Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies

Toshi A Furukawa, Andrea Cipriani*, Lauren Z Atkinson, Stefan Leucht, Yusuke Ogawa, Nozomi Takeshima, Yu Hayasaka, Anna Chaimani, Georgia Salanti*

Summary

Background Previous studies have shown that placebo response rates in antidepressant trials have been increasing since the 1970s. However, these studies have been based on outdated or limited datasets and have used inappropriate statistical methods. We did a systematic review of placebo-controlled randomised controlled trials of antidepressants to examine associations between placebo-response rates and study and patient characteristics.

Methods In this systematic review, we searched for published and unpublished double-blind randomised placebo-controlled trials of first-generation and second-generation antidepressants for acute treatment of major depression in adults (update: Jan 8, 2016). The log-transformed proportions of placebo response, defined as 50% or greater reduction in depression severity score from baseline, were meta-analytically synthesised for each year. We then looked for a structural break point in the secular changes in these characteristics through the years and examined the influence of the study year and other trial and patient characteristics on the response rates through meta-regression.

Findings We identified 252 placebo-controlled trials (26 324 patients on placebo) done between 1978 and 2015. There was a structural break in 1991, and since then, the average placebo response rates in antidepressant trials have remained constant in the range between 35% and 40% (relative risk [RR] 1·00, 95% CI 0·97–1·03, p=0·99, for every 5-year increase). The length of the study and the number of study centres were significant factors (RR 1·03, 95% CI 1·01–1·05 for 1 more week in trial length; 1·32, 1·11–1·57 for multicentre *vs* single-centre trials).

Interpretation Contrary to the widely held belief, the average placebo response rates in antidepressant trials have been stable for more than 25 years. This new evidence should have an effect on the interpretation of the scientific literature and the future of psychopharmacology, both from a clinical and methodological point of view.

Funding Japan Society for Promotion of Science, Great Britain Sasakawa Foundation.

Introduction

It has been widely accepted that placebo response rates in antidepressant trials have been increasing over the past 30 years. The first systematic study[1] on this topic revealed that among the 75 placebo-controlled antidepressant trials up to 2000, there was a positive correlation between the proportion of responders on placebo and the year of publication (r=0·45, 95% CI 0·25–0·61). A very similar association was found in paediatric antidepressant trials (r=0·64, 95% CI 0·02–0·91).[2] Other studies examined pre–post raw change scores[3,4] or their effect sizes[5,6] in the placebo groups and confirmed a significant association between the thus-measured placebo response and the year of study publication.

Since the increasing placebo response was suspected of contributing to increasing numbers of so-called failed antidepressant trials,[7,8] many authors explored factors behind this relation in the hope of finding measures to curtail this increase. Walsh and colleagues[1] found that the length of the trial and the minimum depression severity required at baseline also influenced the placebo response rates in adult antidepressant trials; however, when jointly entered into multiple linear regression, only the year of publication remained significantly associated with the response to placebo.[1] Age, sex, use of placebo run-in or co-medication did not show significant association.[1] The number of recruiting sites and the number of randomised patients were identified as important factors associated with placebo response in paediatric trials.[2] Other patient or study characteristics suggested to be associated with placebo response include baseline severity,[3,6] number of groups or probability of being allocated to placebo,[9,10] dosing schedule,[11] length of trial,[3,12] and inflation of baseline severity.[13] However, all these empirical studies have methodological problems. Many of them are based on samples that are not only out of date now but also biased, because they included only published trials, often only in English language.[1,3,6,12,14] Other reviews included data from the US Food and Drug Administration, which thus precluded publication bias but also limited the number and scope of available studies.[4,5,11] Moreover, in terms of statistical analysis, previous studies examined the association between the proportion of placebo

Lancet Psychiatry 2016; 3: 1059–66

Published Online October 7, 2016 http://dx.doi.org/10.1016/S2215-0366(16)30307-8

See Comment page 1005

*These authors contributed equally

Department of Health Promotion and Human Behavior (Prof T A Furukawa, Y Ogawa MD, N Takeshima MD, Y Hayasaka MD), **Department of Clinical Epidemiology** (Prof T A Furukawa), **Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan; Department of Psychiatry, University of Oxford, Oxford, UK** (A Cipriani PhD, L Z Atkinson MSc); **Department of Psychiatry and Psychotherapy, Technische Universität München, Klinikum rechts der Isar, Munich, Germany** (Prof S Leucht MD); **Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland** (A Chaimani PhD, G Salanti PhD); **and Department of Hygiene and Epidemiology, University of Ioannina, Greece** (A Chaimani PhD, G Salanti)

Correspondence to: Dr Andrea Cipriani, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK **andrea.cipriani@psych.ox.ac.uk**

