “duration of illness” the influence on the drug-placebo difference was not significant, but fewer studies reported this factor and the pattern was the same (Figure 2g).

Industry-sponsoring tended to reduce drug response and tended to increase placebo response, which resulted in smaller effect sizes than those of non-sponsored studies (Figure 2h). This finding is counterintuitive in the sense that the pharmaceutical industry is suspected for designing trials in a way that inflates positive results [48]. However, industry sponsorship probably is a composite of multiple factors.

Studies using PANSS rather than BPRS, studies analyzing other antipsychotics than D<sub>2</sub> receptor antagonists according to NbN, and studies applying operationalized diagnostic criteria had higher placebo-response, similar drug-response and smaller drug-placebo differences than their counterparts (Figures 2i-k). However, this may be confounded by publication year. In early studies operationalized diagnostic criteria and the PANSS were not available yet, and early studies examined D<sub>2</sub> receptor antagonists (mainly haloperidol) more frequently than recent studies.

Finally, studies which examined two or more active drugs had larger drug-placebo differences than those with only one drug. This result stands in contrast the finding that the number of arms in a study (which could also be due to different dose arms of the same drug being examined) did not affect the effect size (Figures 2l and 2s).

~~We next discuss the~~

***Factors that had an impact on either drug-response or placebo-response or both, but not on drug-placebo differences***

Baseline severity data shows a pattern that could be expected based on the findings from a previous individual-patient data meta-analysis [36]. The more ill the patients were at baseline the higher the difference between drug-response and placebo-response, although the effect on the drug-placebo difference was not significant (Figure 2m). In contrast, the factors: duration of the wash-out period, study duration and study origin affected drug-response and placebo-response in the same way so that they had no impact on the drug-placebo difference (Figures 2n, 2o, 2p). This finding is important, because from previous analyses examining placebo-response only, one might have concluded that washout phases need to be long enough to have a low placebo-response. However, our data suggest that longer wash-out phases may not be an optimal solution, because drug-response was also lower in studies with long wash-out periods.

Finally, the factors total dropout rate, percentage randomized to placebo, number of arms, percentage of academic sites, fixed or flexible dosing, and the various risk of bias items had *no important effect on drug-response, placebo-response and drug-placebo differences* (Figure 2q-2z). Reasons why, in contrast to some of our previous analyses, we found an effect of the percentage of trial participants randomized to placebo group [4,5], the percentage of academic sites [5] or the percentage of men in a study [37], could be due to somewhat different definitions, or it could simply be due to our much larger number of trials available for the current analysis.

Given that various factors did not affect drug-response, placebo-response and drug-placebo differences in the same way, it is not surprising that the factors which were significant in the multivariable analyses were also not the same (Table 3). It should be noted that multivariable models are subject to multicollinearity and sometimes lead to over-simplified models by erroneously dropping variables with important contribution to model fit. In contrast, univariable models have a risk of confounding as moderators are correlated. Both results should therefore be considered. To our knowledge the publication by Rutherford et al. [6,47] is the only other one that analysed the interaction between drug-response and placebo-response. Comparison is difficult, because at that time only approximately half of the placebo-controlled studies were available, and because studies comparing antipsychotics with each other were included as well. This is probably the reason why drug-response increased in their analysis while it remained stable in ours.

Our analysis is limited by the fact that we have examined antipsychotics as a class, because not enough studies would have been available for single drugs. Moreover, meta-regression is based on mean values of trials which makes the results prone to ecological bias [49], meaning that an average value representing a large group of individuals is rarely accurate in describing a specific individual from that group. Individual patient data meta-analysis is more powerful in this regard, but unfortunately it is limited by insufficient data availability.

