Chapter 4

Hiding negative antidepressant trials by pooling them: the pooled-trials publication bias

Ymkje Anna de Vries, Annelieke M. Roest,
Erick H. Turner, Peter de Jonge

Submitted

Abstract

Background: Previous studies on reporting bias generally examined whether trials were published in stand-alone publications. In this study we investigated whether pooled-trials publications constitute a specific form of reporting bias. We assessed whether negative trials were more likely to be exclusively published in pooled-trials publications than positive trials. In addition, we examined the research questions, individual trial results, and conclusions presented in these articles.

Methods: Data from a cohort of 74 randomized controlled trials of 12 antidepressants were extracted from an earlier publication and the corresponding Food and Drug Administration reviews. A systematic literature search was conducted to identify pooled-trials publications.

Results: We found 86 pooled-trials publications that reported results of 20 (83%) of 24 trials not published in stand-alone publications. Relative to positive trials, not-positive trials were 9.5 times more likely to be published exclusively in pooled-trials publications ($p<0.001$). Ten (12%) of 86 publications had as primary aim to present data on the trial's primary research question (drug efficacy compared to placebo). Only 3 publications, reporting on 3 (15%) trials, presented individual efficacy data for the primary research question. Additionally, only 3 (3%) of 86 pooled-trials publications had a negative conclusion.

Interpretation: Compared to positive trials, negative trials of antidepressants for depression were much more likely to be reported exclusively in pooled-trials publications. Pooled-trials publications flood the evidence base with often-redundant articles that, instead of addressing the original primary research question, present (positive) results on secondary questions. Therefore, pooled-trials publications inflate the apparent efficacy of antidepressants.

Introduction

The presence of reporting bias has been demonstrated in many medical fields [38, 182, 203]. An important form of reporting bias is study publication bias, which occurs when trials with positive results are more likely to be published than those with negative results [102].

In studies on reporting bias, trials that are published exclusively in pooled-trials publications, which pool data from two or more trials, are usually regarded as unpublished [19, 104] or incompletely published [204]. In contrast, some pharmaceutical companies have argued that these trials have actually been published and should be counted as such [205].

Although pooled-trials publications have been found to provide new information, they may be particularly susceptible to bias, for example because it is often unclear how trials were selected for inclusion [175]. In addition, pooled-trials publications often have a research question that differs from the original research question of the included trials. For example, they may focus on differential efficacy in various patient subgroups, leading to substantial redundancy and the suggestion that many of these articles may represent 'salami publications' [174].

Another potential problem with pooled-trials publications is that, in contrast to positive trials, negative trials may often be published exclusively in pooled-trials publications. A study examining trials for five antidepressants approved between 1989 and 1994 found that positive trials were usually reported in stand-alone publications, while negative trials were frequently "bundled" (often with positive trials) into pooled-trials publications [27]. Consequently, pooled-trials publications may actually further bias the published literature, rather than helping to provide transparent access to trial results.

A previous meta-analysis found that 31% of antidepressant trials for major depressive disorder (MDD) remained unpublished [19]. However, pooled-trials publications were excluded from this study. In the present study, we use the trials included in this meta-analysis to investigate whether the practice of pooling trials for publication constitutes a specific form of reporting bias.

Our first aim was to assess whether unpublished antidepressant trials were actually published in pooled-trials publications and to determine how frequently negative trials were published exclusively in pooled-trials publications compared to positive trials. Our second aim was to evaluate how often the research question of pooled-trials publications corresponded to the original primary research question of the included trials and how often these publications reported individual trial results for this primary outcome. Finally, our third aim was to assess how often pooled-trials publications reached positive conclusions about the trial drug.

Methods

Trials submitted to the FDA

Information on phase 2/3 clinical trial programs for 12 second-generation antidepressants (bupropion sustained release, citalopram, fluoxetine, paroxetine controlled release [CR], duloxetine, escitalopram, mirtazapine, nefazodone, paroxetine, sertraline, venlafaxine immediate release [IR], venlafaxine extended release [XR]) was extracted from an earlier publication by Turner and colleagues [19] and the Food and Drug Administration (FDA) reviews used in that study.

These programs included 74 randomized, double-blind, placebo-controlled trials investigating the short-term treatment of major depression. Because pharmaceutical companies must preregister trials they intend to conduct in support of an application of marketing approval with the FDA, FDA reviews can be used as a registry and results database [105].