Panel: Research in context

Evidence before this study

It has been accepted knowledge that placebo response rates in antidepressant trials have been increasing over the past decades since 1970s. This finding has been associated with the increasing number of so-called "failed" antidepressant trials, and many authors have examined and debated factors behind this association, such as depressive symptoms' severity, inflation of baseline score, and the use of placebo run-in phases in the trial design. Unfortunately, previous reviews have been based on incomplete datasets or used inappropriate statistical methods (for instance, the association between the proportion of placebo responders and the year of publication was examined as correlation coefficients or linear regressions without taking into account the precision of estimated proportion in each included study and erroneously assuming normal distribution for proportions). We did a systematic and comprehensive search for published and unpublished randomised trials of first-generation and second-generation antidepressants for acute treatment of adult patients with major depression (update: January, 2016), to examine the association between proportion of placebo-responders and study and patient characteristics.

Added value of this study

We identified a dataset of 252 placebo-controlled double-blind studies done since 1978 (overall, 26 324 participants randomly assigned to placebo). In our large dataset, there was a structural break in 1991, and the average placebo response rate in antidepressant trials has basically remained constant in the range of 35–40% since 1991. We were able to replicate the reported increase in placebo response rates before 2000. However, we found that this analysis was confounded by methodological factors (such as shorter duration of trials and preponderance of single-centre studies), which were frequent in the very early trials and became less often employed since 1990s. Controlling for these variables, the association between placebo response rate and study year was no longer significant.

Implications of all the available evidence

Placebo response rates in antidepressant trials have, on average, stayed constant over the past 25 years. This new evidence should have an effect on the interpretation of the scientific literature and future psychopharmacology, with both clinical and methodological implications. In terms of study design, some important factors should be reconsidered. Appropriately timed assessments and multicentre design are necessary for trials to provide results directly relevant and applicable to real-world clinical practice. Trialists need to use design characteristics to make their trials as clinically relevant as possible.

responders and the year of publication as correlation coefficients or linear regressions without taking into account the precision of estimated proportion in each included study and erroneously assuming normal distribution for proportions.[1,2,14]

We did a comprehensive and systematic search for published and unpublished double-blind placebo-controlled trials of first-generation and second-generation antidepressants (the search was part of an update of a network meta-analysis about antidepressants in major depression).[15] We focused on the placebo response rates and, using the largest dataset so far, we aimed at addressing the following questions: whether the believed increase in placebo response rates had persisted up to 2015, and if not, whether we could replicate the previous findings,[1] and what patient and study characteristics could influence the studied associations.

Methods

Data sources and search criteria

We included all double-blind randomised trials comparing one of the following active drugs with placebo in the acute treatment of major depression: agomelatine, amitriptyline, bupropion, citalopram, clomipramine, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, nefazodone, paroxetine, reboxetine, sertraline, trazodone, venlafaxine, vilazodone, and vortioxetine. We included all the second-generation antidepressants licensed in Europe, USA, Australia, and Japan, and, of the older drugs, we selected the two tricyclics included in the WHO list of essential medicines (amitriptyline and clomipramine), plus trazodone and nefazodone because they had very distinct effect and tolerability profiles.[15] We included randomised trials with patients aged 18 years or older, of both sexes, and with a primary diagnosis of unipolar major depression according to standard operationalised diagnostic criteria. We searched the relevant electronic databases including Cochrane CENTRAL, CINAHL, Embase, LiLACS, MEDLINE, MEDLINE In-Process, and PsycINFO, supplemented by searches for published, unpublished, and ongoing randomised trials in major drug-approving agencies, clinical trial registries, and pharmaceutical company websites up to Jan 8, 2016. We also contacted the National Institute for Health and Care Excellence (NICE, UK), the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, Germany), and other relevant organisations and experts in the field for any additional information not already identified. For details, please refer to the full study protocol.[15] No date limits or language restrictions were applied to any of the searches.

At least two people (TAF, ACi, LZA, YO, NT, or YH) independently checked reference titles and abstracts thus identified. Full texts of all potentially eligible studies were retrieved, and were further inspected using the same

eligibility criteria. Any disagreements were resolved by discussion or in consultation with a third member of the review team. Two or more people (TAF, ACi, LZA, YO, NT, or YH) then independently extracted data using the pre-defined data extraction sheet. Two or more independent raters also assessed the quality of each study using the Cochrane risk of bias tool.[16] Each study was then classified as having low risk of bias (none of the domains rated at high risk of bias and three or fewer domains rated at unclear risk), moderate risk of bias (one domain rated at high risk of bias or none rated at high risk of bias but four or more at unclear risk), or high risk of bias (all other cases).[15]

As we obtained information from various sources including single trial publications, published meta-analyses, regulatory agencies' documents, and company website information, great efforts were expended to identify and avoid duplication, both before and after data extraction, and to contact original study authors for supplemental information, if needed.