Previous analyses had shown that placebo-response in antipsychotic drug trials has increased over the years [3,5,6], that drug response has remained stable [1] and that as a consequence drug-placebo differences have decreased [1]. The current analysis adds the following information to these analyses: The predictors of drug response are not the same as those of placebo response and of effect sizes. Therefore, it is not sufficient to only understand placebo response in such trials. We also need to understand drug-response, because its interaction with placebo-response forms the drug-placebo difference which counts for patients. ~~Nevertheless, our analysis provided the following major insight: different factors affect~~

~~drug response, placebo response and effect sizes in different ways. Therefore, it is not sufficient to only understand placebo response in such trials. We also need to understand drug response, because their interaction forms the drug-placebo difference which counts for patients.~~

## **FUNDING AND DISCLOSURES**

The meta-analysis was supported by the German Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF) Grant: FKZ 01KG1115 and by the National Institute for Health Research (NIHR) Oxford Health Biomedical Research Centre (grant BRC-1215-20005). The funding body was not involved in the study design, collection, analysis, interpretation of data, and in the decision to submit for publication. The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health. Authors had full access to the study data and complete discretion in the analysis of data and writing of this report. In the last 3 years, Stefan Leucht has received honoraria for consulting or lectures from LB Pharma, Otsuka, Lundbeck, Boehringer Ingelheim, LTS Lohmann, Janssen, Johnson&Johnson, TEVA, MSD, Sandoz, SanofiAventis, Angelini, Recordati, Geodon Richter. Claudia Leucht is Stefan Leucht's spouse. Maximilian Huhn received lecture honoraria from Janssen and Lundbeck. The other authors have no conflicts of interest to declare.

## **AUTHOR CONTRIBUTIONS**

SL, ACh, DM and JMD 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 and DM analysed and interpreted the data. ACi, and JRG contributed to the interpretation of the data. SL, ACh, AC and BH drafted the manuscript and all other authors critically reviewed the manuscript for important intellectual content. All authors approved the final version.

## **ACKNOWLEDGEMENT**

Dr Cipriani is also supported by an NIHR Research Professorship (grant RP-2017-08-ST2-006) and by the NIHR Oxford Cognitive Health Clinical Research Facility. We thank Georgia Salanti for her initial help in developing the statistical method, 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 help in data extraction. Ole Andreassen, Guy Chouinard, Michael Jann, Herbert Meltzer, Ofer Agid, Bret Rutherford, Merck, Dainippon Sumitomo, Novartis, Sunovion, Eli Lilly, Johnson&Johnson and Pfizer have replied to our requests for explanations or data. We could also again use the data that authors had sent us for requests in the context of previous meta-analyses of our group.

## SUPPLEMENTARY INFORMATION

Supplementary Information accompanies this paper.

## REFERENCES

- 1 Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B, et al. 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*. 2017;appiajp201716121358.
- 2 049 S. 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.fda.gov>. 2010.
- 3 Leucht S, Chaimani A, Leucht C, Huhn M, Mavridis D, Helfer B, et al. 60 years of placebo-controlled antipsychotic drug trials in acute schizophrenia: Meta-regression of predictors of placebo response. *Schizophr Res*. 2018.
- 4 Mallinckrodt CH, Zhang L, Prucka WR, Millen BA. Signal detection and placebo response in schizophrenia: parallels with depression. *Psychopharmacol Bull*. 2010;43:53-74.
- 5 Agid O, Siu CO, Potkin SG, Kapur S, Watsky E, Vanderburg D, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970-2010. *Am J Psychiatry*. 2013;170(11):1335-44.
- 6 Rutherford BR, Pott E, Tandler JM, Wall MM, Roose SP, Lieberman JA. Placebo response in antipsychotic clinical trials: a meta-analysis. *JAMA psychiatry*. 2014;71(12):1409-21.
- 7 Kirsch I. Are drug and placebo effects in depression additive? *Biol Psychiatry*. 2000;47(8):733-5.
- 8 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100.
- 9 Carpenter WT, Buchanan RW. Schizophrenia. *N Engl J Med*. 1994;330:681-90.

