

Association of Single-Blind Placebo Run-in Periods With the Placebo Response in Randomized Clinical Trials of Antidepressants: A Systematic Review and Meta-analysis

Amelia J. Scott, PhD; Louise Sharpe, PhD; Veronica Quinn, PhD; Ben Colagiuri, PhD

[+ Supplemental content](#)

IMPORTANCE Single-blind placebo run-in (PRI) periods are common in randomized clinical trials (RCTs) of treatment for depression. They aim to increase sensitivity to detect drug effects; however, the association of PRI periods with study outcomes remains unclear. This is concerning given the costs of PRI periods to patients and investigators.

OBJECTIVE To examine the association of the use of PRI periods with the placebo response, drug response, and drug-placebo difference among RCTs of antidepressants.

DATA SOURCES MEDLINE, Embase, the Cochrane Central Register of Controlled Trials, and PsycINFO, as well as repositories of unpublished studies, were systematically searched up to July 2021.

STUDY SELECTION Included studies were double-blind, placebo-controlled RCTs of antidepressant medication among adults with depressive disorders.

DATA EXTRACTION AND SYNTHESIS Data were extracted into a coding sheet, including the characteristics of studies, the characteristics of PRI periods, and the outcomes of studies.

MAIN OUTCOMES AND MEASURES Study outcomes were the primary depression symptom measure reported by the RCT. These outcomes were used to calculate effect sizes (Hedges g) of the within-group drug response and placebo response as well as the drug-placebo difference. Random-effects meta-analysis was used to calculate effect sizes, and subgroup analyses were used to compare outcomes depending on use of PRI periods.

RESULTS A total of 347 trials (representing 89 183 participants) were included; 174 studies (50%) reported using a single-blind PRI period. Response outcome data were available for 189 studies. Studies using PRI periods reported a smaller placebo response ($g = 1.05$ [95% CI, 0.98-1.11]; $I^2 = 82\%$) than studies that did not use a PRI period ($g = 1.15$ [95% CI, 1.09-1.21]; $I^2 = 81\%$; $P = .02$). Subgroup analysis showed a larger drug response size among studies that did not use a PRI period ($g = 1.55$ [95% CI, 1.49-1.61]; $I^2 = 85\%$) than those that did use a PRI period ($g = 1.42$ [95% CI, 1.36-1.48]; $I^2 = 81\%$; $P = .001$). The drug-placebo difference did not differ by use of PRI periods ($g = 0.33$ [95% CI, 0.29-0.38]; $I^2 = 47\%$ for use of a PRI period vs $g = 0.34$ [95% CI, 0.30-0.38]; $I^2 = 54\%$ for no use of PRI periods; $P = .92$). The likelihood of response to drug vs placebo also did not differ between studies that used a PRI period (odds ratio, 1.89 [95% CI, 1.76-2.03]) and those that did not use a PRI period (odds ratio, 1.77 [95% CI, 1.65-1.89]; $P = .18$).

CONCLUSIONS AND RELEVANCE This study suggests that RCTs using PRI periods yield smaller within-group changes across both placebo and drug groups compared with RCTs without PRI periods. The reduction in effect size across groups was equivalent in magnitude. Consequently, PRI studies do not observe larger drug-placebo differences, suggesting that they do not increase trial sensitivity. As such, given the resources and probable deception required and risk to external validity, the practice of using PRI periods in RCTs of antidepressants should be ended.

JAMA Psychiatry. doi:10.1001/jamapsychiatry.2021.3204
Published online November 10, 2021.

Author Affiliations: School of Psychology, University of Sydney, Sydney, New South Wales, Australia (Scott, Sharpe, Colagiuri); Department of Psychology, eCentreClinic, Macquarie University, Sydney, New South Wales, Australia (Scott, Quinn).

Corresponding Author: Amelia J. Scott, PhD, Department of Psychology, eCentreClinic, Macquarie University, Balacava Road, Macquarie Park NSW Australia, 2109 (amelia.scott@mq.edu.au).

Depression is a common condition affecting approximately 1 in 10 adults.¹ In health care settings, the detection and management of depression have steadily improved over time.² Prescribing practices and decision-making in such settings rely on evidence from randomized clinical trials (RCTs). Thus, it is crucial that these RCTs are conducted in a way that maximizes their validity and minimizes costs.

A common practice within RCTs of antidepressants is the use of single-blind placebo run-in (PRI) periods. In typical PRI periods, all participants are given placebo prior to randomization. Then, those who demonstrate improvement within this period are excluded, and the remaining participants are randomly assigned to the treatment or control group. Placebo run-in periods have been used for many decades^{3,4}; however, their theoretical and/or empirical basis is unclear. Early studies suggest that the use of PRI periods was based largely on intuition³; some presumed that eliminating early placebo responders would lower placebo response rates after randomization, thereby increasing the difference between placebo and active treatments.