Statistical analyses

We examined the proportions of placebo response in the placebo groups of all the identified trials in the following manner. We first tabulated the proportion of placebo response by half-decades since the earliest trial. The log-transformed proportion of responders was meta-analytically synthesised for each half decade, using the random effects model and the method of moments[17] to estimate the between-study component of variance τ^2 with the *metan* command in Stata version 14.1 (StataCorp, College Station, TX, USA). We defined response as a 50% or greater reduction in the total score from baseline to week 8 on a standardised observer-rating scale for depression. We used the Hamilton Rating Scale for Depression (HRSD) or, if HRSD was not used, another standardised and validated observer-rating scale such as the Montgomery-Asberg Depression Rating Scale (MADRS). Since we were interested in the acute treatment of major depression, assessments had to be done between 4 and 12 weeks, and the ones closest to 8 weeks were prioritised. We excluded longer-term studies from the statistical analyses if they did not provide data for the 4–12 week period. When the number of responders was not reported but baseline mean and endpoint mean and its standard deviation on the depression rating scale were provided, we calculated the number of responders by using a validated imputation method.[18] To abide by the intention-to-treat principle, we used the number of patients randomised as the denominator of our primary analysis and we assumed any participants with an unknown outcome to be non-responders in accordance with a validated method to estimate relative treatment effects.[19]

Next we examined if there was any structural break in this time series—ie, when a time series abruptly changes at a point in time, by applying the command *estat sbsingle* in Stata to our series of meta-analytically pooled response rates for each year from 1978 to 2015. The structural break test helps us to determine when and whether there is a significant change in the data series. We then ran a meta-regression of log of placebo response rate by the year of study completion. We gave preference to this variable over the year of publication because, by definition, the latter is unavailable for unpublished studies. We next ran meta-regression of log of placebo response rates by year of the study completion while controlling for various participant and study design characteristics. We recorded depression severity as measured by different versions of HRSD and MADRS into HRSD-17 scores by using the conversion table based on the item response theory.[20] We

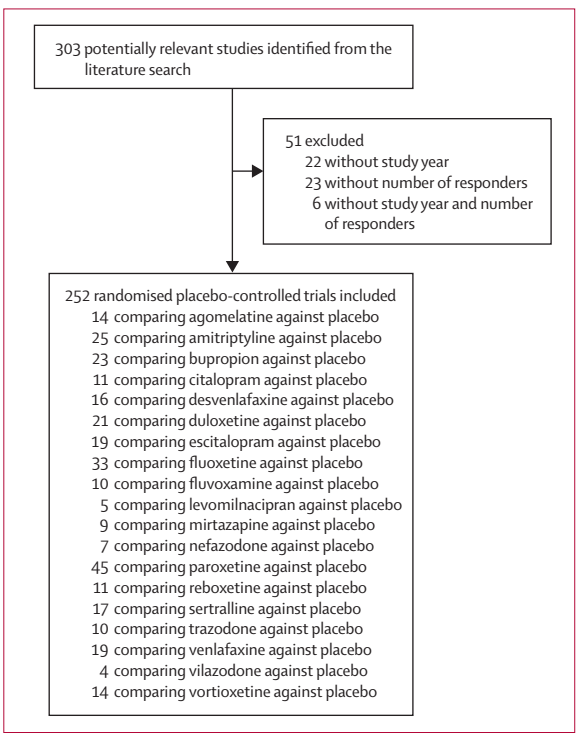

Figure 1: **Study selection**
*The total added number exceeds 252 because 60 trials involved in more than one active drug.

	Number of studies	Range of proportions of responders	Weighted mean proportion of responders (95% CI)	I^2, %	Tau
1978–85	27	0·00–0·41	0·30 (0·27–0·33)	12·0%	0·087
1986–90	32	0·08–0·50	0·28 (0·25–0·31)	32·1%	0·171
1991–95	34	0·14–0·54	0·35 (0·32–0·39)	71·6%	0·232
1996–2000	39	0·21–0·53	0·38 (0·35–0·41)	69·3%	0·181
2001–05	47	0·18–0·58	0·40 (0·37–0·42)	72·8%	0·176
2006–10	46	0·23–0·70	0·37 (0·34–0·40)	78·3%	0·207
2011–15	27	0·17–0·63	0·36 (0·33–0·40)	84·9%	0·232
All years	252	0·00–0·70	0·36 (0·35–0·37)	74·1%	0·211