- 10 Leucht S, Tardy M, Komossa K, Heres S, Kissling W, Salanti G, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet*. 2012;379(9831):2063-71.
- 11 Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382(9896):951-62.
- 12 Honigfeld G. Clozapine: antipsychotic activity in treatment-resistant schizophrenics. *AdvTher*. 1984;1:77-97.
- 13 Samara MT, Cao H, Helfer B, Davis JM, Leucht S. Chlorpromazine versus every other antipsychotic for schizophrenia: a systematic review and meta-analysis challenging the dogma of equal efficacy of antipsychotic drugs. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*. 2014;24(7):1046-55.
- 14 Hartling L, Abou-Setta AM, Dursun S, Mousavi SS, Pasichnyk D, Newton AS. Antipsychotics in Adults With Schizophrenia: Comparative Effectiveness of First-Generation Versus Second-Generation Medications: A Systematic Review and Meta-analysis. *Annals of internal medicine*. 2012.
- 15 Dold M, Samara MT, Li C, Tardy M, Leucht S. Haloperidol versus first-generation antipsychotics for the treatment of schizophrenia and other psychotic disorders. *Cochrane Database Syst Rev*. 2015;1:CD009831.
- 16 McMahon RP, Kelly DL, Boggs DL, Li L, Hu Q, Davis JM, et al. Feasibility of reducing the duration of placebo-controlled trials in schizophrenia research. *Schizophr Bull*. 2008;34(2):292-301.
- 17 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. Wiley and Sons: Chichester, UK; 2011.
- 18 Wu TX, Li YP, Liu GJ, Bian Z, Li J, Zhang J, et al. 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. 2006.
- 19 Woodhead M. 80% of China's clinical trial data are fraudulent, investigation finds. *BMJ*. 2016;355:i5396.
- 20 Adams CE, Coutinho E, Davis JM, Duggan L, Essali A, Fenton M, et al. Cochrane Schizophrenia Group. *The Cochrane Library*. Chichester, UK: John Wiley & Sons Ltd; 2011.
- 21 Klein DF, Davis JM. *Diagnosis and drug treatment of psychiatric disorders*. Williams and Wilkins: Baltimore; 1969.
- 22 Adams CE, Awad G, Rathbone J. Chlorpromazine versus placebo for schizophrenia. *Cochrane Database of Systematic Reviews*. 2007(2).
- 23 Joy CB, Adams CE, Laurie S. Haloperidol versus placebo for schizophrenia. *Cochrane Database of Systematic Reviews*. 2007(2).
- 24 Shen X, Xia J, Adams CE. Flupenthixol versus placebo for schizophrenia. *Cochrane Database Syst Rev*. 2012;11:CD009777.
- 25 Matar HE, Almerie MQ. Oral fluphenazine versus placebo for schizophrenia. *Cochrane Database Syst Rev*. 2007(1):CD006352.
- 26 Hartung B, Wada M, Laux G, Leucht S. Perphenazine for schizophrenia. *Cochrane Database of Systematic Reviews*. 2005(1):CD003443.
- 27 Omori IM, Wang J. Sulpiride versus placebo for schizophrenia. *Cochrane Database Syst Rev*. 2009(2):CD007811.

