OPEN ACCESS

University of BRISTOL

University of Bristol - Explore Bristol Research
General rights

Mental Health

Impact of placebo arms on outcomes in antidepressant trials: systematic review and meta-regression analysis

Georgia Salanti,[1] Anna Chaimani,[2,3,4] Toshi A Furukawa,[5] Julian PT Higgins,[6] Yusuke Ogawa,[5] Andrea Cipriani[7,8]* and Matthias Egger[1]

[1]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland, [2]School of Medicine, Paris Descartes University, Paris, France, [3]INSERM, UMR1153 Epidemiology and Statistics, Sorbonne Paris Cité Research Center (CRESS), Paris, France, [4]French Cochrane Center, Hôpital Hôtel-Dieu, 1 place du Parvis Notre-Dame, 75004 Paris, France, [5]Departments of Health Promotion and Human Behavior and of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan, [6]School of Social and Community Medicine, University of Bristol, Bristol, UK, [7]Department of Psychiatry, University of Oxford, Oxford, UK and [8]Oxford Health NHS Foundation Trust, Warneford Hospital, Warneford Lane, Oxford, UK

*Corresponding author. Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX, UK.
E-mail: andrea.cipriani@psych.ox.ac.uk

Abstract

Background: There is debate in the literature as to whether inclusion of a placebo arm may alter characteristics of antidepressant trials. However, previous research has focused on response rates of various antidepressants on average only, ignoring potential differences among drugs or other aspects of trial findings. Little is known about the impact of a placebo arm on all-cause dropout and dropout due to adverse events.

Methods: We carried out a systematic review of published and unpublished double-blind randomized controlled trials (RCTs) for the acute treatment of unipolar major depression (update: January 2016). The probability of being allocated to placebo (π) was the exposure of interest, and we examined its influence on responders (efficacy), all-cause dropouts (acceptability) and dropouts due to adverse events (tolerability), while accounting for differences in drugs, trials and patient characteristics in multivariate random effects meta-regression.

Results: We included 421 studies (68 305 participants) comparing 16 antidepressants or placebo; π ranged from 20% to 50%. Response rate was lower [risk ratio (RR) 0.87; 95% confidence interval (CI) 0.83, 0.92] and all-cause dropout rate higher (RR 1.19; 95% CI 1.08, 1.31) for the same antidepressants in placebo-controlled trials compared with head-to-head trials. The probability of responding decreased by 3% (95% CI 2–5%) for every 10% increase in π, whereas the risk of all-cause dropout increased by 4% (95% CI 1–7%). Tolerability was unaffected by π. Response rate was inversely correlated with dropouts due to any cause (correlation coefficient −0.48; 95% CI −0.58, −0.36) and due to adverse events (−0.34; 95% CI −0.44, −0.23).

Conclusions: For the same antidepressant, response rate was on average smaller and dropouts higher when placebo was included; however, no association was found with dropouts due to adverse events. Decreased patient expectations, larger dropout rates and use of inappropriate statistical methods to impute missing data may explain this phenomenon. The findings call for caution in the integration of randomized evidence involving placebo arms.

Key words: Systematic review, meta-analysis, meta-regression, antidepressants, placebo, randomized controlled trial

---

Key Messages

- Previous reviews found that in antidepressant trials, response to active intervention gradually decreased among head-to-head studies, multi-arm placebo-controlled studies and two-arm placebo-controlled studies, in this order. However, these studies did not adjust their results for differences in trial or patient characteristics and, more importantly, they did not account for differences between individual drugs and investigated the impact of the inclusion of placebo on response rate only.
- By synthesizing 706 active treatment arms, using appropriate multivariate meta-regression techniques, our study is the largest to date to provide evidence on the impact of inclusion of placebo arm on the response to 16 different antidepressants and also on the likelihood of dropout due to any cause and due to adverse events.
- This study found that therapeutic response to the same antidepressant arm was on average smaller and dropouts more likely when the probability of receiving placebo increased. By contrast, there was no influence on dropout rate due to adverse events.
- For the same drug and the same probability of receiving placebo, larger dropout rates were associated with lower response rates to the treatment.
- The probability of receiving placebo in the clinical trial alters the characteristics of the trial by inducing different response and dropout rates with the same antidepressants among its participants.
- Decreased patient expectations in placebo-controlled trials and the widespread use of the 'last observation carried forward' approach to record missing outcome data might explain this phenomenon. The probability of receiving placebo should be considered when interpreting and synthesizing results from randomized controlled trials in major depression.

Background

The expectations and preferences patients may have regarding a treatment can influence their response to that treatment. In open trials, patients who are allocated to the non-preferred treatment may experience 'resentful demoralization' and consequently show lower adherence to the assigned treatment.[1] In major depression, strong expectations of improvement were associated with both a higher probability of complete response and reduced severity of depression at the end of a multi-arm trial of psychotherapy and pharmacotherapy.[2] A trial including a preference arm showed that in mild-to-moderate depression, taking into account patients' preferences for pharmacotherapy or psychotherapy was associated with additional benefit.[3] Although many instruments have been developed to assess patient expectations, measurement is complex and the validity and reliability of different approaches are unclear.[4]