Table 1: **Proportion of placebo responders by half-decades**

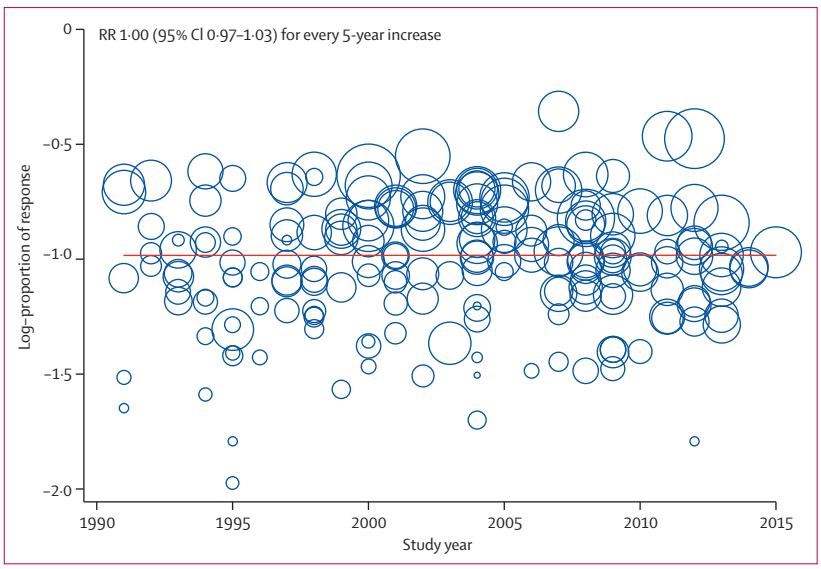

Figure 2: **Meta-regression of log-proportion of response to placebo by study year since 1991 (193 studies)**
RR=relative risk.

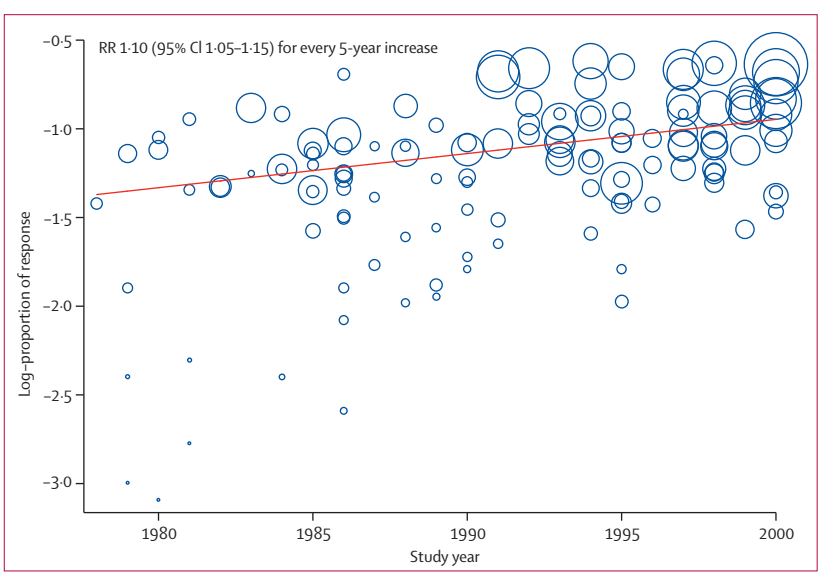

Figure 3: **Meta-regression of log-proportion of response to placebo by study year up to 2000 (132 studies)**
RR=relative risk.

See Online for appendix

entered each variable individually in addition to the study year first, and when the participant or study characteristics proved to exert significant influence, they were entered together to see whether the association between log of placebo response rates and the year of the study was affected. For meta-regression, we used the *metareg* command in Stata with the method of moments estimator for heterogeneity. The percentage of heterogeneity standard deviation that is explained by additional covariates was monitored and reported.

The influential study by Walsh and colleagues[1] was published in 2002 and included trials up to 2000. We examined whether we could replicate their finding by limiting our sample to trials up to the year 2000, while taking a statistically more appropriate dependent variable and employing meta-analytic methods with proper weighting. We first ran simple meta-regression by year of the study and then examined if additional study characteristics affected the association.

We did the following sensitivity analyses to test the robustness of our primary findings: (1) exclusion of studies that imputed the number of responders based on the continuous depression severity scores;[18] (2) use of the number of participants analysed instead of the number of participants randomised as denominator for calculations of the response rate; (3) use of the year of publication instead of the year of study completion; (4) exclusion of unpublished studies; (5) exclusion of studies at high or moderate risk of bias; (6) use of log-odds of placebo response instead of log of placebo response rate as the dependent variable; (7) use of pre–post change on the rating scale as the dependent variable.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. TAF and ACi had full access to all the data in the study and had final responsibility for the decision to submit for publication

Results

The literature search identified 303 placebo-controlled trials. However, in 51 of them either the year of the study or any information about efficacy was missing, so the remaining 252 studies constituted the dataset for our analysis (134 randomised trials with published data only, 74 with both published and unpublished data, and 44 studies with unpublished data only; figure 1). The characteristics of the included studies are summarised in the appendix. Between 1978 and 2015, placebo response rates ranged widely from 0% to 70% but the weighted mean proportion of responders appeared to converge towards 35% to 40%, especially after the early 1990s (table 1). It must be noted that overall the *I*-squared value remained high (74·1% overall), indicating substantial heterogeneity in placebo response rates among the included trials throughout these years.