- 28 Fenton M, Rathbone J, Reilly J, Sultana A. Thioridazine for schizophrenia. *Cochrane Database Syst Rev*. 2007(3):CD001944.
- 29 Leucht S, Arbter D, Engel RR, Kissling W, Davis JM. How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *MolPsychiatry*. 2009;14(4):429-47.
- 30 Kay SR, Fiszbein A. The positive and negative symptom scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*. 1987;13(2):261-75.
- 31 Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *PsycholRep*. 1962;10:790-812.
- 32 Leucht S, Rothe P, Davis JM, Engel RR. Equipercetile linking of the BPRS and the PANSS. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*. 2013;23(8):956-9.
- 33 Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *BiolPsychiatry*. 2006;59(11):1001-05.
- 34 Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology*. 2006;59(1):7-10.
- 35 Salanti G, Higgins JP, White IR. Bayesian synthesis of epidemiological evidence with different combinations of exposure groups: application to a gene-gene-environment interaction. *StatMed*. 2006;25(24):4147-63.
- 36 Furukawa TA, Levine SZ, Tanaka S, Goldberg Y, Samara M, Davis JM, et al. Initial severity of schizophrenia and efficacy of antipsychotics: participant-level meta-analysis of 6 placebo-controlled studies. *JAMA psychiatry*. 2015;72(1):14-21.
- 37 Rabinowitz J, Werbeloff N, Caers I, Mandel FS, Stauffer V, Menard F, et al. Determinants of antipsychotic response in schizophrenia: implications for practice and future clinical trials. *J Clin Psychiatry*. 2014;75(4):e308-16.
- 38 Mattila T, Wohlfarth T, Koeter M, Storosum J, van den Brink W, de Haan L, et al. 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*. 2014;24(7):1067-77.
- 39 Rabinowitz J, Levine SZ, Barkai O, Davidov O. 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*. 2009;35(4):775-88.
- 40 Levine SZ, Goldberg Y, Samara M, Davis JM, Leucht S. Joint modeling of dropout and outcome in three pivotal clinical trials of schizophrenia. *Schizophr Res*. 2015;164(1-3):122-6.
- 41 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273:408-12.
- 42 Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*. 2008;336(7644):601-5.
- 43 Porta N, Bonet C, Cobo E. Discordance between reported intention-to-treat and per protocol analyses. *J Clin Epidemiol*. 2007;60(7):663-9.
- 44 Egger M, Davey mG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-34.

- 45 Heres S, Davis J, Maino K, Jetzinger E, Kissling W, Leucht S. 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*. 2006;163(2):185-94.
- 46 Zohar J, Stahl S, Moller HJ, Blier P, Kupfer D, Yamawaki S, et al. A review of the current nomenclature for psychotropic agents and an introduction to the Neuroscience-based Nomenclature. *Eur Neuropsychopharmacol*. 2015;25(12):2318-25.
- 47 Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ. International consensus study of antipsychotic dosing. *Am J Psychiatry*. 2010;167(6):686-93.
- 48 Heres S, Davis J, Maino K, Jetzinger E, Kissling W, Leucht S. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: An exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *AmJPsychiatry*. 2006;163(2):185-94.
- 49 Simmonds MC, Higgin JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. *Stat Med*. 2007;26:2982-99.
- 50 Tzimos A, Samokhvalov V, Kramer M, Ford L, Gassmann-Mayer C, Lim P, et al. in *American Journal of Geriatric Psychiatry* Vol. 16 31-43 (USA, 2008).



**Table 1: Summary of characteristics of included studies**

Explanatory variable	Mean	Standard Deviation	Range	median	IQR	Percentages
Publication year	2000	13.3	(1969,2016)			
Total participants	227	181.1	(12,729)			
Number of sites			(1,112)	25	(1,45)	
Mean dose in chlorpromazine equivalents	584.3	248.9	(200,1740)			
Minimum severity entry criterion						No: 28(26.9%) Yes: 73(70.2%) Unclear: 3(2.9%)
Mean age (years)	38.6	4.8	(25.9,69.8)			
Duration of illness (years)	13.8	4.0	(2.5,32.0)			
Industry sponsorship						Non-sponsored: 31(29.8%) Sponsored/Mixed: 65(62.5%) Unclear: 8(7.7%)
Scale						PANSS: 33(31.7%) BPRS: 68(65.4%) Other: 3(2.9%)
Operationalized criteria						Op: 88 (84.7%) Not op:15 (14.4%) Unclear: 1(1.0%)
Number of medications						2 medications:33(31.7%) >2 medications:71 (68.3%)
Baseline severity (PANSS total)	94.6	5.1	(80,112.5)			
Duration of wash-out period (days)			(0,84)	3	(2,5)	
Study duration (weeks)			(3,26)	6	(6,6)	
Country						USA: 45(43.3%) Other: 59(56.7%)
Total dropout rate	39.3%	17.1%	(0,82.6%)			
Percentage randomized to placebo	28.3%	8.0%	(14.1,50.4)			
Number of arms						2 arms: 11(10.6%) >2 arms: 93(89.4%)
Percentage of academic sites	57.6%	36.3%	(0,100%)			
Dosing schedule						Fixed: 78(75%) Flexible: 26(25%)