The role of placebo in randomized controlled trials (RCTs) has been long debated from both methodological and ethical perspectives.[5] Although a standard requirement for licensing approval in regulatory settings, the use of placebo in phase III studies has been challenged.[6] In the field of depression it has been suggested that placebo arms are needed because equivalence between a new drug and standard treatment is not evidence of efficacy unless the new drug is also more effective than placebo.[7] However, the high rate of placebo responders in antidepressant trials has added a layer of complexity to the process of designing trials and interpreting results.[8] Trial participants are informed that they will receive either one of several active treatments, or placebo, with the probability of placebo ranging from 0% in a head-to-head trial to 50% in a two-arm placebo-controlled trial. Several studies have shown that administration of a placebo that simulates an active

treatment can mimic the effects of the pharmacological intervention, depending on contextual factors and mediated through psychological and neurobiological mechanisms ('the powerful placebo').[9] Conversely, in blinded clinical trials the knowledge that an inactive treatment might be received could reduce the response to an active treatment, and abolish the placebo response. Patient expectations could also be responsible for high rates of adverse events, even in patients who do not receive an active treatment ('the nocebo phenomenon').[10]

Many factors can be associated with placebo response in antidepressant trials, such as baseline severity,[11] dosing schedule[12] and length of trial.[13] However, the inclusion of a placebo arm is an important issue not only from a clinical viewpoint, but also from a methodological one, because it can introduce heterogeneity between trials and violate the assumptions underlying meta-analyses.[14] Two previous studies investigated this issue. Papakostas and Fava studied 182 placebo-controlled trials and found that a higher probability of receiving placebo reduced the response to the active intervention.[15] Sinyor and colleagues[16] synthesized 90 head-to-head and placebo-controlled trials and found that response to active intervention gradually decreased between head-to-head studies, multi-arm placebo-controlled studies and two-arm placebo-controlled studies. However, these studies did not include a large sample of trials, they focused only on efficacy, did not adjust results for differences in trial or patient characteristics[16] and did not account for the different drugs used in the trials.[15,16] Additionally, as publication bias is a well-known threat in antidepressant trials, it is not known how this might have affected the conclusions of previous studies. Therefore to properly address this question, we conducted a systematic review and meta-regression analysis including unpublished data, and examined not only whether the probability of receiving placebo modifies response to treatment but also whether it affects acceptability and tolerability in antidepressant trials.

Methods

This study is based on a systematic review and network meta-analysis of the comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression.[17] The protocol was registered in PROSPERO (CRD42012002291) and published.[18] The results are reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines. Eligible trials were identified from seven electronic databases (CENTRAL, CINAHL, EMBASE, LiLACS, MEDLINE, MEDLINE In-Process and PSYCINFO) from inception to 8 January 2016. The reference lists of relevant papers were also

scrutinized. Files of the national drug licensing agencies in six countries (USA, UK, The Netherlands, Sweden, Japan and Australia), the European Medicines Agency and several trial registries were searched for published, unpublished and ongoing RCTs (for further details on the search strategy, see reference[18]).

Inclusion and exclusion criteria

The following criteria had to be satisfied for studies to be eligible for the meta-analysis and meta-regression analysis: (i) randomized controlled trials (RCTs) reported as double-blind, comparing one active drug with another or with placebo in the acute phase treatment of major depression (studies where sequence generation was not clearly random, or where the allocation was clearly not concealed, were excluded); (ii) patients aged 18 years or older, of both sexes, with a primary diagnosis of non-psychotic, unipolar major depression, according to any standard diagnostic criteria (DSM, Feighner or ICD-10 criteria), with no serious concomitant medical illness; and (iii) any second-generation antidepressant (agomelatine, bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, nefazodone, paroxetine, reboxetine, sertraline, trazodone, venlafaxine, vilazodone and vortioxetine) and two first-generation drugs included in the WHO list of essential medicines (amitriptyline, clomipramine). Full details about inclusion and exclusion criteria are reported in the published protocol, with full details about the rationale for selecting the investigational drugs.[18]

Data extraction and risk of bias

Two persons independently reviewed the titles and abstracts retrieved by the search. The full text of potentially eligible articles was obtained, and any disagreements about eligibility were resolved via discussion with a third reviewer. The structured data extraction sheet covered the study and participant characteristics, intervention details and outcome measures. Severity scores at baseline were transformed to Hamilton Depression Rating Scale (17-item) scores when necessary (Supplementary Appendix 1, available as Supplementary data at *IJE* online). The results for response reported in the original publication were compared with those in any unpublished study reports, and a decision tree was used for data extraction (Supplementary Appendix 1, available as Supplementary data at *IJE* online).[18] Where necessary, authors were contacted in order to obtain further information. Two independent raters assessed generation of allocation sequence, allocation concealment, blinding of study personnel and participants, blinding of outcome assessor and other domains, including sponsorship

using the Cochrane risk of bias tool. Inter-rater agreement on risk of bias judgements was assessed in a sample of the eligible articles. Analyses of completers only or of last observations carried forward (LOCF) were considered at high risk of attrition bias. Studies supported by industry were considered at high risk of sponsorship bias, and arms of drugs manufactured by the sponsor of the trial were considered to be at higher risk of bias than comparator arms (see Supplementary Appendix 1, available as Supplementary data at *IJE* online for more details).