Some early reviews have examined the association of the use of PRI periods with the outcomes in RCTs of antidepressants.^{3,4} In 1994, Trivedi and Rush³ analyzed 101 RCTs and found that the use of PRI periods did not alter the drug-placebo difference in response, nor specifically the response rates within the drug and placebo groups. In 2004, Lee and colleagues⁴ reexamined this issue among 43 RCTs of selective serotonin reuptake inhibitors, focusing on continuous depression outcomes as a more sensitive indication of symptom change. Lee and colleagues⁴ also found that the drug-placebo difference was not associated with the use of PRI periods; however, they noted a numerically larger effect in studies using PRI periods. On this basis, the authors concluded that their analyses were potentially underpowered and speculated that the use of PRI periods may continue in order to achieve small advantages. Since then, the use of PRI periods in RCTs of antidepressants has persisted, although the prevalence of the use of PRI periods is unknown, as is whether the use of PRI periods has changed over time.⁵⁻⁷

There is thus a need to further clarify the association of PRI periods with study outcomes. When examining within-group change, it may be that small differences emerge in 1 or more trial groups that may affect the drug-placebo difference. The study by Lee and colleagues⁴ did not examine within-group changes, whereas Trivedi and Rush³ examined only the proportion of people reporting an improvement of 50% or more, which may have been insensitive to detect small effects. Given that almost 20 years have passed since the study by Lee and colleagues⁴ and given claims of an increasing placebo response over time,⁸ the effect of PRI periods may have also changed.

Whether or not PRI periods improve trial sensitivity is nontrivial given the high costs to investigators and participants. Placebo run-in periods also present a notable departure from what would be expected in clinical practice and therefore minimize the external validity of RCTs. The ethical implications of PRI periods are also questionable.⁹ This practice first

Key Points

Question Is the use of a placebo run-in (PRI) period associated with different outcomes among randomized clinical trials (RCTs) of antidepressants?

Findings In this systematic review and meta-analysis of 347 RCTs of antidepressants comprising 89 183 participants, the use of PRI periods was associated with a lower placebo and drug response but was not associated with the drug-placebo difference.

Meaning This study suggests that the use of PRI periods is common in RCTs of antidepressants, despite offering no apparent benefits to RCT outcomes; given the risks and costs of PRI periods, their practice should be ceased.

involves the exclusion of treatment-seeking participants who are eligible at the beginning of the trial and who may benefit from full participation. In addition, a form of deception is required to conduct a single-blind PRI period. Within double-blind RCTs, participants provide informed consent to be assigned to receive either drug or placebo. Conversely, a single-blind PRI period involves deceiving participants into believing that they may receive a real drug during this phase, while all participants are given placebo.⁹ Ethical guidelines stipulate that deception should be used in research only when methodologically necessary; therefore, PRI periods cannot be justified if they do not improve trial sensitivity.

Given these issues, there is a need to carefully examine the use of PRI periods in RCTs of antidepressants and whether study outcomes differ depending on their use. To achieve this, we systematically reviewed and analyzed the frequency, nature, and outcomes associated with the use of PRI periods in RCTs of antidepressants for depressive disorders. Specifically, we sought to (1) estimate the prevalence of the use of PRI periods, (2) understand the characteristics of single-blind PRI periods and the number of participants excluded during the PRI period, and (3) examine the association of PRI periods with the placebo response, drug response, and drug-placebo difference, using continuous outcome measures.

Methods

This systematic review and meta-analysis was prospectively registered on the PROSPERO register (CRD42020139768; see eTable 5 in the [Supplement](#) for changes to protocol after registration) and conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline.¹⁰ Included studies were double-blind, placebo-controlled RCTs of currently approved and regulated antidepressant medication for adult participants who met recognized diagnostic criteria for a depressive disorder. We excluded studies that included individuals with bipolar depression, nonoriginal RCTs (eg, secondary, subgroup, or long-term analyses), RCTs with nonparallel designs (eg, crossover designs), RCTs that included nondrug treatment groups (eg, psychotherapy and acupuncture), and studies investigating

adjunctive medication (eg, drug A and drug B, drug A and placebo).

Search Strategy and Study Selection

We systematically searched the literature for published and unpublished records. The electronic databases MEDLINE, Embase, PsycINFO, and the Cochrane Register of Controlled Trials were searched up to July 2021, with no lower date limits. Search terms were constructed according to 3 stems relating to (1) depressive disorders, (2) antidepressant medications, and (3) RCTs. Searches were limited to English-language records. See the eMethods in the Supplement for a full description of study search terms. The US National Library of Medicine (ClinicalTrials.gov), European Clinical Trials Register, and the reference lists of existing reviews of unpublished data¹¹ were also searched for unpublished records.

After search and retrieval, titles and abstracts of retrieved records were reviewed by 1 researcher (A.J.S.), with a random 10% reviewed by a second researcher. Full texts were reviewed independently by 2 researchers, with disagreements resolved through consensus.