Consistent with Turner and colleagues [19], who extracted the FDA's regulatory decision (i.e., whether the primary endpoint(s) were judged to be positive or not), 38 trials were considered positive and 36 trials were considered not-positive in the current study.

We retrieved the references of 50 journal articles that reported the results of trials registered with the FDA from Turner and colleagues [19]. One article presented the pooled results of two identically designed trials of paroxetine CR [189]. This article was regarded as a pooled-trials publication in the current study. In addition, we found a matching stand-alone publication [206] for one trial (UK-06) regarded as unpublished by Turner and colleagues [19].

Trials published in pooled-trials publications only

We assessed whether trials that were unpublished according to Turner and colleagues [19] (i.e. trials not published in stand-alone form) were published in pooled-trials publications. Pooled-trials publications were defined as publications in which the individual patient data of two or more trials were analyzed. This included publications described as individual patient data meta-analyses, but it did not include other meta-analyses (based on aggregate data).

A systematic literature search was conducted in PubMed, EMBASE and the Cochrane Central Register of Controlled Trials, restricted to articles in English, until February 10, 2015. The search strategy included the name of the drug, terms related to depression and "placebo". Terms were customized to the search strategies of each database; for example, when searching PubMed for relevant citalopram publications the search syntax was: citalopram [Title] AND depress* [Title/abstract] AND placebo.

After identifying the pooled-trials publications, matches for each trial were identified using the following parameters: drug name, active comparator (when applicable), dosage groups, sample sizes, trial duration, and names of investigators. We included only pooled-trials publications for which individual trials could be matched to FDA-registered trials.

From each publication, we extracted the primary research question and whether individual trial results were reported. Research questions were categorized as "primary efficacy", i.e. the research question of the pooled-trials publication was the same as the original trial's primary research question (efficacy of the drug compared to placebo), or "not primary efficacy". The second category consisted of publications on secondary efficacy outcomes (e.g. anxiety, sleep problems, efficacy compared to an active comparator), predictors of efficacy (e.g. efficacy in subgroups, baseline severity), and other efficacy or safety outcomes.

In addition, AR and YV classified each pooled-trials publication as positive, neutral, or negative, based on the abstract. Publications were judged to be positive when the abstract claimed that the antidepressant was more effective than placebo or an active comparator, equal in efficacy to an active comparator, safer or better tolerated than placebo or an active comparator, equal in safety/tolerability to placebo or an active comparator, or simply "safe" or "well-tolerated". Publications were judged to be neutral when the publication was primarily methodological in orientation, for example assessing differences between analytical approaches. Differences were resolved by consensus.

Statistical analysis

We examined whether not-positive trials were more likely to be published exclusively in a pooled-trials publication than positive trials. Because of small cell sizes, p values were obtained using Fisher's exact test. In addition, risk ratios and their 95% confidence intervals are reported. Analyses were performed using Stata software, version 13.1.

Results

Pooled-trials publications

Twenty-four of 74 FDA-registered antidepressant trials were not published in stand-alone publications. Of these, 20 (83.3%) were included in 86 pooled-trials publications reporting results on 10 antidepressants (see Table 4.1).

As shown in Figure 4.1, of the 36 trials judged not-positive by the FDA, 18 (50%) were exclusively published in pooled-trials publications, compared to 2 (5.3%) FDA positive trials. Consequently, all positive trials were published in some form (either stand-alone or

pooled), as were 89% of the not-positive trials. Compared to positive trials, not-positive trials were 9.5 times more likely to be published exclusively in pooled-trials publications (risk ratio: 9.50; 95% CI: 2.37-38.06; Fisher's exact p<0.001).