Outcomes and exposure

The outcome for efficacy was response rate, defined as the percentage of patients who had a reduction of at least 50% on the total score between baseline and week 8 on a standardized observer-rating scale for depression (e.g. Hamilton Depression Rating Scale or Montgomery-Åsberg Depression Rating Scale). When only changes in scores were reported, we imputed responder rates using a standardized method.[19] Dropout due to any cause and dropout due to adverse events were also recorded.

The exposure of interest was the probability of being allocated to placebo, denoted as π and estimated using equal allocation ratios. We analysed π as a dichotomous variable as $\pi = 0\%$ or $\pi > 0\%$ (arms from head-to-head versus placebo-controlled trials), as a trichotomous variable ($\pi = 0\%$, $0\% > \pi < 50\%$, $\pi = 50\%$ corresponding to arms from head-to-head trials, placebo-controlled trials with more than two active arms and two-arm placebo-controlled trials) and as a continuous variable (the number of placebo arms over the number of all study arms).

Statistical analysis

The preliminary analysis aimed to identify trial and patient characteristics that differed between active arms with $\pi = 0\%$ and $\pi > 0\%$ (Supplementary Appendix 2, available as Supplementary data at *IJE* online). We used random-effects meta-regression with each characteristic as the dependent variable and dichotomous π as covariate.[20] If a precision measure of the characteristic was not available or relevant, we employed a conventional regression analysis. The characteristics associated with π in this analysis were considered potential confounders of the association between π and study outcomes. In the main analysis, we fit multivariable multivariate random-effects meta-regression models for the log-transformed response and dropout with $\pi = 0\%$, $0\% > \pi > 50\%$ and $\pi = 50\%$ and by entering π as a continuous exposure variable (Supplementary Appendix 2, available as Supplementary data at *IJE* online).[21] We controlled for possible confounding by drug by including it as

a covariate. We entered other potential confounders into the model to evaluate their independent impact on response and dropout. We fit the final models for each outcome by including the exposure variable π, the drugs and the trial or patient characteristics that were independent predictors of response and dropout.

In a sensitivity analysis we excluded arms of amitriptyline and trazodone, which are difficult to blind due to side effects. We also re-ran the main analysis models using the logit-transformed responses and dropout. To explore the association between dropout and response rates, we estimated overall correlation coefficients using a random-effects model.[22] We also ran a multivariate meta-analysis of log-transformed response rates using the dropout rate as covariate (on top of placebo and active drugs). Heterogeneity was measured using the random-effects standard deviation τ. Statistical analyses were performed in Stata version 13.1 and R version 3.0.2.[23,24] For full details about the statistical models, see Supplementary Appendix 2, available as Supplementary data at *IJE* online.

Results

A total of 28 541 citations were reviewed (Supplementary Appendix 5, available as Supplementary data at *IJE* online). We included 421 studies (68 305 participants) 73 of which are unpublished (Supplementary Appendix 1, available as Supplementary data at *IJE* online): in total, 706 arms of active drugs that were examined both in 169 head-to-head trials ($n = 29\ 841$) and in 252 placebo-controlled studies ($n = 38\ 464$) (see Supplementary Appendix 4, available as Supplementary data at *IJE* online for reference list of included studies). The 706 arms studied 16 different antidepressants. About half of the active drug arms (340) belonged to 169 studies with 0% probability of receiving placebo (π). The placebo-controlled studies had two to five arms, so that π ranged between 20% and 50%. Supplementary Appendix 6 (available as Supplementary data at *IJE* online) shows the distribution of the arms in the various types of studies with and without placebo. Studies with $\pi = 0\%$ were comparable to those with a 20% or higher probability of placebo in terms of year of publication, use of rescue medication, risk of bias and frequency of reporting response and all-cause dropout rates (Table 1). Most head-to-head studies included only two arms (158 studies, 93.5%) whereas many placebo-controlled trials had three arms or more (197, 78.2%). The response rate was available in 386 studies and 650 arms (92.1%). The dropout rate was reported in 378 studies and 644 arms (91.2%), and dropout for adverse events in 354 studies and 596 arms (84.4%). Placebo-controlled trials were more likely to report dropouts for adverse events.

Table 1. Characteristics of the included studies. *P*-values are obtained from random-effect meta-regressions or simple regressions as described in the Supplementary Appendix, available as Supplementary data at *IJE* online (see Supplementary Appendix 1, available as Supplementary data at *IJE* online, for the definition of the covariates)

Trials	Studies with $\pi = 0\%$ (head-to-head trials) ($n = 169$, 100%)[a]	Studies with π between 20% and 50% (placebo-controlled trials)($n = 252$, 100%)[a]	*P*-value
Median year of study completion [range]	1999 [1980, 2015]	1999.5 [1978, 2014]	0.71
Use of rescue medication	66 (39.1%)	76 (30.2%)	0.06
Use of placebo run-in phase	71 (42.0%)	148 (58.7%)	0.001
No. of arms			
2	158 (93.5%)	55 (21.8%)	<0.001
3	9 (5.3%)	142 (56.3%)	
>3	2 (1.2%)	55 (21.8%)	
Risk of bias			
Generation of random sequence			0.41
Low risk of bias	30 (17.8%)	53 (21.0%)	
Unclear risk of bias	139 (82.2%)	199 (79.0%)	
Concealment of allocation			0.18
Low risk of bias	19 (11.2%)	40 (15.9%)	
Unclear risk of bias	150 (88.8%)	212 (84.1%)	
Blinding of assessors			0.04
Stated but not tested	34 (20.1%)	32 (12.7%)	
Unclear risk of bias	135 (79.9%)	220 (87.3%)	
Response rates available	154 (91.1%)	232 (92.1%)	0.73
Unpublished report available and presents adequate response data	50 (29.6%)	121 (48.0%)	<0.001
Response rate presented in published report	128 (83.1%)	163 (70.3%)	0.005
Published and unpublished reports on response are in agreement	139 (90.3%)	197 (84.9%)	0.126
All-cause dropout reported	153 (90.5%)	225 (89.3%)	0.68
Response and all-cause dropout reported	141 (83.4%)	213 (84.5%)	0.76
Dropout for adverse events reported	134 (79.3%)	220 (87.3%)	0.03
Funding			<0.001
High risk of sponsorship bias (Industry funding or unclear)	159 (94.0%)	205 (81.3%)	
Low risk of sponsorship bias (not-for profit funding or no funding)	10 (6.0%)	47 (18.7%)	