The structural break test suggested that the break date was the year 1991 (p=0·04), which coincided well with the observed values reported in table 1. When pooled through meta-analysis, the summary response rate for this time period was 0·37 (95% CI 0·36–0·39) with a heterogeneity standard deviation of 0·203. We ran meta-regression to examine the trend after this date and also to explore possible confounders in the association. Figure 2 shows meta-regression of log of proportion of placebo response by the year of the study completion after 1991. Since 1991, the relative risk in the proportion

of placebo response was 1·00 (95% CI 0·97–1·03, p=0·99) for every 5-year increase in the study year. When individual participant or study design characteristics were examined, the length of the study and the number of study centres emerged as significant factors: the longer studies had greater response rates, and the multicentre trials had greater response rates than single-centre studies (RR 1·03, 95% CI 1·01–1·05 for 1 more week in trial length; 1·32, 1·11–1·57 for multicentre trials *vs* single-centre trials). When both variables were entered together in the meta-regression, both remained significant predictors (RR 1·03, 95% CI 1·00–1·05, for trial duration; 1·31, 1·10–1·56 for multicentre trials) whereas the study year remained non-significant (0·99, 0·96–1·01, for every 5-year increase). This multivariable model explained 2·2% of the total heterogeneity standard deviation.

We examined if we could replicate Walsh and colleagues' findings[1] by limiting the studies up to the year 2000. The summary response rate from simple meta-analysis was 0·34 (95% CI 0·32–0·35) with a heterogeneity standard deviation 0·223. Meta-regression of log of proportion of placebo responders by the study year up to 2000 clearly indicated the increasing placebo response (RR 1·10, 95% CI 1·05–1·15, p<0·0001, for every 5-year increase; figure 3, table 2). When possible confounders were examined individually in meta-regression, the number of study centres, the dosing schedule, the length of the trial, and the number of patients randomised emerged as possible confounders (table 2). In addition to the same tendencies observed for the studies after 1991 for the number of weeks and the number of study centres, studies using fixed dose regimens showed lower placebo response rates than those using flexible dose regimens, and the smaller groups tended to show lower response rates. When all these four variables were entered into multivariable meta-regression, the mean response rate for a hypothetical 4-week, single-centre trial with fixed dose regimen in the earliest year of publication (1978) was estimated to be 0·19 (95% CI 0·16–0·23). The response rate increases for a multicentre trial, for a flexible dosing schedule, and for 1 more week in the length in the trial were still significant (table 2). However, the effect of the study year and the number of patients randomised were no longer significant. This multivariable model was able to explain 28·4% of the total heterogeneity and 18·4% of the residual heterogeneity of the meta-regression model with the study year alone. When we examined the secular changes in these characteristics through the years, we found that the studies were much shorter and more often single-centred in the 1980s than in 1990s and later, whereas the flexible dose regimen became increasingly unpopular in the 2000s and later (table 3).

All sensitivity analyses revealed no effect of year among studies after 1991, providing support for robustness of our main findings (appendix).

Discussion

In our study, we found that the average placebo response rate in antidepressant trials has remained constant in the range of 35–40% since 1991, year at which we found a structural break. We were able to replicate the reported increase in placebo response rates before 2000, but this association was no longer significant when controlling for methodological factors (such as shorter duration of trials and preponderance of single-centre studies), which