Percentage of men	66.3%	20.6%	(0,100%)			
Risk of bias: randomization						Low risk: 48(46.2%) Unclear: 56 (53.9%)
Risk of bias: allocation concealment						Low risk: 33 (31.7%) Unclear: 71(68.3%)
Risk of bias: blinding						Low risk: 57 (54.8%) Unclear: 47 (45.2%)
Risk of bias: missing outcome data						Low risk: 73 (70.2%) Unclear: 19 (18.3%) High risk: 12(11.5%)
Drug Mechanism						M1: 16 (15.4%) Other mechanism: 88(84.6%)

IQR = interquartile range, PANSS = Positive and Negative Syndrome Scale, BPRS = Brief Psychiatric Rating Scale, op. = operationalized, non-op. = not operationalized. M1 – M5 are drug mechanisms of action according to the “Neuroscience-based Nomenclature (NbN)” [46]: M1 = receptor antagonists (D2) clopenthixol, fluphenazine, haloperidol, perphenazine, pimozide, pipotiazine, sulpiride, trifluoperazine. M2 = receptor antagonists (D2, 5-HT2) chlorpromazine, iloperidone, loxapine, lurasidone, olanzapine, sertindole, thioridazine, ziprasidone, zotepine. M3 = receptor partial agonists (D2, 5-HT1A) aripiprazole, brexpiprazole, cariprazine. M4= receptor antagonists (D2, 5-HT2, NE alpha2) asenapine, paliperidone, risperidone. M5= receptor antagonist (D2, 5-HT2) and reuptake inhibitor (NET) quetiapine. A few old drugs have not been classified by NbN yet.

**Table 2: Univariable meta-regressions of potential predictors of drug-response, placebo-response and drug-placebo differences (SMDs)**

	Drug response				Placebo response				SMD				Coefficient corresponds to	N,n (for SMD)
Explanatory variable	Coefficient	95%CI	p-value	% heterogeneity explained	Coefficient	95%CI	p-value	% heterogeneity explained	Coefficient	95%CI	p-value	% heterogeneity explained		
Moderators with a significant effect on drug-placebo differences (SMDs)														
Publication year	0.26	(-0.96,1.48)	0.670	0.4%	2.74	(1.60,3.87)	<0.001	9.5%	-0.10	(-0.13,-0.06)	<0.001	22.2%	10-year increase	105, 22741
Total participants	0.17	(-0.71,1.06)	0.700	0.4%	1.56	(0.72,2.40)	<0.001	5.9%	-0.03	(-0.05,0.00)	0.041	0.0%	100 patients more	105, 22741
Number of sites	-0.09	(-0.81,0.62)	0.797	4.1%	1.13	(0.45,1.81)	0.001	2.7%	-0.03	(-0.05,0.00)	0.020	0.0%	10-site increase	96, 20941
Mean dose	0.56	(-0.11,1.23)	0.103	2.9%	-0.78	(-1.47,-0.08)	0.028	7.6%	0.03	(0.00,0.06)	0.025	5.6%	100 units increase	91, 19957
Min. severity entry criterion	0.93	(-2.83,4.70)	0.624	1.1%	7.47	(3.93,11.01)	<0.001	7.1%	-0.21	(-0.34,-0.09)	0.001	11.1%	With (vs without)	89, 18586
Average age <sup>a</sup>	-10.69	(-14.42,-6.96)	<0.001	15.8%	-4.72	(-9.31,-0.13)	0.044	3.6%	-0.22	(-0.38,-0.07)	0.006	11.1%	10-year increase	100, 22567
Duration of illness <sup>a,b</sup>	-12.14	(-19.11,-5.18)	0.001	12.4%	-9.92	(-18.16,-1.68)	0.019	3.0%	-0.11	(-0.33,0.12)	0.343	22.2%	10-year increase	60, 14278
Industry sponsorship	-0.43	(-3.51,2.65)	0.783	13.1%	1.88	(-1.49,5.24)	0.271	6.7%	-0.18	(-0.28,-0.07)	0.001	16.7%	Sponsored (vs not)	97, 22397
Scale	-1.18	(-4.57,2.22)	0.493	0.3%	5.96	(2.66,9.26)	0.001	5.6%	-0.18	(-0.30,-0.06)	0.003	22.2%	PANSS (vs BPRS)	101, 22589
Operationalized criteria	-0.22	(-5.14,4.69)	0.929	0.9%	-6.18	(-10.97,-1.39)	0.012	2.5%	0.32	(0.14,0.49)	<0.001	11.1%	No (vs yes)	103, 22151
Drug mechanism	-1.75	(-6.19,2.68)	0.434	0.1%	3.01	(-1.57,7.59)	0.195	0.4%	-0.19	(-0.34,-0.05)	0.011	5.6%	M2-M5 (vs M1)	101, 22315
Number of medications	0.57	(-2.86,4.00)	0.742	0.4%	-2.19	(-5.64,1.25)	0.209	0.2%	0.14	(0.04,0.24)	0.008	5.6%	≥two vs one	105, 22741