π: probability of receiving placebo.

[a]Percentages are calculated out of the total number of trials reporting on response (154 and 232, respectively).

The risk of bias was unclear in most placebo-controlled and head-to-head trials. Outcome assessors were reported to be blinded more often in head-to-head studies than in placebo-controlled trials (Table 1). In a sample of 155 articles, inter-rater agreement on risk of bias judgements ranged from 89% to 98.1% (Supplementary Appendix 1, available as Supplementary data at *IJE* online).

The arms from head-to-head trials and from placebo-controlled trials were similar in terms of mean age, baseline severity score of participants, use of LOCF approach and sample size (Table 2). The mean percentage of female participants was higher in arms with zero probability of receiving placebo. Most trial arms were at high risk of

attrition bias, both in placebo-controlled and in head-to-head trials. Appropriate imputation of missing data was more frequent in arms from placebo-controlled trials than arms from head-to-head studies. Head-to-head studies were more frequently sponsored by industry. However, active arms were more likely to be associated with the sponsor of the trial in a placebo-controlled trial than in a head-to-head study (Table 2).

For all drugs except amitriptyline, the mean response rate was higher in the head-to-head trials compared with the placebo-controlled trials (Figure 1a). For most drugs there was a negative association with the probability of receiving placebo. The lowest mean response was

1459

Table 2. Characteristics of the included arms. P-values are obtained from random-effect meta-regressions or simple regressions as described in the Supplementary Appendix, available as Supplementary data at *IJE* online (see Supplementary Appendix 1, available as Supplementary data at *IJE* online for the definition of the covariates).

Active arms	Arms from studies with $\pi = 0\%$ ($n = 340$, 100%)	Arms from studies with π between 20% and 50% ($n = 366$, 100%)	P-value
Mean age (from 496 arms)	44.5	43.1	0.34
Mean percentage of females (from 332 arms)	64.8	61.2	0.007
Mean baseline depression score (from 637 arms)	24.4	24.3	0.24
Mean sample size	88	104	0.27
Attrition bias			
Low risk of bias (appropriate imputations or dropout <1%)	7 (2.1%)	43 (11.7%)	<0.001
High risk of bias (incomplete data analysis)	77 (22.6%)	59 (16.1%)	
High risk of bias (LOCF approach)	227 (66.8%)	239 (65.3%)	
Unclear risk of bias	29 (8.5%)	25 (6.8%)	
High risk of sponsorship bias (the industry funding the trial manufactures the drug examined in the arm)	186 (54.7%)	270 (73.8%)	0.001
Active intervention in the arm**			
Agomelatine	9 (2.6%)	17 (4.6%)	0.94
Amitriptyline	46 (13.5%)	26 (7.1%)	<0.001
Bupropion	7 (2.1%)	18 (4.9%)	0.44
Citalopram	20 (5.9%)	16 (4.4%)	0.01
Duloxetine	7 (2.1%)	28 (7.7%)	0.05
Escitalopram	23 (6.8%)	22 (6.0%)	0.02
Fluoxetine	63 (18.5%)	44 (12.0%)	0.08
Fluvoxamine	13 (3.8%)	12 (3.3%)	<0.001
Mirtazapine	18 (5.3%)	17 (4.6%)	0.04
Nefazodone	7 (2.1%)	8 (2.2%)	0.42
Paroxetine	51 (15.0%)	56 (15.3%)	0.005
Reboxetine	5 (1.5%)	10 (2.7%)	0.87
Sertaline	24 (7.1%)	25 (6.8%)	0.04
Trazodone	13 (3.8%)	9 (2.5%)	0.02
Venlafaxine	33 (9.7%)	33 (9.0%)	0.01
Vortioxetine	1 (0.3%)	25 (6.8%)	0.001

π, probability of receiving placebo; LOCF, last observation carried forward.

**P-values are obtained from X^2 tests.

observed in two-arm placebo-controlled trials ($\pi = 50\%$), a higher response in multi-arm placebo-controlled trials ($20\% \leq \pi < 50\%$) and the highest response in head-to-head studies ($\pi = 0\%$). Conversely, all-cause dropout rates were lower in head-to-head studies (Figure 1b). For several drugs (citalopram, escitalopram, mirtazapine, nefazodone and paroxetine) there was a positive association between π and dropout from all causes. By contrast, for dropout due to adverse events no clear pattern emerged (Suppplementary Appendix 7, available as Supplementary data at *IJE* online).