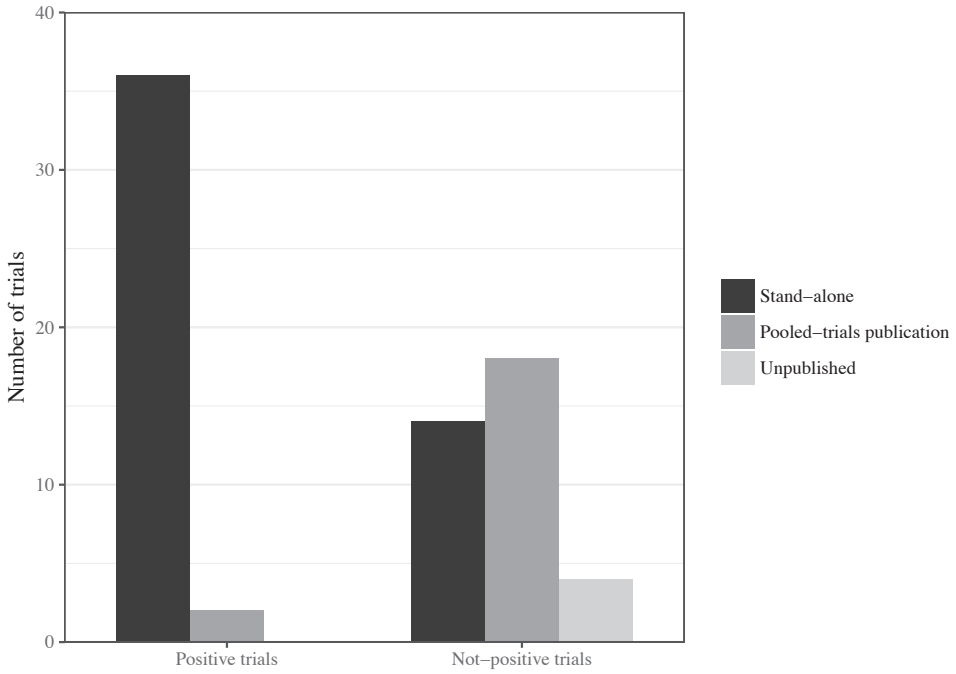

Figure 4.1: *Publication status of positive and not-positive FDA trials*

Research questions of trials published in pooled-trials publications and presentation of efficacy data

Ten out of 86 (11.6%) pooled-trials publications had the same research question as the included trial's original primary research question (drug efficacy compared to placebo) (Table 4.1). These 10 publications together included 7 (35%) of 20 trials published exclusively in pooled-trials publications. Only 3 (15%) of these publications presented individual efficacy data for the primary research question, reporting efficacy results for 3 not-positive trials.

Other pooled-trials publications reported on secondary efficacy outcomes (number of publications = 22; number of included trials = 9), predictors of efficacy (number of publications = 23; number of included trials = 8), other efficacy data (number of publications = 10; number of included trials = 5), or safety outcomes (number of publications = 21; number of included trials = 12) (Table 4.1).

Table 4.1: *Pooled-trials publications*

				Research question				
				Efficacy				
Drug	Trial	FDA	N	Primary	Secondary	Predictors	Other	Safety
Bupropion	205	NP	1	0 (0)	0	0	0	1
	212	NP	1	0 (0)	0	0	0	1
	Total		1	0 (0)	0	0	0	1
Citalopram	89306	NP	1	0 (0)	0	0	0	1
	Total		1	0 (0)	0	0	0	1
Duloxetine	HMAT-A	NP	38	4 (2)	6	12	4	12
	HMAQ-B	NP	31	2 (1)	6	12	1	10
	Total		38	4 (2)	6	12	4	12
Escitalopram	MD-02	NP	16	0 (0)	6	4	4	2
	Total		16	0 (0)	6	4	4	2
Mirtazapine	003-020	P	3	0 (0)	3	0	0	0
	003-021	NP	3	0 (0)	3	0	0	0
	003-003	NP	1	0 (0)	1	0	0	0
	003-008	NP	1	0 (0)	1	0	0	0
	Total		3	0 (0)	3	0	0	0
Nefazodone	7	NP	0	0 (0)	0	0	0	0
	004A	NP	0	0 (0)	0	0	0	0
	Total		0	0 (0)	0	0	0	0
Paroxetine	01-001	NP	2	0 (0)	0	0	0	2
(IR and CR)	03-003	NP	4	3 (0)	0	0	1	0
	07	NP	0	0 (0)	0	0	0	0
	09	NP	2	0 (0)	0	0	0	2
	UK-12	NP	2	0 (0)	0	0	0	2
	448	NP	2	1 (0)	0	1	0	0
	449	P	2	1 (0)	0	1	0	0
	Total		8	4 (0)	0	1	1	2
Sertraline	315	NP	1	0 (0)	0	1	0	0
	101	NP	1	0 (0)	0	0	0	1
	310	NP	0	0 (0)	0	0	0	0
	Total		2	0 (0)	0	1	0	1
Venlafaxine	303	NP	7	2 (1)	1	2	0	2
(IR and XR)	367	NP	12	1 (0)	6	3	1	1
	Total		17	2 (1)	7	5	1	2
All drugs	Total		86	10 (3)	22	23	10	21

The FDA column indicates the Food and Drug Administration decision (P: positive; NP: not-positive). N indicates the number of pooled-trials publications. For pooled-trials publications examining primary efficacy, the number in parentheses indicates the number of such publications that included individual trial results. CR: controlled release; IR: immediate release; XR: extended release.