Data Extraction

Data were extracted into a predesigned coding sheet, including relevant trial and sample characteristics. We also coded whether studies used a PRI period or not. For studies using a PRI period, we recorded its length, the criteria for excluding participants, the rationale for using the PRI period, and the number of participants excluded. Relevant outcome data were also extracted, namely, the mean (SD) scores or the mean (SD) difference in scores on the primary depression outcome between baseline and the primary end point. Response rate data in each treatment group (ie, those reporting $\geq 50\%$ reduction in the primary outcome) were also extracted for secondary analysis. When outcome data were incomplete, further data were requested from contactable authors (ie, where a corresponding email was provided or obtained) on 2 occasions.

Data Analysis

We first conducted a descriptive analysis of study characteristics. Potential factors associated with the use of a PRI period were examined using binary logistic regression analyses, including year of publication, sample size, trial length, study location, industry sponsorship status, study setting, type of depressive disorder, antidepressant class, and risk of bias judgment. For these analyses, SPSS, version 25 (IBM Corp) was used.

Meta-analysis

A meta-analysis was conducted in Comprehensive Meta-analysis, version 3 (Biostat Inc) using random-effects models. We calculated treatment effect size using Hedges g , which quantifies the mean difference between groups divided by the pooled SD and includes a correction for small sample size.¹² We calculated the Hedges g for 3 outcomes of interest: the within-group change (ie, from pretreatment to

posttreatment) for the drug and placebo groups, as well as the drug-placebo difference. To calculate the within-group Hedges g for studies reporting only baseline and posttreatment mean (SD) values, the correlation between observed scores is needed. Because this was not reported in any studies, fixed pre-post correlation was assumed in this meta-analysis. Previous meta-analyses have suggested imputing values ranging from 0.2¹³ to 0.7.¹⁴ We therefore adopted a value of 0.5 for both within-group effects and conducted sensitivity analyses using a pre-post correlation of 0.2 and 0.7. No change in the pattern of results was identified. For the between-group effects, we approached multiplicity (by multiple comparisons between separate drug groups and a single placebo group) by adjusting the sample size of the placebo group, as recommended.¹⁵

Heterogeneity was examined with the Cochran Q value and its P value,¹⁶ although the Q test can be overly sensitive with larger samples of studies.¹⁷ Heterogeneity was also reported using the I^2 statistic and its 95% CIs, which report the percentage of variance in observed effects due to variance in true effects rather than sampling error.¹⁸ The τ^2 value was reported as a measure of between-study variance.¹⁹ Finally, 95% prediction intervals were calculated to assess the dispersion of effect sizes and identify where 95% of true effects would be expected to fall.²⁰

Moderator Analysis

Subgroup analysis was used to determine whether the Hedges g values differed depending on the use of PRI periods. We used the mixed-effects model, where studies within subgroups were pooled using the random-effects model and between-group differences were tested using the fixed-effects model.¹⁹

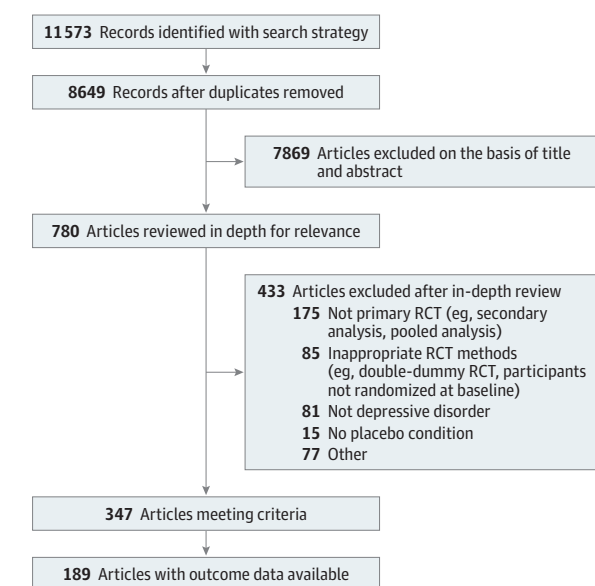
Secondary Analyses

We repeated the primary moderator analyses to determine whether the pooled response rate outcomes (reported as the pooled prevalence of response and the odds ratio between treatment and placebo) differed depending on the use of PRI periods. To further explore heterogeneity, the moderator analyses for studies that used a PRI period vs those that did not use a PRI period were repeated within a range of subgroups, including drug class, industry sponsorship, risk of bias rating, and publication status. Finally, meta-regression was also used to determine whether the number of participants excluded during the PRI period was associated with the placebo response, treatment response, or between-group effect size.

Risk of Bias

Study quality and risk of bias were evaluated using the Cochrane Risk of Bias tool.²¹ This tool assesses the potential risk of bias arising from the randomization process, deviations from the intervention, missing outcome data, measurement of the outcome, and selection of the reported result. Risk of bias regarding selection of the reported result was modified for studies published prior to 2004, given that earlier studies were not expected to preregister or publish RCT protocols. Across each domain and overall, risk of bias was rated as either being low, having some concerns, or being high.²¹

Figure 1. Search and Retrieval Process



RCT indicates randomized clinical trial.