The probability of response to an active drug decreased by 13% (RR 0.87, 95% CI 0.83 to 0.92) if the drug was tested in a placebo-controlled trial rather than in a head-to-head trial (Table 3). There was also a monotonic association between response and the number of arms in placebo-controlled trials. The probability of response decreased by 16% when $\pi = 50\%$ (RR 0.84, 95% CI 0.77, 0.92) and 12% (RR 0.88, 95% CI 0.83, 0.93) when $20\% \leq \pi < 50\%$ compared with $\pi = 0\%$. It decreased by 3% (RR 0.97, 95% CI 0.95, 0.98) for every 10% increase in the probability of being allocated to placebo.

The risk of dropout increased by 19% (RR 1.19, 95% CI 1.08, 1.31) when the arm was part of a placebo-controlled trial rather than a head-to-head study (Table 3). This risk increased by 4% (RR 1.04, 95% CI 1.01, 1.07) for every 10% increase in the probability of being allocated to placebo. No important differences were found for dropout for adverse events. The results are presented graphically in Supplementary Appendix 8, available as Supplementary data at *IJE* online.
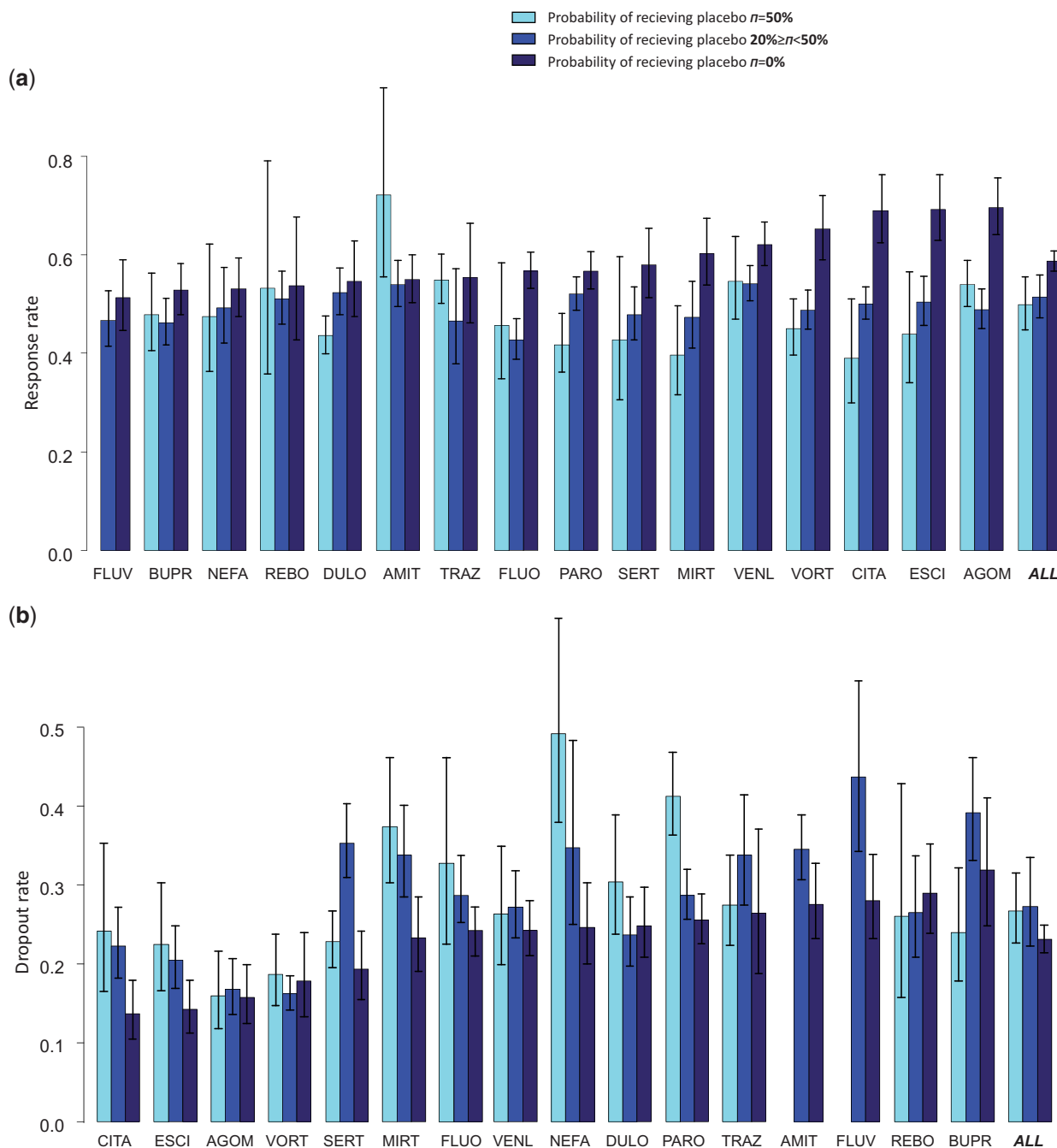
Figure 1. Response to treatment (1a) and dropout rate for any reason (1b) as estimated in active arms grouped by probability of receiving placebo. Drugs are ordered by response and dropout rates estimated in the head-to-head trials ($\pi = 0\%$). The bars and confidence intervals for ALL (all drugs) are estimated from the multivariate model after adjusting for differences between active drugs. AGOM, agomelatine; AMIT, amitriptyline; BUPR, bupropion; CITA, citalopram; DULO, duloxetine; ESCI, escitalopram; FLUO, fluoxetine; FLUV, fluvoxamine; LEVO, levomilnacipran; MILN, milnacipran; MIRT, mirtazapine; NEFA, nefazodone; PARO, paroxetine; REBO, reboxetine; SERT, sertraline; TRAZ, trazodone; VENL, venlafaxine; VORT, vortioxetine. The raw percentages are shown in Table 1 in Supplementary Appendix 3, available as Supplementary data at *IJE* online.