Only three pooled-trials publications (3%) reported a negative conclusion (Table 4.2 in the Appendix). One of these publications examined the general safety profile of duloxetine and two examined the risk of suicidality with paroxetine treatment. All three concluded that the drug was associated with more adverse events than placebo. An additional 8 pooled trials publications (9%) had neutral conclusions (predictors of efficacy = 5; other efficacy = 3) while the remaining 75 (87%) were positively framed.

Discussion

To our knowledge, this study is the first to show that pooled-trials publication bias constitutes a specific form of reporting bias, which distorts the apparent efficacy of antidepressants. Although 24 of 74 antidepressant trials were not published in stand-alone articles, we showed that only four trials were completely unpublished, while the other 20 trials were included in pooled-trials publications. Trials lacking positive results were approximately 10 times more likely to be exclusively published in pooled-trials publications than trials with positive results.

Importantly, only 12% of all pooled-trials publications (including 7 (35%) of 20 trials) examined the original primary research question (efficacy of drug compared to placebo). Furthermore, individual trial results for this primary research question were provided in a pooled-trials publication for only 15% of trials. Finally, only 3% of pooled-trials publications had a negative conclusion. Therefore, although these trials have technically been published, the negative efficacy results are obscured, thus inflating the drugs' apparent efficacy.

For some drugs, particularly duloxetine, the number of pooled-trials publications was very high. Trials HMAT-A and HMAQ-B (which both had a negative result on the primary efficacy outcome) were included in 38 and 31 publications, respectively. Our study thus replicates a prior report on the 'salami slicing' of duloxetine trials [174].

Additionally, we found many pooled-trials publications for venlafaxine (17 publications for immediate- and extended-release combined) and escitalopram (16 publications). Several of these publications seemed redundant. For instance, three pooled-trials publications compared the efficacy of escitalopram to citalopram; three examined the effects of age and gender on the efficacy of venlafaxine; and eight compared the efficacy of venlafaxine to selective serotonin reuptake inhibitors.

It is noteworthy that duloxetine (approved in 2004) and escitalopram (approved in 2002) are the two newest antidepressants in our study, although venlafaxine ER was approved somewhat earlier in 1997. This suggests that the practice of pooling trials in many separate publications is a relatively new one, perhaps developing concurrently with physicians' growing skepticism of advertising and sales representatives and greater trust in peer-reviewed publications [207].

Consistent with this, a previous study examining five antidepressants approved between 1989 and 1994 identified at most six pooled-trials publications for a single drug [27]. In light of the growing concern that the medical literature may function as a marketing tool for pharmaceutical companies [207, 208, 209, 210], pooled-trials publications, consisting primarily of secondary analyses of previously collected data, may provide an easy and inexpensive way to keep a drug 'in the spotlight' and enhance its sales.

Others have also noted that antidepressant meta-analyses (including pooled-trials publications and aggregate data meta-analyses) are massively produced, frequently have some kind of industry involvement, and almost never include any negative statements if one or more authors are industry employees [211]. In our study, 76 (88%) of 86 pooled-trials publications had at least one author who was employed by a pharmaceutical company and, as noted, only three publications had negative conclusions. In all three cases, the publication concluded that the antidepressant was associated with more adverse events than placebo, a finding that is not unexpected.

A significant proportion of pooled-trials publications examined safety and tolerability. Because many adverse events occur infrequently, individual trials often lack sufficient statistical power for signal detection; pooling trials addresses this issue. The link between antidepressants and suicidality, for instance, was convincingly established only by pooling trials across indications and drugs [190, 191].

However, pooled-trials publications can also mislead. For instance, bupropion SR was only approved at dosages of 300 – 400 mg/day, but a pooled-trials publication assessing its safety and tolerability pooled all dosage groups (50 – 400 mg/day) [212]. Since adverse events are often dose-dependent, this publication is likely to paint an overly optimistic picture of the safety and tolerability of bupropion SR. Furthermore, there is ongoing concern that meta-analyses of harm outcomes may be particularly threatened by selective outcome reporting [196]. For non-systematic pooled analyses, this concern is further increased by the possibility of selective inclusion of trials [175].