Moderators with a significant effect on drug-response and/or placebo-response but without a significant impact on drug-placebo differences (SMDs)														
Baseline severity	<b>4.88</b>	<b>(2.08,7.69)</b>	<b>0.001</b>	<b>16.6%</b>	2.52	(-0.59,5.63)	0.111	8.1%	0.10	(0.00,0.21)	0.061	16.7%	10-unit PANSS increase	85, 21259
Duration of wash-out period	<b>-1.53</b>	<b>(-2.48,-0.58)</b>	<b>0.002</b>	<b>0.0%</b>	<b>-1.41</b>	<b>(-2.33,-0.49)</b>	<b>0.003</b>	<b>4.0%</b>	0.02	(-0.02,0.06)	0.328	5.6%	10-day increase	89, 18586
Study duration	<b>-8.12</b>	<b>(-13.01,-3.23)</b>	<b>0.001</b>	<b>14.9%</b>	<b>-8.80</b>	<b>(-13.97,-3.63)</b>	<b>0.001</b>	<b>10.6%</b>	0.08	(-0.12,0.27)	0.432	16.7%	10-week increase	96, 22443
Country	2.75	(-0.40,5.90)	0.087	1.0%	<b>3.64</b>	<b>(0.48,6.81)</b>	<b>0.024</b>	<b>2.1%</b>	0.00	(-0.10,0.10)	0.986	0.0%	Non-USA or mixed (vs only USA)	105, 22741
Moderators without a significant effect on neither drug-response, nor placebo-response, nor drug-placebo differences (SMDs)														
Total dropout rate	-0.77	(-1.75,0.20)	0.119	5.8%	-0.78	(-1.81,0.25)	0.134	1.5%	-0.02	(-0.05,0.02)	0.330	5.6%	10% increase	89, 21514
Percentage randomized to placebo	0.14	(-1.91,2.20)	0.890	0.5%	-0.99	(-3.06,1.08)	0.345	0.1%	0.01	(-0.05,0.08)	0.721	0.0%	10% increase	105, 22741
Number of arms	-0.31	(-5.60,4.99)	0.909	0.5%	2.21	(-3.12,7.54)	0.412	0.2%	0.00	(-0.19,0.19)	0.996	0.0%	Three or more (vs two)	105, 22741
Percentage of academic sites	-0.32	(-1.09,0.46)	0.418	25.1%	-0.65	(-1.35,0.06)	0.071	12.3%	0.00	(-0.02,0.03)	0.879	5.6%	10% increase	59, 9379
Dosing schedule	3.29	(-0.27,6.85)	0.070	1.3%	-0.82	(-4.52,2.89)	0.663	0.5%	0.10	(-0.03,0.23)	0.140	5.6%	Flexible (vs fixed)	105, 22741
Percentage of men	0.52	(-0.33,1.37)	0.225	5.2%	0.20	(-0.67,1.07)	0.649	3.4%	0.00	(-0.04,0.03)	0.888	11.1%	10% increase	91, 21119
Risk of bias: randomization	1.26	(-1.92,4.44)	0.433	0.1%	-0.13	(-3.36,3.09)	0.934	0.5%	-0.02	(-0.12,0.07)	0.626	0.0%	Unclear risk (vs low)	105, 22741
Risk of bias: allocation concealment	1.29	(-2.16,4.75)	0.460	0.1%	-1.11	(-4.57,2.36)	0.528	0.4%	0.03	(-0.07,0.13)	0.573	0.0%	Unclear risk (vs low)	105, 22741
Risk of bias: blinding	1.78	(-1.39,4.95)	0.268	0.1%	1.47	(-1.75,4.68)	0.367	0.1%	-0.01	(-0.11,0.09)	0.850	0.0%	Unclear or high risk (vs low)	105, 22741
Risk of bias: missing outcome data	0.14	(-3.39,3.68)	0.936	0.5%	-2.70	(-6.27,0.88)	0.138	0.6%	0.10	(-0.01,0.21)	0.084	5.6%	Unclear or high risk (vs low)	105, 22741