After excluding amitriptyline and trazodone in the sensitivity analysis, the differences between arms from placebo-controlled and head-to-head trials became slightly more pronounced (Table 2 in Supplementary Appendix 3, available as Supplementary data at *IJE* online). Re-analysis of the data using the logit-transformed response and dropout rates led to similar results (Table 3 in Supplementary Appendix 3, available as Supplementary data at *IJE* online). Response rate was inversely correlated with dropout rate: the summary correlation coefficient was −0.48 (95% CI −0.58, −0.36) (Supplementary Appendix 9, available as Supplementary data at *IJE* online). The correlation between the two outcomes was not affected by the probability of receiving placebo.

Table 3. Results from multivariate meta-regression for the impact of placebo-controlled trials in response and dropout with active antidepressant drugs. Results are adjusted for the type of antidepressant. Heterogeneity standard deviations did not differ materially across models (less than 1% change) and were $\tau = 0.20$ (response), $\tau = 0.36$ (all-cause dropout), $\tau = 0.43$ (dropout due to adverse events)

	Response to active treatment	All-cause dropout	Dropout due to adverse events
Number of active arms (studies)	647 (386)	641 (378)	580 (350)
Risk ratios (95% CI): $\pi = 0\%$	1	1	1
$\pi > 0\%$	0.87 (0.83, 0.92)	1.19 (1.08, 1.31)	1.07 (0.94, 1.23)
$\pi = 0\%$	1	1	1
$20\% \leq \pi < 50\%$	0.88 (0.83, 0.93)	1.19 (1.08, 1.31)	0.98 (0.78, 1.22)
$\pi = 50\%$	0.84 (0.77, 0.92)	1.15 (0.99, 1.34)	1.08 (0.94, 1.25)
for π increase by 10%	0.97 (0.95, 0.98)	1.04 (1.01, 1.07)	1.06 (0.73, 1.56)

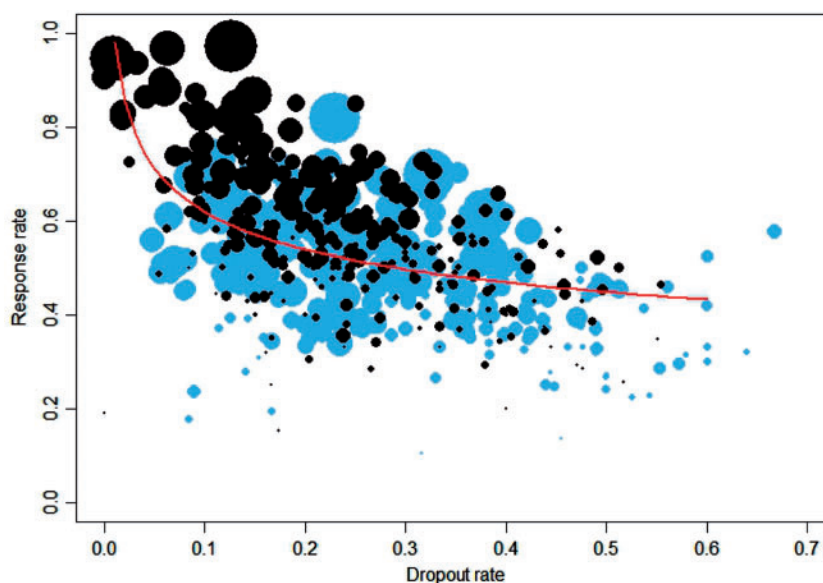π, probability of being allocated to placebo arm.

Figure 2. Response rate in active arms versus dropout rate due to any cause. Data in grey are from arms with probability of receiving placebo $\pi > 0\%$ and data in black are from arms with probability of receiving placebo $\pi = 0\%$. The line corresponds to the multivariate regression line (exponentiated).

We added the variables that had different distributions in placebo-controlled and head-to-head studies (sponsorship, blinding and attrition bias, availability of the unpublished report and the use of a placebo run-in phase) in a multivariate model. The percentage of female participants was not associated with response or all-cause dropout (for a 10% increase in the percentage of female participants, we obtained RR for response: 1.01 95% CI (0.98, 1.04) and RR for dropout 0.99 (95% CI 0.94, 1.04) and was omitted from the multivariate analysis in order to increase the sample size (374 arms had missing values). Only the availability of an unpublished report and the use of placebo run-in phase had an independent effect on response or dropout (Table 4 in Supplementary Appendix 3, available as Supplementary data at *IJE* online). The probability

of receiving placebo remained the most important predictor of response and dropout after including the two variables in the model (Table 5 in Supplemementary Appendix 3, available as Supplementary data at *IJE* online).