Limitations

A limitation of the current study is that we may have missed some pooled-trials publications that actually did include FDA-registered trials, because some publications provided too little information to allow matching. However, this would not decrease the impact of pooled-trials publication bias, since individual trial results are never included in these publications.

A second limitation is that we did not count pooled-trials publications that did report individual trial results for the original primary outcome but focused on a secondary research question, as we felt that it was unlikely that these results would be found by researchers or clinicians interested in the primary outcome. However, such publications

were few in number, so the impact of this is minor.

A final limitation of this study is that we assessed the presence of pooled-trials publication bias in a narrow field, namely antidepressant trials for MDD. Nevertheless, it would not be surprising to find such bias in other fields of medicine, because reporting bias has been found to occur throughout psychiatry [20, 33, 176], medicine [38, 182, 203], and science in general [34].

Conclusions

Although meta-analyses on reporting bias have been criticized for their decision to exclude pooled-trials publications, our study shows that these publications are biased toward positive conclusions. As these publications rarely include individual trial results, they appear to serve primarily to heighten the (positive) visibility of a drug, rather than to clearly and transparently report negative trial results.

Journal editors could request that pooled-trials publications also present individual trial results and reference articles that present the primary efficacy results for all included trials. Additionally, editors, peer reviewers, and readers should be aware of the potential for bias and redundancy with pooled-trials publications [174] and perhaps ask whether they enhance or merely distort and bloat the evidence base.

In summary, the practice of pooling trials increases the apparent efficacy of antidepressants by flooding the literature with publications that highlight positive results and obscure negative results. Together with study publication bias, selective outcome reporting, and spin, pooled-trials publication bias is a form of reporting bias that should be taken into account in future research.

Appendix

Table 4.2: *Supplemental table of studies*

Publication	Included trials	Research question	Abstract
Bupropion			
Settle (1999) [212]	205, 212	Safety	Positive
Citalopram			
Pedersen (2006) [213]	89306	Safety	Positive
Duloxetine			
Bailey (2006) [214]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Bech (2006) [215]	HMAT-A	Other efficacy	Positive
Brecht (2008) [216]	HMAT-A, HMAQ-B	Secondary efficacy	Positive
Brunton (2010) [217]	HMAT-A, HMAQ-B	Safety	Negative
Cookson (2006) [218]	HMAT-A, HMAQ-B	Secondary efficacy	Positive
Delgado (2005) [219]	HMAT-A	Safety	Positive
Dodd (2014) [220]	HMAT-A, HMAQ-B	Predictors of efficacy	Neutral
Dunner (2003) [221]	HMAT-A, HMAQ-B	Secondary efficacy	Positive
Dunner (2005) [222]	HMAT-A, HMAQ-B	Safety	Positive
Fishbain (2008) [223]	HMAT-A	Other efficacy	Positive
Greist (2004) [224]	HMAT-A, HMAQ-B	Safety	Positive
Gueorguieva (2011) [225]	HMAT-A, HMAQ-B	Primary efficacy	Positive
Harada (2015) [226]	HMAT-A	Primary efficacy	Positive
Hudson (2005) [227]	HMAT-A, HMAQ-B	Safety	Positive
Kornstein (2006) [228]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Lewis-Fernandez (2006) [229]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Mallinckrodt (2003) [230]	HMAT-A, HMAQ-B	Secondary efficacy	Positive
Mallinckrodt (2004) [231]	HMAT-A, HMAQ-B	Other efficacy	Neutral
Mallinckrodt (2005) [232]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Mallinckrodt (2006) [233]	HMAT-A	Primary efficacy	Positive
Mallinckrodt (2008) [234]	HMAT-A, HMAQ-B	Secondary efficacy	Positive
Nelson (2005) [235]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Nelson (2006) [236]	HMAT-A	Safety	Positive
Nelson (2010) [237]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Nelson (2011) [238]	HMAT-A, HMAQ-B	Predictors of efficacy	Neutral
Nelson (2013) [239]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Nemeroff (2002) [240]	HMAT-A, HMAQ-B	Primary efficacy	Positive
Perahia (2005) [241]	HMAT-A, HMAQ-B	Safety	Positive
Perahia (2006) [242]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Pritchett (2007) [243]	HMAT-A	Other efficacy	Positive
Schacht (2014) [244]	HMAT-A, HMAQ-B	Predictors of efficacy	Neutral
Shelton (2007) [245]	HMAT-A, HMAQ-B	Predictors of efficacy	Positive
Stewart (2006) [246]	HMAT-A, HMAQ-B	Safety	Positive
Thase (2005) [247]	HMAT-A, HMAQ-B	Safety	Positive
Thase (2007) [248]	HMAT-A, HMAQ-B	Secondary efficacy	Positive