	Number of studies	Study year RR (for every 5-year increase, 95% CI)	Covariate RR (95% CI) [interpretation]	Tau
Model with year as covariate				
Year	132	1·10 (1·05–1·15)	..	0·196
Models with year and one additional covariate				
Additional covariate				
Age	119	1·10 (1·05–1·15)	1·00 (0·92–1·08) [for every 10-year increase in mean age]	0·198
Sex	57	1·10 (1·02–1·18)	1·03 (1·00–1·07) [for every 10% increase in proportion of women]	0·208
Region	130	1·10 (1·05–1·15)	0·94 (0·77–1·14) [North America *vs* Europe]; 1·03 (0·81–1·31) [cross-continental *vs* Europe]	0·202
Number of study centres	118	1·07 (1·02–1·11)	1·37 (1·21–1·56)[multicentre *vs* single-centre]	0·180
Setting*	66	1·06 (1·00–1·14)	0·86 (0·70–1·06) [secondary or tertiary care *vs* primary care]	0·205
Patient status†	119	1·10 (1·05–1·15)	1·00 (0·87–1·16) [outpatients *vs* inpatients]	0·202
Baseline eligibility threshold‡	108	1·09 (1·04–1·14)	0·96 (0·85–1·07) [for every five point increase in HAMD17 threshold score]	0·193
Baseline severity‡	111	1·09 (1·04–1·14)	0·91 (0·83–1·00) [for every five point increase in baseline HAMD17]	0·181
Number of arms	132	1·10 (1·05–1·15)	1·13 (0·78–1·63) [one arm more]	0·198
Placebo run-in	106	1·10 (1·04–1·15)	0·89 (0·75–1·06) [present *vs* absent]	0·192
Flexible *vs* fixed dose regimen	131	1·10 (1·06–1·15)	1·13 (1·01–1·27) [flexible *vs* fixed]	0·187
Length of trial (in weeks)	124	1·05 (1·00–1·11)	1·05 (1·02–1·08) [1 week more]	0·182
Rescue medication	57	1·07 (0·99–1·15)	0·90 (0·76–1·07) [present *vs* absent]	0·191
Number of randomised patients	132	1·08 (1·03–1·13)	1·11 (1·02–1·22) [100 participants more]	0·196
Multivariable meta-regression (using all significant variables)				
Study year	112§	1·04 (0·99–1·09)§	..	0·160§
Number of study centres	..	..	1·41 (1·22–1·62)	..
Flexible *vs* fixed dose regimen	..	..	1·17 (1·04–1·31)	..
Length of trial	..	..	1·04 (1·01–1·06)	..
Number of randomised patients	..	..	0·97 (0·89–1·07)	..

RR=relative risk. *Studies done in both primary and secondary or tertiary care were classified as ones in primary care. †Studies recruiting both inpatients and outpatients were classified as ones recruiting inpatients. ‡Converted into Hamilton Rating Scale for Depression-17 using the conversion table.[20] §Results for all five variables.

Table 2: **Factors influencing proportion of placebo responders before 2000**

	Number of studies	Mean length of trial, weeks (range)	Multicentre trials, n (%)	Flexible dosing, n (%)
1978–1985	27	4·6 (4–6)	13 (48·2%)	23 (85·2%)
1986–1990	32	5·8 (4–8)	11 (34·4%)	24 (75·0%)
1991–1995	34	6·7 (4–12)	22 (64·7%)	25 (73·5%)
1996–2000	39	8·0 (4–12)	34 (87·2%)	29 (74·4%)
2001–2005	47	7·8 (6–12)	44 (93·6%)	20 (42·6%)
2006–2010	46	7·6 (6–12)	45 (97·8%)	18 (39·1%)
2011–2015	27	8·1 (6–12)	25 (92·6%)	3 (11·1%)
All years	252	7·1 (4–12)	194 (77·0%)	142 (56·4%)

Table 3: **Secular changes in study characteristics of placebo-controlled trials of antidepressants**

were frequent in the very early trials and became less often used since 1990s.

This is the largest systematic review of antidepressant randomised trials done so far and sheds new light in the current heated debate about diminishing efficacy of antidepressant drugs. First, the average placebo response rates have stayed constant since 1991 in the range of 35–40%. Second, the myth of increasing placebo response was due to the exceptionally low rates in 1980s and before, which were confounded especially by the shorter duration of the trials and the preponderance of single-centre studies. These findings are quite contrary to the heretofore almost unanimously held belief in the medical literature that the antidepressant trials are fraught by increasing placebo response rates (interestingly, the meta-analysis by Rief and colleagues[6] found that only doctor-reported outcomes, such as HAMD scores, showed the rise in placebo response of studies published before 2000, but this was not true for patient-reported outcomes).[1–6,14]

Based on our large dataset, however, there is no doubt now that the average placebo-response rates have been staying constant since 1990s, as far as antidepressant trials in depression are concerned. These results are robust since they remained consistent in all the sensitivity analyses. It might well be the case, however, that the association is different in other psychiatric disorders.[21,22]

The second insight may do away with much speculation in the literature about the causes of increasing placebo response. Of the variables found to be significant in the scientific literature, when we examined, in our much larger dataset and with appropriate statistical methods, the length of trial, the number of study centres, the number of randomised patients, the baseline severity, the number of groups, and the dosing schedule, only the length of trial and the number of study centres were found to exert consistent influence independently of each other. Previous findings of the association between increasing placebo response rates and study year became non-significant when we controlled for the confounding factors. In other words, the reported increase before 2000 is a methodological artifact because, when the relevant trial design characteristics were held relatively constant after 1990s, the placebo response rates stayed stable. Publication status did not influence this association since exclusion of unpublished studies did not change the findings (appendix).