Studies with higher dropout rates showed lower average response rates (Figure 2). In the multivariate random-effects meta-analysis adjusted for drug differences and the inclusion of placebo, the coefficient between log-transformed response rate and dropout rate for any cause was -0.17 (95% CI -0.22, -0.13). This means that, for the same drug, an arm with a 10% dropout rate will have a 31% greater response rate compared with an arm with a 50% dropout rate. The correlation between log-transformed response rate and dropout rate due to adverse events was a bit smaller than that for dropout for any cause (-0.34, 95% CI

−0.44, −0.23) (Supplementary Appendix 10, available as Supplementary data at *IJE* online).

Discussion

In this large meta-epidemiological study, we found that the therapeutic response to antidepressants was on average smaller and dropouts were more likely when a placebo arm was included in a trial. For the same drug and the same probability of receiving placebo, greater dropout rates from all causes were associated with a smaller response to the treatment. Of note, dropout rates due to adverse events were not related to the probability of receiving placebo.

There are several explanations for the association between placebo arms, response rates and dropout observed in this study. First, patient expectations may influence the response to a treatment independently of the efficacy of the drug.[25] This unspecific, contextual response to treatment might be greater in patients enrolled in head-to-head trials, who know that they will receive an active drug, than in patients enrolled in placebo-controlled trials (Supplementary Appendix 11, available as Supplementary data at *IJE* online). The overall, absolute response to the drug observed in the study will consequently be larger in a head-to-head trial than in a placebo-controlled trial, even if the specific effect of the drug remains the same. Note that the contextual response can be measured in a placebo-controlled study (as the response to placebo) but not in a head-to-head study. The impact of differential contextual response might be amplified by patient differences; it is possible that patients recruited in placebo-controlled trials differ from those recruited in head-to-head studies in their attitude and pre-onceptions about psychopharmacological treatments. Interestingly, no association of response or dropout rates with the probability of receiving placebo was observed for amitriptyline and trazodone, possibly because the side effects made blinding difficult.

Second, the LOCF approach to imputing missing outcome data may have affected estimates of treatment responses.[26] In our study, more patients randomized to an active drug in a placebo-controlled trial dropped out than patients enrolled in a head-to-head trial of the same drug. Carrying their poor early response forward will result in estimates of response that are biased downwards, resulting in underestimation of the absolute response to active drugs in placebo-controlled studies. In our study, a high dropout rate was associated with poor response independently of the presence of a placebo arm. The use of LOCF may thus act jointly with patient expectations to produce the greater response in head-to-head studies. As only 9% of the included studies (46 studies) used appropriate methods to impute missing outcome data (multiple imputations or mixed-effect model repeated measure) examining only

newer drugs, a sensitivity analysis using only those studies would not be very informative.

A third explanation for the larger response rates in the head-to-head trials could be biases in conduct, analysis or reporting of these trials, driven by commercial interests.[27] This seems unlikely because funding by industry was not associated with important differences in response or dropout rates, in line with our previous analysis of new-generation antidepressants.[28] Furthermore, our literature search was comprehensive, and we included a substantial amount of unpublished data, including data from drug companies and licensing authorities, which are known to show smaller effects of antidepressants than published trials.[18]

Dropout due to any cause includes dropout due to inefficacy; this might explain why dropout in active arms in placebo-controlled studies is larger than dropout in active-controlled studies. As with response, the knowledge that there is a probability of being allocated to the placebo arm impacts on patients' improvement in symptoms and on their decision to leave the trial. In contrast, adverse events (such as nausea, sexual dysfunction or weight gain) are the direct results of the pharmacological substances and hence unlikely to be influenced by patient expectations. This can explain why dropouts due to adverse events are not modified by the presence of a placebo arm in trials.

This study is based on the largest systematic collection of published and unpublished antidepressant trials ever compiled.[17] Our analysis is the first to account for differences between the studied antidepressants and several patient and trial characteristics via multivariate meta-regression. However, our study has some limitations. The meta-regression analyses provide observational evidence that might be subject to residual confounding, or confounding by variables not included in our analyses. We did not have access to individual patient data and could not test our hypothesis that differences in response are due to the LOCF approach. Finally, we used the total number of arms in a trial to estimate the probability of receiving placebo by assuming equal randomization ratios in all study arms, ignoring other randomization ratios.

Our findings have potentially important implications for the evaluation of antidepressants. The Food and Drug Administration (FDA) issued guidelines for the evaluation of antidepressants, which have not been updated since 1977.[29] According to these guidelines, five different types of controls may be used in a clinical trial: (i) placebo concurrent control; (ii) dose-comparison concurrent control; (iii) no-treatment concurrent control; (iv) active-treatment concurrent control; and (v) historical control. A clear preference is, however, given to placebo-controlled studies over active-control trials: '...because alternative study designs,

especially active-control studies, may not be informative, exposing subjects to risk but without being able to collect useful information'.[29] Our study supports the notion that in the case of antidepressants, comparative effectiveness research may not only be more relevant for clinical and reimbursement decisions,[30] but also may be less biased than placebo-controlled trials research. In the case of antidepressants, we need phase III, superiority trials with an active comparator, chosen among the most effective and better-tolerated treatments on the market.[6]

In conclusion, a higher probability of receiving placebo in a randomized trial increased the chances of dropout and decreased the absolute response of patients to active antidepressants. This might be explained by decreased patient expectations and the use of the LOCF approach in combination with larger dropout rates in placebo-controlled trials. The synthesis of placebo-controlled and head-to-head trials in meta-analyses is common practice. Our study suggests that the strength of association between probability of receiving placebo and response might be different across the 16 drugs examined. This means that not only the response to the drug but also the relative differences in response between drugs is influenced by the probability of receiving placebo. Consequently, careful attention is needed when results from studies with different probabilities of receiving placebo are combined to estimate relative treatment effects. Future research should replicate our analysis in other areas of health care, especially in areas where response to placebo and patient expectations are important, for example in pain research.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

Acknowledgements

Author Contributions

the ICMJE criteria for authorship and agree with the results and conclusions of this Article.