continued

Table 4.2: *Supplemental table of studies*

Publication	Included trials	Research question	Abstract
Viktrup (2004) [249]	HMAT-A, HMAQ-B	Safety	Positive
Wernicke (2007) [250]	HMAT-A, HMAQ-B	Safety	Positive
Wise (2006) [251]	HMAT-A, HMAQ-B	Safety	Positive
Escitalopram			
Baldwin (2007) [252]	SCT-MD-02	Safety	Positive
Baldwin (2009) [94]	SCT-MD-02	Other efficacy	Neutral
Bandelow (2007) [253]	SCT-MD-02	Secondary efficacy	Positive
Demyttenaere (2008) [254]	SCT-MD-02	Secondary efficacy	Positive
Demyttenaere (2011) [255]	SCT-MD-02	Other efficacy	Neutral
Gorman (2002) [256]	SCT-MD-02	Secondary efficacy	Positive
Kennedy (2009) [257]	SCT-MD-02	Secondary efficacy	Positive
Kilts (2009) [258]	SCT-MD-02	Predictors of efficacy	Positive
Lader (2005) [259]	SCT-MD-02	Secondary efficacy	Positive
Lam (2006) [260]	SCT-MD-02	Predictors of efficacy	Positive
Llorca (2005) [261]	SCT-MD-02	Other efficacy	Positive
Papakostas (2011) [262]	SCT-MD-02	Predictors of efficacy	Positive
Pedersen (2005) [263]	SCT-MD-02	Safety	Positive
Stein (2006) [264]	SCT-MD-02	Other efficacy	Neutral
Stein (2011) [265]	SCT-MD-02	Secondary efficacy	Positive
Wade (2006) [266]	SCT-MD-02	Predictors of efficacy	Neutral
Mirtazapine			
Bech (2001) [267]	003-020, 003-021	Secondary efficacy	Positive
Fawcett (1998) [268]	003-003, 003-008, 003-020, 003-021	Secondary efficacy	Positive
Stahl (1997) [269]	003-020, 003-021	Secondary efficacy	Positive
Paroxetine IR and CR			
Carpenter (2011) [270]	448, 449, UK-12, 01, 09	Safety	Negative
Dunbar (1991) [271]	03-003	Primary efficacy	Positive
Feighner (1992) [272]	03-003	Primary efficacy	Positive
Feighner (1993) [273]	03-003	Primary efficacy	Positive
Kraus (2010) [274]	448, 449, UK-12, 01, 09	Safety	Negative
Montgomery (1992) [275]	03-003	Other efficacy	Positive
Dunner (2005) [276]	448, 449	Predictors of efficacy	Positive
Golden (2002) [189]	448, 449	Primary efficacy	Positive
Sertraline			
Berti (1995) [277]	315	Predictors of efficacy	Positive
Fisch (1992) [278]	101	Safety	Positive
Venlafaxine IR and XR			
Danjou (1995) [279]	303	Safety	Positive
Entsuah (1995a) [280]	303	Predictors of efficacy	Positive

continued

Table 4.2: *Supplemental table of studies*

Publication	Included trials	Research question	Abstract
Entsuah (1995b) [281]	303	Predictors of efficacy	Positive
Entsuah (2001) [282]	367	Predictors of efficacy	Positive
Entsuah (2002) [283]	367	Other efficacy	Positive
Gibbons (2012a) [76]	303, 367	Primary efficacy	Positive
Gibbons (2012b) [284]	303, 367	Safety	Positive
Mallick (2003) [285]	367	Secondary efficacy	Positive
Mendlewicz (1995) [286]	303	Primary efficacy	Positive
Nemeroff (2008) [287]	367	Secondary efficacy	Positive
Rudolph, (1998) [288]	303	Secondary efficacy	Positive
Shelton (2005) [289]	367	Secondary efficacy	Positive
Silverstone (2002) [290]	367	Predictors of efficacy	Positive
Stahl (2002) [291]	367	Secondary efficacy	Positive
Thase (2001) [292]	367	Secondary efficacy	Positive
Thase (2005) [293]	367	Predictors of efficacy	Positive
Trivedi (2004) [294]	367	Secondary efficacy	Positive