There are several strengths of our study. First, we were able to include 252 placebo-controlled studies of antidepressants published between 1978 and 2015, in contrast to the 75 studies included in the study by Walsh and colleagues (up to 2000)[1] or 107 studies included in the study by Undurraga and colleagues (up to 2010).[14] We cannot rule out the possibility that our search still failed to identify some unpublished studies; however, the total number of participants on placebo in the current study was three-to-four times larger than in the previous studies (26 323 *vs* 6285[1] or 9925[14]). Second, we used methods that were more appropriate statistically throughout the analyses by taking the log of response rate instead of raw rates, by weighting the studies by their sample sizes, and by using the objective method to detect the change point in the longitudinal time series.

This study has some limitations. We included trials of an a-priori defined list of first-generation and second-generation antidepressants but we were not able to include all medications known to act as antidepressants to date. However, this approach was inevitable as there is no uniformly accepted list of effective antidepressants and some compounds might be approved for major depression in some countries but not in others. Moreover, by specifying the drugs to focus on, we think we were able to conduct a more systematic and comprehensive search for both published and unpublished trials, which is demonstrated by the size of the current dataset that far exceeds those of any previous study on the topic. To be clinically informative, we included only licensed drugs, therefore we did not include studies of drugs whose development failed. This failure could have happened because of very high response rates. The non-inclusion of placebo response rates from failed development programmes could have biased our sample of eligible studies toward a lower rate of placebo response; however, we carried out a comprehensive search of more than 250 published and unpublished trials and we do not think that the potential omission of a handful of such studies would materially affect the overall results.

The substantial heterogeneity as suggested by high *I*-squared values suggests that the placebo response rates remain highly variable through the years. To address this methodological issue, we therefore used the random effects meta-analytic methods.[23] Only access to individual patient data would allow researchers to better evaluate moderators and mediators of treatment effect; however, this is not always easy and often very time consuming.[24] Moreover, this analysis should be carried out in a systematic way and across multiple drugs, and not limited to a small sample of available studies about one drug, as previously done in the field of antidepressant trials.[25]

When we replicated Walsh and colleagues' findings before 2000, very early studies were on average smaller than those published later. This difference means that in the meta-regression analysis the study year was correlated with study weights, potentially challenging the assumptions of a joint analysis. When the sample size was examined as a univariable explanatory variable, it was a significant predictor but lost its significance when entered together with other covariates in the multivariable model (table 2). We also ran a post-hoc non-weighted multivariable regression, which indicated that neither the study year nor the sample size was a significant predictor, differently from what was previously reported.[26]

In conclusion, even if the average response to placebo has remained constant over the past 25 years, 35–40% is still a high proportion of patients. Assuming this is the best estimate we have of an average response to placebo, important implications for clinicians and researchers can be drawn. The expectation of improvement, classic conditioning, and the contact with a health-care environment with supportive and therapeutic features contribute to the objective response observed in patients with major depression who are randomly assigned to placebo.[27] These non-pharmacological aspects, however, are usually not provided to the same extent in standard clinical practice. Clinicians should create a specific context and level of therapeutic contact, to enhance non-specific effects of treatment and gain greater treatment response.[28] In terms of research, innovation in psychopharmacology is urgently needed not only for drug discovery and development, but also in terms of clinical trials' design. The key question is whether there is still need to have placebo-controlled phase 3 studies.[29] Clinical research organisations have been running trials in depression for most pharmaceutical companies since the early 1990s and they are financially incentivised to have as many visits as possible to increase the duration and therefore the cost per patient of the trial. Moreover, they tend to use as many centres as possible so that the study is completed sooner. To evaluate the next putative antidepressants, future studies should require the development of more efficient study designs to improve signal detection in drug development studies and an increased antidepressant response in clinical treatment.[30] It is time for academics, pharmaceutical industry, and regulators to create new models of drug discovery and drive innovation in the methodology of drug development.[31]

Contributors

TAF and ACi conceived the study and drafted the manuscript. GS, ACi, ACh, and TAF designed the methods. TAF, ACi, LZA, YO, NT, and YH selected the articles and extracted the data. TAF, ACh, and GS analysed the data. TAF, ACi, and SL interpreted the results. All authors critically revised the manuscript. All authors read and met the International Committee of Medical Journal Editors (ICMJE) criteria for authorship and agree with the results and conclusions of this Article.

Declaration of interests

TAF has received lecture fees from Eli Lilly, Janssen, Meiji, MSD, Otsuka, Pfizer, and Tanabe-Mitsubishi, and consultancy fees from Sekisui Chemicals and Takeda Science Foundation; has received royalties from Igaku-Shoin and Nihon Bunka Kagaku-sha publishers; and has received grant or research support from Mochida and Tanabe-Mitsubishi; he is diplomate of the Academy of Cognitive Therapy. ACi was expert witness for Accord Healthcare for a patent issue about quetiapine extended release. SL has received honoraria for lectures from Eli Lilly, Lundbeck (Institute), Pfizer, Janssen, BMS, Johnson and Johnson, Otsuka, Roche, Sanofi-Aventis, ICON, Abbvie, AOP Orphan, and Servier; has received honoraria for consulting or advisory boards from Roche, Janssen, Lundbeck, Eli Lilly, Otsuka, and TEVA; for the preparation of educational material and publications from Lundbeck Institute and Roche. Eli Lilly has provided medication for a clinical trial led by SL as principal investigator. The other authors declare no competing interests.