Publication Bias

We examined publication bias for the drug-placebo difference by inspecting funnel plot symmetry (which plots each study effect size against its SE), along with the Egger weighted regression as a statistical test of asymmetry.²² The trim-and-fill analysis of Duval and Tweedie²³ was used to estimate the number of missing studies and the estimated effect size if these studies were included in the meta-analysis.

Results

After systematic search, 11 573 articles were identified (including 2924 duplicates). A total of 780 remained after title and abstract review. After full-text review, 347 studies (representing 89 183 participants) met eligibility criteria and were retained for data extraction (Figure 1). See Table 1 for summary statistics of included studies and eTable 1 in the Supplement for characteristics of individual studies.

Frequency and Nature of PRI Periods

Of the 347 studies included, 174 (50%) used a single-blind PRI period. Several study features were associated with the use of PRI periods, including year of publication, study location, study sponsorship, risk-of-bias rating, and drug class (Table 2).

Length of PRI Period

Most studies specified a fixed PRI period (102 of 174 [59%]), whereas 69 (40%) reported a variable PRI period (eg, 7-14 days). Three studies did not specify the length of the PRI period. Among studies reporting fixed length, the length of PRI periods ranged from 3 to 84 days (median, 7 days; mean [SD], 8.7 [8.0] days).

Table 1. Summary of Included Study Characteristics

Characteristic	No. (%) (N = 347)
Year of study	
1960-1970	1 (0.3)
1971-1980	3 (1)
1981-1990	45 (13)
1991-2000	104 (30)
2001-2010	104 (30)
2011-2021	81 (23)
Location	
International	42 (12)
North America	230 (66)
Europe	49 (14)
Asia	13 (4)
South America	7 (2)
Africa	2 (0.6)
Depressive disorder	
Major	312 (90)
Persistent	10 (3)
Premenstrual dysphoric disorder	12 (4)
Mixed or other	11 (3)
Trial sponsorship	
Un-sponsored	51 (15)
Partial sponsorship	23 (7)
Fully sponsored	237 (68)
Not specified	36 (10)
Recruitment setting	
Inpatient	18 (5)
Outpatient	307 (89)
Mixed or not specified	23 (7)
No. of active treatment groups	
1	158 (46)
2	146 (42)
3	36 (10)
4	6 (2)
Trial length, wk	
<4	4 (1)
4-5	27 (8)
6-7	108 (31)
8-9	115 (33)
10-11	14 (4)
12	62 (18)
>12	17 (5)
Primary antidepressant class	
Atypical	45 (13)
Azapirone	2 (0.6)
Monoamine oxidase inhibitor	14 (4)
Norepinephrine reuptake inhibitor	12 (4)
Serotonin antagonist and reuptake inhibitor	9 (3)
Serotonin and norepinephrine reuptake inhibitor	62 (18)
Selective serotonin reuptake inhibitor	145 (42)
Tricyclic	47 (14)
Tetracyclic	11 (3)

Table 2. Logistic Regression Analyses of Factors Associated With Studies' Use of Placebo Run-in Periods

Factor	OR (95% CI)	Wald statistic	P value
Year of publication	0.94 (0.92-0.96)	28.19	<.001
Sample size	1.00 (0.99-1.00)	0.32	.58
Trial length, wk	0.99 (0.96-1.02)	0.53	.47
Study location			
International	1 [Reference]	NA	NA
North America	3.90 (1.87-8.15)	13.15	<.001
Europe	1.79 (0.73-4.37)	1.61	.21
Asia	1.25 (0.32-4.90)	0.11	.75
Pharmaceutical involvement			
No sponsorship	1 [Reference]	NA	NA
Partial sponsorship	1.68 (0.61-4.64)	1.01	.32
Fully sponsored	2.73 (1.43-5.20)	9.31	.002
Study setting			
Inpatient	1 [Reference]	NA	NA
Outpatient	0.84 (0.32-2.18)	0.13	.72
Depressive disorder			
Major	1 [Reference]	NA	NA
Persistent	1.53 (0.42-5.53)	0.42	.52
Premenstrual dysphoric disorder	2.04 (0.60-6.91)	1.31	.25
Mixed or other	0.58 (0.17-2.03)	0.64	.39
Antidepressant class			
Selective serotonin reuptake inhibitor	1 [Reference]	NA	NA
Serotonin and norepinephrine reuptake inhibitor	1.11 (0.60-2.05)	0.12	.73
Tricyclic	0.67 (0.35-1.31)	1.37	.24
Atypical	0.18 (0.08-0.39)	18.04	<.001
Monoamine oxidase inhibitor	0.70 (0.23-2.11)	0.40	.52
Other	0.44 (0.20-0.94)	4.52	.03
Risk of bias			
Low risk	1 [Reference]	NA	NA
Some concerns	1.03 (0.48-2.20)	0.01	.94
High risk	2.47 (1.24-4.94)	6.56	.01

Abbreviations: NA, not applicable; OR, odds ratio.