Statistically significant findings are presented by bold print. N = number of studies, n= number of patients, PANSS = Positive and Negative Syndrome Scale, BPRS = Brief Psychiatric Rating Scale. <sup>a</sup>These analyses were performed after the exclusion of an outlier study [50]. 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. <sup>b</sup>The effect of duration of illness on SMD was not significant, but because it is another chronicity measure related to mean age, and because the trend was the same we put the results in this section. Coefficient = coefficient of the meta-regression. 95% CI = 95% confidence interval.

“Coefficient corresponds to” = for example publication year: a 10-year increase in publication year on the average increases the drug response by 0.26 PANSS units, the placebo response by 2.74 PANSS units, and the SMD by -0.10 units.

M1 – M5 are drug mechanisms of action according to the “Neuroscience-based Nomenclature (NbN)” [46]: M1 = receptor antagonists (D2) clopenthixol, fluphenazine, haloperidol, perphenazine, pimozide, pipotiazine, sulpiride, trifluoperazine. M2 = receptor antagonists (D2, 5-HT<sub>2</sub>) chlorpromazine, iloperidone, loxapine, lurasidone, olanzapine, sertindole, thioridazine, ziprasidone, zotepine. M3 = receptor partial agonists (D2, 5-HT<sub>1A</sub>) aripiprazole, brexpiprazole, cariprazine. M4= receptor antagonists (D2, 5-HT<sub>2</sub>, NE alpha<sub>2</sub>) asenapine, paliperidone, risperidone. M5= receptor antagonist (D2, 5-HT<sub>2</sub>) and reuptake inhibitor (NET) quetiapine. A few old drugs have not been classified by NbN yet.

#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.

**Table 3: Multivariable meta-regression models for drug response, placebo response and drug-placebo differences (SMD) (backward stepwise algorithm)**

Drug response					
Moderator	Moderator	Moderator	p-value	Moderator	Interpretation
Average age <sup>a</sup>	-5.41	-9.73,-1.10	0.015	10 years increase	For every 10-year increase in study average age, drug response decreases on average by 5.41 units
Baseline severity <sup>a</sup>	3.77	0.93,6.62	0.010	10 units increase	For every 10-point increase in baseline PANSS score, drug response increases on average by 3.77 units
Summary of the model in Table 3a: 82 studies with 21093 participants, heterogeneity standard deviation 6.21, the model explained 22% of the heterogeneity					
Placebo response					
Explanatory variable	Coefficient	95% CrI	p-value	Coefficient corresponds to	Interpretation
Total participants <sup>a</sup>	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 <sup>a</sup>	-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
<u>Summary of the model:</u> 95 studies with 22300 participants, heterogeneity standard deviation 7.02, the model explained 12% of the heterogeneity					
<b>Drug-Placebo differences (Standardised Mean difference, SMD)</b>					
<b>Moderator</b>	<b>Coefficient</b>	<b>95% CrI</b>	<b>p-value</b>	<b>Coefficient corresponds to</b>	<b>Interpretation</b>
<b>Placebo response<sup>a</sup></b>	-0.15	-0.21,-0.09	<0.001	10 units increase	For every 10-point increase in mean PANSS change score in placebo arm, the SMD decreases on average by 0.15 units
<b>Industry sponsorship<sup>a</sup></b>	-0.13	-0.22,-0.04	0.006	10 years increase	For studies including at least one sponsored drug in comparison with non-sponsored studies, the average SMD is smaller by 0.13 units
<u>Summary of the model in Table 1:</u> 92 studies with 22236 participants, heterogeneity standard deviation 0.10, the model explained 42% of the heterogeneity					