Acknowledgments

This study was supported in part by JSPS KAKENHI (Grant-in-Aid for Scientific Research) Grant Number 26670314 to TAF and by Great Britain Sasakawa Foundation Butterfield Award 4939 to ACi and TAF. ACi is supported by the NIHR Oxford Cognitive Health Clinical Research Facility and ACh by the International Short Visit Program of the Swiss National Foundation (Grant Number IZK0Z3_167541).

References

1 Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* 2002; **287:** 1840–47.

2 Bridge JA, Birmaher B, Iyengar S, Barbe RP, Brent DA. Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. *Am J Psychiatry* 2009; **166:** 42–49.

3 Khan A, Bhat A, Kolts R, Thase ME, Brown W. Why has the antidepressant-placebo difference in antidepressant clinical trials diminished over the past three decades? *CNS Neurosci Ther* 2010; **16:** 217–26.

4 Khin NA, Chen YF, Yang Y, Yang P, Laughren TP. Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. *J Clin Psychiatry* 2011; **72:** 464–72.

5 Fountoulakis KN, Moller HJ. Efficacy of antidepressants: a re-analysis and re-interpretation of the Kirsch data. *Int J Neuropsychopharmacol* 2011; **14:** 405–12.

6 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.

7 Montgomery SA. The failure of placebo-controlled studies. ECNP Consensus Meeting, September 13, 1997, Vienna. European College of Neuropsychopharmacology. *Eur Neuropsychopharmacol* 1999; **9:** 271–76.

8 Rutherford BR, Roose SP. A model of placebo response in antidepressant clinical trials. *Am J Psychiatry* 2013; **170:** 723–33.

9 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.

10 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.

11 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.

12 Rutherford BR, Cooper TM, Persaud A, Brown PJ, Sneed JR, Roose SP. Less is more in antidepressant clinical trials: a meta-analysis of the effect of visit frequency on treatment response and dropout. *J Clin Psychiatry* 2013; **74:** 703–15.

13 Landin R, DeBrota DJ, DeVries TA, Potter WZ, Demitrack MA. The impact of restrictive entry criterion during the placebo lead-in period. *Biometrics* 2000; **56:** 271–78.

14 Undurraga J, Baldessarini RJ. Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. *Neuropsychopharmacology* 2012; **37:** 851–64.

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

16 Higgins JP, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. http://www.cochrane-handbook.org (accessed June 28, 2016).

17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7:** 177–88.

18 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.

19 Spineli LM, Higgins JP, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. *Clin Trials* 2013; **10:** 378–88.

20 Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol* 2006; **16:** 601–11.

21 Stein DJ, Baldwin DS, Dolberg OT, Despiegel N, Bandelow B. Which factors predict placebo response in anxiety disorders and major depression? An analysis of placebo-controlled studies of escitalopram. *J Clin Psychiatry* 2006; **67:** 1741–46.

22 Agid O, Siu CO, Potkin SG, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970–2010. *Am J Psychiatry* 2013; **170:** 1335–44.

23 Nikolakopoulou A1, Mavridis D, Salanti G. Demystifying fixed and random effects meta-analysis. *Evid Based Ment Health* 2014; **17:** 53–57.

24 Drazen JM. Sharing individual patient data from clinical trials. *N Engl J Med* 2015; **372:** 201–02.

25 Mancini M, Wade AG, Perugi G, Lenox-Smith A, Schacht A. Impact of patient selection and study characteristics on signal detection in placebo-controlled trials with antidepressants. *J Psychiatr Res* 2014; **51:** 21–29.

26 Weimer K, Colloca L, Enck P. Placebo effects in psychiatry: mediators and moderators. *Lancet Psychiatry* 2015; **2:** 246–57.

27 Fava M. Implications of a biosignature study of the placebo response in major depressive disorder. *JAMA Psychiatry* 2015; **72:** 1073–74.

28 Cipriani A, Geddes JR. Placebo for depression: we need to improve the quality of scientific information but also reject too simplistic approaches or ideological nihilism. *BMC Med* 2014; **12:** 105.

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

30 Rutherford BR, Roose SP. A model of placebo response in antidepressant clinical trials. *Am J Psychiatry* 2013; **170:** 723–33.

31 Norman TC, Bountra C, Edwards AM, Yamamoto KR, Friend SH. Leveraging crowdsourcing to facilitate the discovery of new medicines. *Sci Transl Med* 2011; **3:** 88mr1.