Criteria for Exclusion

In total, 98 of 174 trials (56%) reported criteria for excluding participants after the PRI period. Thirty trials excluded those who no longer met initial study eligibility criteria; 38 excluded participants with an improvement of 20% or more on the primary outcome; 16 studies used similar improvement criteria, although of a larger magnitude (eg, $\geq 25\%$ or $\geq 30\%$); and 9 excluded participants who were rated as “much” or “very much” improved on the Clinical Global Impression Scale. Sixteen trials excluded participants whose scores were below a particular severity cutoff, and 4 excluded participants who no longer met diagnostic criteria for the depressive disorder under study.

Reasons Provided

Of the 174 studies that used a single-blind PRI period, most (149 [86%]) reported using a PRI period without providing a reason. Of the 25 trials providing 1 or more reasons for the use of PRI periods, 22 reported that it was to identify and exclude placebo responders. Three studies reported that their PRI period was also used to wash out existing psychotropic medication,

whereas 2 reported that the PRI period was used to monitor and exclude participants who were not adherent to the study drug alongside placebo responders. Two studies reported that their PRI period would assist to maximize drug-placebo differences.

Number of Participants Excluded

Only 35 of 174 trials (20%) reported the number of participants excluded after the PRI period. Of the 35 trials, between 0 and 101 participants were excluded (for a total of 719 participants). The mean (SD) proportion of participants excluded was 9.5% (7.7%) (range, 0%-32%).

Association of Use of PRI Period With Study Outcomes

Data were available for 189 studies (53 091 participants). A total of 148 studies used the Hamilton Rating Scale for Depression,²⁴ 38 used the Montgomery-Åsberg Depression Rating Scale,²⁵ and 4 specified other primary outcomes. Table 3 displays the overall effect sizes and the results of subgroup analyses for the drug, placebo, and between-group effect sizes. Figure 2 displays these effect sizes and 95% CIs.

Table 3. Overall Effect Size and Subgroup Analyses Depending on Use of PRI Periods

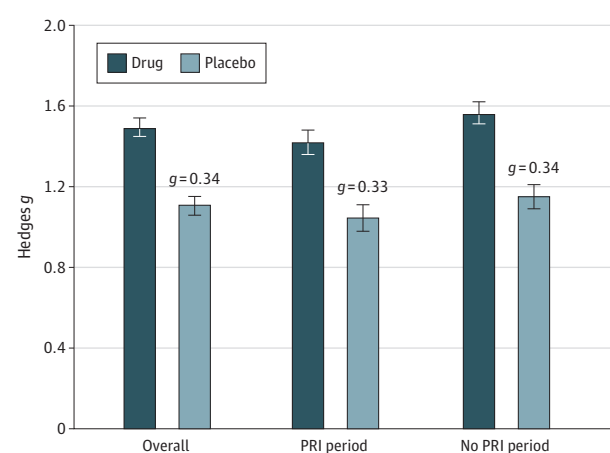
			Heterogeneity				
Characteristic	No.	Hedges <i>g</i> (95% CI)	Q value ^a	<i>I</i> ² statistic (95% CI)	τ ² Value	95% Prediction interval	<i>P</i> value ^b
Drug							
Overall	277	1.49 (1.44-1.53)	1755.64	84.27 (82.60-85.79)	0.10	0.86-2.11	.001
No PRI period	148	1.55 (1.49-1.61)	988.03	85.12 (82.94-87.02)	0.12	0.86-2.24	
PRI period	129	1.42 (1.36-1.48)	681.47	81.22 (78.04-83.94)	0.08	0.87-1.97	
Placebo							
Overall	189	1.10 (1.06-1.15)	1006.62	81.29 (78.86-83.67)	0.07	0.58-1.64	.02
No PRI period	106	1.15 (1.09-1.21)	548.57	80.68 (77.07-83.75)	0.07	0.61-1.69	
PRI period	83	1.05 (0.98-1.11)	439.18	81.56 (77.31-84.64)	0.07	0.53-1.57	
Drug vs placebo							
Overall	277	0.34 (0.31-0.37)	563.70	51.03 (43.76-57.37)	0.03	0.01-0.66	.92
No PRI period	148	0.34 (0.30-0.38)	322.57	54.43 (45.68-62.38)	0.03	0.00-0.68	
PRI period	129	0.33 (0.29-0.38)	243.76	47.49 (34.96-56.94)	0.02	0.05-0.61	

Abbreviation: PRI, placebo run-in.

^a In all cases, the Cochran Q value was statistically significant (all $P < .001$).

^b Corresponding to the comparison between PRI and non-PRI studies.

Figure 2. Within-Group and Between-Group Effect Sizes Calculated Using Study Outcomes Depending on Use of Placebo Run-in (PRI) Periods



Error bars indicate 95% CIs.

Drug Response

The overall within-group Hedges *g* for drug response was 1.49 (95% CI, 1.44-1.53; $I^2 = 84.27$; $\tau^2 = 0.10$). Results of subgroup analysis were significant, revealing a larger drug response size among studies that did not use a PRI period ($g = 1.55$ [95% CI, 1.49-1.61]) vs those that did ($g = 1.42$ [95% CI, 1.36-1.48]) ($P = .001$).