<sup>a</sup>Statistically significant moderators.

## Legends of Figures

### Figure 1: Meta-regression placebo-response versus drug-response

### Figure 2: Moderators of placebo-response – univariable meta-regressions

The figures in this panel correspond to the following moderators: a) Publication year, b) Number of participants, c) Number of sites, d) Mean dose in chlorpromazine equivalents, e) Minimum scale-derived severity threshold as inclusion criterion, f) Average participant age in years, g) Average duration of illness in years, h) Industry sponsorship, i) Scale, j) Operationalized criteria or not, k) Drug mechanism, l) Number of medications, m) Baseline severity (PANSS total score at baseline) n) Duration of wash-out phases, o) Study duration, p) Country, q) Total dropout rate, r) percentage randomized to placebo, s) number of arms risk, t) Percentage of academic sites, u) Dosing schedule, v) Percentage of men, w) Risk

of bias concerning randomization method, x) risk of bias concerning allocation concealment, y) risk of bias concerning blinding, z) risk of bias concerning missing outcomes.

M1 – M5 are drug mechanisms of action according to the “Neuroscience-based Nomenclature (NbN)”

[46]: M1 = receptor antagonists (D2) clopenthixol, fluphenazine, haloperidol, perphenazine, pimozide, pipotiazine, sulpiride, trifluoperazine. M2 = receptor antagonists (D2, 5-HT2) chlorpromazine,

iloperidone, loxapine, lurasidone, olanzapine, sertindole, thioridazine, ziprasidone, zotepine. M3 =

receptor partial agonists (D2, 5-HT1A) aripiprazole, brexpiprazole, cariprazine. M4= receptor antagonists (D2, 5-HT2, NE alpha2) asenapine, paliperidone, risperidone. M5= receptor antagonist (D2, 5-HT2) and reuptake inhibitor (NET) quetiapine. A few old drugs have not been classified by NbN yet.

The B values and their 95% confidence intervals at the bottom of the graphs are the coefficients for placebo response ( $B_{pla}$ ), drug response ( $B_{drug}$ ) and drug-placebo differences ( $B_{SMD}$ ). The numbers in square brackets in the figures describe to what the coefficient refers to. For example, in Figure 2a publication year: “ [10 years increase],  $B_{pla}=2.74$  (1.60,3.88),  $B_{drug}$  0.26 (-0.96,1.48),  $B_{SMD}$  (-0.10 (-0.13,-0.06) ” 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, 0.26 (-0.96,1.48) higher drug response and -0.10 (-0.13,-0.06) lower drug-placebo difference. Or in Figure 2g), as an example for a dichotomous moderator: [With (vs not)],  $B_{pla}=7.47$  (3.93,11.01),  $B_{drug}=0.93$  (-2.83,4.70),  $B_{SMD}=-0.21$  (-0.34,-0.09) ” means a study which had a minimum baseline severity as an inclusion criterion had on average 7.47 (95% confidence interval 3.93,11.01) PANSS points higher placebo response, -0.21 (-0.34,-0.09) lower drug response and -0.21 (-0.34,-0.09) lower drug-placebo difference compared to a study without such a criterion.

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

<sup>1</sup>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, <sup>2</sup>this meta-regression was also statistically significant when the only outlier study of a duration of 26 weeks was excluded