References

1. Bradley C. Designing medical and educational intervention studies. A review of some alternatives to conventional randomized controlled trials. *Diabetes Care* 1993;**16**:509–18.

2. Sotsky SM, Glass DR, Shea MT *et al*. Patient predictors of response to psychotherapy and pharmacotherapy: findings in the NIMH treatment of depression collaborative research program. *Am J Psychiatry* 1991;**148**:997–1008.

3. Mergl R, Henkel V, Allgaier AK *et al*. Are treatment preferences relevant in response to serotonergic antidepressants and cognitive-behavioral therapy in depressed primary care patients? Results from a randomized controlled trial including a patients' choice arm. *Psychother Psychosom* 2010;**80**:39–47.

4. Bowling A, Rowe G, Lambert N *et al*. The measurement of patients' expectations for health care: a review and psychometric testing of a measure of patients' expectations. *Health Technol Assess* 2012;**16**:1–509.

5. Committee for Medicinal Products for Human Use (European Medicines Agency). *Reflection Paper on the Need for Active Control in Therapeutic Areas Where Use of Placebo is Deemed Ethical and One or More Established Medicines are Available*. EMA/759784/2010. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/01/WC500100710.pdf (5 April 2017, date last accessed).

6. Geddes JR, Cipriani A. Time to abandon placebo control in pivotal phase III trials? *World Psychiatry* 2015;**14**:306–07.

7. Kupfer DJ, Frank E. Placebo in clinical trials for depression: complexity and necessity. *JAMA* 2002;**287**:1853–54.

8. Furukawa TA, Cipriani A, Atkinson LZ *et al*. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *Lancet Psychiatry* 2016;**3**:1059–66.

9. Peciña M, Bohnert AS, Sikora M *et al*. Association between placebo-activated neural systems and antidepressant responses: neurochemistry of placebo effects in major depression. *JAMA Psychiatry* 2015;**72**:1087–94.

10. Barsky AJ, Saintfort R, Rogers MP, Borus JF. Nonspecific medication side effects and the nocebo phenomenon. *JAMA* 2002;**287**:622–27.

11. Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG. Meta-analysis of the placebo response in antidepressant trials. *J Affect Disord* 2009;**118**:1–8.

12. Khan A, Khan SR, Walens G, Kolts R, Giller EL. Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the food and drug administration summary basis of approval reports. *Neuropsychopharmacology* 2003;**28**:552–57.

13. Rutherford BR, Cooper TM, Persaud A, Brown PJ, Sneed JR, Roose SP. Less is more in antidepressant clinical trials. *J Clin Psychiatry* 2013;**74**:703–15.

14. Furukawa TA, Cipriani A, Leucht S *et al*. Is placebo response in antidepressant trials rising or not? A reanalysis of datasets to conclude this long-lasting controversy. *Evid Based Ment Health* 2018;**21**:1–3.

15. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol* 2009;**19**:34–40.

16. Sinyor M, Levitt AJ, Cheung AH *et al*. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. *J Clin Psychiatry* 2010;**71**:270–79.

17. Cipriani A, Furukawa TA, Salanti G *et al*. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;**391**:1357–66.

18. Furukawa TA, Salanti G, Atkinson LZ *et al*. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* 2016;**6**:e010919.

19. Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 2005;**20**:49–52.

20. Higgins JPT, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004;**23**:1663–82.

21. Mavridis D, Salanti G. A practical introduction to multivariate meta-analysis. *Stat Methods Med Res* 2013;**22**:133–58.

22. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*. Chichester, UK: Wiley, 2009.

23. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP, 2015.

24. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2013.

25. Hróbjartsson A. What are the main methodological problems in the estimation of placebo effects? *J Clin Epidemiol* 2002;**55**: 430–35.

26. Cook RJ, Zeng L, Yi GY. Marginal analysis of incomplete longitudinal binary data: a cautionary note on LOCF imputation. *Biometrics* 2004;**60**:820–28.

27. Perlis RH, Perlis CS, Wu Y, Hwang C, Joseph M, Nierenberg AA. Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. *Am J Psychiatry* 2005; **162**:1957–60.

28. Cipriani A, Furukawa TA, Salanti G *et al*. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;**373**:746–58.

29. Zimmerman M. The FDA's failure to address the lack of generalisability of antidepressant efficacy trials in product labelling. *Br J Psychiatry* 2016;**208**:512–14.

30. Li T, Vedula SS, Scherer R, Dickersin K. What comparative effectiveness research is needed? A framework for using guidelines and systematic reviews to identify evidence gaps and research priorities. *Ann Intern Med* 2012;**156**:367–77.