Placebo Response

The overall within-group Hedges *g* for placebo response was 1.10 (95% CI, 1.06-1.15; $I^2 = 82.29$; $\tau^2 = 0.07$). Results of subgroup analysis were significant, revealing a larger placebo response size among studies that did not use a PRI period ($g = 1.15$ [95% CI, 1.09-1.21]) vs those that did ($g = 1.05$ [95% CI, 0.98-1.11]) ($P = .02$).

Drug-Placebo Difference

The overall effect size between drug and placebo groups was 0.34 (95% CI, 0.31-0.37; $I^2 = 51.03$, $\tau^2 = 0.03$). Results of subgroup analysis were nonsignificant, with equivalent effect sizes being observed among studies that used a PRI period ($g = 0.33$ [95% CI, 0.29-0.38]) compared with those that did not use a PRI period ($g = 0.34$ [95% CI, 0.30-0.38]) ($P = .92$).

Secondary Analyses

Moderator analysis was repeated using response rate outcomes; between studies that used a PRI period and those that did not use a PRI period, there was no significant difference for the pooled drug response, placebo response, or drug-placebo difference. For studies that used a PRI period, the odds ratio was 1.89 (95% CI, 1.76-2.03), whereas for studies that did not use a PRI period, the odds ratio was 1.77 (95% CI, 1.65-1.89; $P = .18$) (eTable 2 in the Supplement). Within all subgroups examined, the pattern of findings remained unchanged (eTable 3 in the Supplement). Meta-regression analyses showed no significant association between the proportion of participants excluded and the drug ($\beta = 1.03$, SE = 0.85; $P = .22$) or placebo response ($\beta = -0.19$, SE = 1.23; $P = .88$). The association between the proportion of participants excluded and the drug-placebo difference was also not significant ($\beta = -1.39$, SE = 0.70; $P = .05$).

Risk of Bias and Publication Bias

See eTable 4 in the Supplement for individual studies' risk of bias and a description of results. The funnel plot indicated missing studies on the bottom left side of the mean effect (eFigure in the Supplement), which was supported by significant Egger test results. Trim-and-fill analysis suggested 57 missing studies, with an adjusted effect size of $g = 0.26$ (95% CI, 0.22-0.29).

Discussion

The use of single-blind PRI periods is common among RCTs of antidepressants, despite little evidence of their benefit. To our knowledge, this study is the first to detect an association between use of PRI periods and study outcomes. Our results show that the within-group antidepressant response differs depending on use of PRI periods, in which smaller within-group symptom improvements are observed in both the drug and placebo groups among trials using PRI periods. However, this reduction in effect size was small, and its magnitude was equivalent in both the drug and placebo groups. Consequently, the drug-placebo difference observed when comparing studies that used a PRI period and those that did not use a PRI period was unchanged ($g = 0.33$ vs 0.34 ; $P = .92$). These findings suggest that using PRI periods offers no advantages in terms of increasing sensitivity to detect drug-placebo differences in RCTs of antidepressants.

The likely explanation of our results is that excluding early placebo responders reduces the overall placebo response after randomization (Figure 2). This finding is partly consistent with original intuition but crucially neglects the observation that drug responses consist of drug-specific effects plus placebo responses (ie, placebo effects along with other nonspecific effects).²⁶ The use and reasoning behind PRI periods highlight an ongoing misconception that placebo responses do not contribute additively to symptom improvement within active treatment groups. Given that we identified that the drug response was smaller in studies that excluded early placebo responders, our results support this “additive model,” whereby active treatment responses consist of the placebo response plus the drug response. This finding is reassuring given that the additivity assumption is inherent in the design of double-blind, placebo-controlled RCTs, which isolate the drug effect by examining the difference between the drug and placebo response.²⁷

To our knowledge, the present study is the first synthesis of information about what occurs during the PRI period and what study features are associated with the use of PRI periods. Although numerous factors were associated with the use of PRI periods, the use of PRI periods has decreased over time. Their decreasing use may suggest a growing recognition that this approach is limited, which these results confirm. Despite this finding, a meaningful proportion (22%) of RCTs still used PRI periods in the last 10 years, suggesting that this practice remains common in some settings. Most studies (86%) did not provide their reason for conducting PRI periods, the criteria for exclusion, or further information about those excluded. Among studies that provided a rationale, it was most commonly to “remove placebo responders,” which further illustrates the misconceptions already discussed. If studies do continue to use PRI periods, better reporting around these practices is needed (including the rationale, criteria for exclusion, and number excluded).

Two studies in our systematic review and meta-analysis reported using a PRI period to exclude participants who demonstrated nonadherence with treatment. This reason-

ing is also reflected in some guidelines that support the use of PRI periods.²⁸ However, such a practice likely reduces the external validity of RCTs, given that nonadherence is inevitable within routine care. In addition, nonadherence is likely to occur across all phases of the RCT.²⁹ There may be important circumstances whereby investigators seek to understand the effect of adherence to treatment (eg, investigating dose-response relationships), and the use of a PRI period in these situations may be somewhat justified. However, for those seeking to investigate the effects of adherence to treatment, we recommend planned sensitivity analyses excluding participants who were nonadherent as a more straightforward approach.

Given that we collected data on the length of PRI periods and the number of participants excluded, we can estimate the cumulative cost associated with this practice. Assuming a mean length of 8.7 days across the 174 trials using PRI periods, this accumulates to more than 4 years within the RCT and 1090 years of participant time. Assuming a mean proportion of 9.5% of 45 841 participants excluded in all of the trials that used a PRI period, more than 4350 participants may have been excluded prior to randomization.

Placebo run-in periods also raise concerns about the external validity of RCTs of antidepressants. Most agree that the generalizability of RCTs of antidepressants is already limited owing to factors such as the prohibitive exclusion criteria adopted in trials and the intensive monitoring of study participants.³⁰ Along with these factors, the use of PRI periods may further limit the generalizability of results to the broader treatment-seeking population.

Finally, PRI periods raise ethical concerns about the deception of participants,⁹ who may be unaware they are receiving placebo prior to randomization. Deception practices in research are not inherently unethical in circumstances in which it is necessary to answer certain research questions, and such knowledge gains outweigh the ethical risks. However, the use of a PRI period does not appear to meet such standards. As a result, we strongly recommend that ethical review boards challenge the practice of using PRI periods on this basis.

Limitations

This systematic review and meta-analysis has some limitations; one was that data were not available from all trials, particularly earlier RCTs. Despite this, the available evidence suggests that the lack of effect of the use of PRI periods on study outcomes has not changed over time.³ In addition, high rates of heterogeneity were observed regarding both the overall effects and those within studies that did and did not use a PRI. While secondary analyses (eTable 3 in the Supplement) within subgroups showed lower heterogeneity, these overall results suggest that there remain sources of between-study variance that are not accounted for by our primary analyses.

We observed a large placebo response ($g = 1.05$ [95% CI, 0.98-1.11]) in studies using PRI periods, suggesting that PRI periods do not eliminate placebo responses from RCTs. There are

a few reasons why this may be the case. First, large placebo responses in studies that use PRI periods could indicate that the exclusion criteria in PRI phases are insufficient. Second, given that antidepressants can take 2 to 8 weeks to produce a benefit,³¹ it may be that many participants do not expect to improve within the PRI phase. Third, nonspecific therapeutic effects (eg, the therapeutic alliance) that result from trial participation may produce a benefit only after some time. In addition, there is no known information about the success of blinding during the PRI period and whether participants are truly “masked” to this practice while undertaking the trial. Future research should explore these possibilities. Fourth, although PRI periods do not achieve the desired goal of eliminating placebo responders from RCTs, there is a need for alternative research paradigms that may further control for placebo effects.

Conclusions

This systematic review and meta-analysis provides evidence to suggest that the use of PRI periods has no scientific basis in trials of antidepressants. At the same time, PRI periods carry numerous costs and risks. These findings suggest that the use of PRI periods should not continue in RCTs of antidepressants, and future research is needed to investigate the use of PRI periods in other settings and for other disorders. Although our results reveal that the use of PRI periods in RCTs of antidepressants carries no benefits, they also provide strong reassurance that eliminating this practice comes at no cost to trial outcomes. Our results highlight the importance of ongoing investigations into the conduct of RCTs and the risks associated with practices based merely on precedent or intuition.

ARTICLE INFORMATION

Accepted for Publication: September 13, 2021.

Published Online: November 10, 2021.
doi:10.1001/jamapsychiatry.2021.3204

Author Contributions: Dr Scott had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: All authors.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Scott.

Critical revision of the manuscript for important intellectual content: Sharpe, Quinn, Colagiuri.

Statistical analysis: Scott, Quinn, Colagiuri.

Obtained funding: Sharpe, Colagiuri.

Administrative, technical, or material support: Scott, Sharpe.

Supervision: Colagiuri.

Conflict of Interest Disclosures: Dr Sharpe reported receiving grants from the Australian Research Council during the conduct of the study. No other disclosures were reported.

Funding/Support: This research was supported by Australian Research Council Discovery Project grant DP150104026 awarded to Drs Colagiuri and Sharpe.

Role of the Funder/Sponsor: The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: Shannon Webb, PhD, University of Sydney School of Psychology, contributed to article screening. Tessa Rooney, BPsych Hons, University of Sydney School of Psychology, contributed to data extraction and risk of bias assessment. They were compensated for their contributions.

REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593-602. doi:10.1001/archpsyc.62.6.593

2. Luo Y, Kataoka Y, Ostinelli EG, Cipriani A, Furukawa TA. National prescription patterns of antidepressants in the treatment of adults with major depression in the US between 1996 and 2015: a population representative survey based analysis. *Front Psychiatry*. 2020;11:35. doi:10.3389/fpsyt.2020.00035
3. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11(1):33-43. doi:10.1038/npp.1994.63
4. Lee S, Walker JR, Jakul L, Sexton K. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? a meta-analytic evaluation. *Depress Anxiety*. 2004;19(1):10-19. doi:10.1002/da.10134
5. Hantsoo L, Ward-O'Brien D, Czarkowski KA, Gueorguieva R, Price LH, Epperson CN. A randomized, placebo-controlled, double-blind trial of sertraline for postpartum depression. *Psychopharmacology (Berl)*. 2014;231(5):939-948. doi:10.1007/s00213-013-3316-1
6. Raby WN, Rubin EA, Garawi F, et al. A randomized, double-blind, placebo-controlled trial of venlafaxine for the treatment of depressed cocaine-dependent patients. *Am J Addict*. 2014;23(1):68-75. doi:10.1111/j.1521-0391.2013.12065.x
7. Hedayati SS, Gregg LP, Carmody T, et al. Effect of sertraline on depressive symptoms in patients with chronic kidney disease without dialysis dependence: the CAST randomized clinical trial. *JAMA*. 2017;318(19):1876-1890. doi:10.1001/jama.2017.17131
8. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA*. 2002;287(14):1840-1847. doi:10.1001/jama.287.14.1840
9. Mann H. Deception in the single-blind run-in phase of clinical trials. *JRB*. 2007;29(2):14-17.
10. Moher D, Altman DG, Liberati A, Tetzlaff J. PRISMA statement. *Epidemiology*. 2011;22(1):128. doi:10.1097/EDE.0b013e3181fe7825
11. Zhou X, Teng T, Zhang Y, et al. Comparative efficacy and acceptability of antidepressants, psychotherapies, and their combination for acute treatment of children and adolescents with depressive disorder: a systematic review and

network meta-analysis. *Lancet Psychiatry*. 2020;7(7):581-601. doi:10.1016/S2215-0366(20)30137-1

12. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *J Educ Statistics*. 1981;6(2):107-128. doi:10.3102/10769986006002107

13. Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. *Epidemiol Psychiatr Sci*. 2017;26(4):364-368. doi:10.1017/S2045796016000809

14. Wilson EJ. Meta-analytic procedures for social research. *J Business Res*. 1994;29:163-165. doi:10.1016/0148-2963(94)90020-5

15. Higgins JP, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons; 2019. doi:10.1002/9781119536604

16. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558. doi:10.1002/sim.1186

17. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-560. doi:10.1136/bmj.327.7414.557

18. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007;335(7626):914-916. doi:10.1136/bmj.39343.408449.80

19. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to Meta-analysis*. John Wiley & Sons; 2021. doi:10.1002/9781119558378

20. Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017;8(1):5-18. doi:10.1002/jrsm.1230

21. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. doi:10.1136/bmj.l4898

22. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629

23. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*.

- 2000;56(2):455-463. doi:[10.1111/j.0006-341X.2000.00455.x](https://doi.org/10.1111/j.0006-341X.2000.00455.x)
24. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol*. 1967;6(4):278-296. doi:[10.1111/j.2044-8260.1967.tb00530.x](https://doi.org/10.1111/j.2044-8260.1967.tb00530.x)
25. Neumann NU, Schulte RM. *Montgomery-Åsberg Depression Rating Scale*. Perimed Fachbuch-Verlagsgesellschaft GmbH; 1989.
26. Evers AWM, Colloca L, Blease C, et al. Implications of placebo and nocebo effects for clinical practice: expert consensus. *Psychother Psychosom*. 2018;87(4):204-210. doi:[10.1159/000490354](https://doi.org/10.1159/000490354)
27. Kirsch I. Are drug and placebo effects in depression additive? *Biol Psychiatry*. 2000;47(8):733-735. doi:[10.1016/S0006-3223\(00\)00832-5](https://doi.org/10.1016/S0006-3223(00)00832-5)
28. Laursen DRT, Paludan-Müller AS, Hróbjartsson A. Randomized clinical trials with run-in periods: frequency, characteristics and reporting. *Clin Epidemiol*. 2019;11:169-184. doi:[10.2147/CLEP.S188752](https://doi.org/10.2147/CLEP.S188752)
29. Pampallona S, Bollini P, Tibaldi G, Kupelnick B, Munizza C. Patient adherence in the treatment of depression. *Br J Psychiatry*. 2002;180(2):104-109. doi:[10.1192/bjp.180.2.104](https://doi.org/10.1192/bjp.180.2.104)
30. Zimmerman M, Clark HL, Multach MD, Walsh E, Rosenstein LK, Gazarian D. Have treatment studies of depression become even less generalizable? a review of the inclusion and exclusion criteria used in placebo-controlled antidepressant efficacy trials published during the past 20 years. *Mayo Clin Proc*. 2015;90(9):1180-1186. doi:[10.1016/j.mayocp.2015.06.016](https://doi.org/10.1016/j.mayocp.2015.06.016)
31. Tylee A, Walters P. Onset of action of antidepressants. *BMJ*. 2007;334(7600):911-912. doi:[10.1136/bmj.39197.619190.80](https://doi.org/10.1136/bmj.39197.619190.80)