

# A Meta-Analysis Comparing Open-Label versus Placebo-Controlled Clinical Trials for Aripiprazole Augmentation in the Treatment of Major Depressive Disorder: Lessons and Promises

Chi-Un Pae<sup>1,2</sup> ✉, Ho-Jun Seo<sup>1</sup>, Boung Chul Lee<sup>3</sup>, Jeong-Ho Seok<sup>4</sup>, Hong Jin Jeon<sup>5</sup>, Jong-Woo Paik<sup>6</sup>, Kyung-Phil Kwak<sup>7</sup>, Byung-Joo Ham<sup>8</sup>, Changsu Han<sup>8</sup>, and Soo-Jung Lee<sup>1</sup>

<sup>1</sup>Department of Psychiatry, The Catholic University of Korea College of Medicine, Seoul, Republic of Korea

<sup>2</sup>Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

<sup>3</sup>Department of Neuropsychiatry, College of Medicine, Hallym University, Seoul, Republic of Korea

<sup>4</sup>Department of Psychiatry, Yonsei University College of Medicine, Seoul, Republic of Korea

<sup>5</sup>Department of Psychiatry, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

<sup>6</sup>Department of Psychiatry, Kyunghee University College of Medicine, Seoul, Republic of Korea

<sup>7</sup>Department of Neuropsychiatry, School of Medicine, Dongguk University, Gyeongju, Republic of Korea

<sup>8</sup>Department of Psychiatry, College of Medicine, Korea University, Seoul, Republic of Korea

**Objective** The present study is to provide whether open-label studies (OLS) may properly foresee the efficacy of randomized, placebo-controlled trials (RCTs) using OLSs and RCTs data for aripiprazole in the treatment of MDD, with the use of meta-analysis approach.

**Methods** A search of the studies used the key terms “depression and aripiprazole” from the databases of PubMed/PsychInfo from Jan 2005 through July 2013. The data were selected and verified for publication in English-based peer-reviewed journals based on rigorous inclusion criteria. Extracted data were delivered into and run by the Comprehensive Meta Analysis program v2.

**Results** The pooled SMDs for the primary efficacy measure was statistically significant, pointing out the significant reduction of depressive symptoms after aripiprazole augmentation (AA) to current antidepressant treatment in OLSs (pooled SMD=-2.114, z=-9.625, p<0.001); similar results were also found in RCTs (pooled SMD=-2.202, z=-6.862, p<0.001). The meta-regression analysis revealed no influence of the study design for treatment outcome.

**Conclusion** There was no difference in the treatment effects of aripiprazole as an augmentation therapy in both OLSs and RCTs, indicating that open-label design may be a potentially useful predictor for treatment outcomes of controlled-clinical trials. The proper conduction of OLSs may provide informative, useful and preliminary clinical data and factors to be involved in controlled-clinical trials, by which we may have better understanding on the role of AA (e.g., dosing issues, proper duration of treatment, specific population for AA) implicated in the treatment of MDD in clinical practice.

**Psychiatry Investig 2014;11(4):371-379**

**Key Words** Open-label study, Randomized-controlled clinical trials, Aripiprazole, Augmentation, Depression.

## INTRODUCTION

Major depressive disorder (MDD) is a common and debilitating illness resulting in functional disability, decrease in quality of life, and increase in healthcare costs.<sup>1</sup> MDD is also

the third leading cause of moderate to severe disability and of disease burden worldwide.<sup>2</sup> A large array of different class of antidepressants are currently available, however, there has been a controversy regarding class and individual differences in efficacy for treatment of MDD patients among various antidepressants,<sup>3,4</sup> although potentially differential points may exist between benefits, acceptability, and acquisition cost.<sup>5,6</sup> Despite of sufficient availability of antidepressants with different classes to date, only 30% of patients treated with first antidepressant treatment show a symptomatic remission and suffer significant functional impairment.<sup>7,8</sup> Such inadequate antidepressant efficacy results in suffering from significant residual symptoms, functional incapacity, increased utilisation of med-

**Received:** September 2, 2013 **Revised:** September 20, 2013

**Accepted:** September 20, 2013 **Available online:** October 20, 2014

✉ **Correspondence:** Chi-Un Pae, MD, PhD

Department of Psychiatry, Bucheon St. Mary's Hospital, The Catholic University of Korea College of Medicine, 327 Sosa-ro, Wonmi-gu, Bucheon 420-717, Republic of Korea

**Tel:** +82-32-340-7067, **Fax:** +82-32-340-2255, **E-mail:** pae@catholic.ac.kr

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ical services, and frequent recurrence and relapse.<sup>9-11</sup>

Most guidelines have suggested that such non- or partial responders should be considered for a switch, combination or augmentation of treatment.<sup>3,12-14</sup> Among such treatment strategies, augmentation is the use of a non-antidepressant agent to broaden or enhance the therapeutic effectiveness of an antidepressant by affecting different neurotransmitter systems combining agents with different mechanisms of action and/or indications. Traditional augmentation agents, lithium and triiodothyronine (T3), as well as buspirone, dopamine agonists, and stimulants have been commonly used for such patient population with limited evidence (e.g., weak efficacy, tolerability issues, shortage of controlled clinical trials and official approval issues).<sup>15</sup>

Atypical antipsychotics such as olanzapine, quetiapine extended release (XR), and aripiprazole have clearly demonstrated efficacy as an augmentation agent for MDD patients through a number of small-scale, open-label studies or randomized, placebo-controlled clinical trials (RCTs). Among such AAs, aripiprazole has been the first approved by the U.S. FDA as an augmentation therapy to antidepressants for treating MDD in November 2007.

Industry-sponsored registration studies involving RCTs have been a gold-standard and cornerstone of modern research concerning medical therapies for human to prove certain medications' efficacy and safety, however, such rigorous research design usually excludes those with substantial medical comorbidity and do not allow concomitant medication treatments, which is commonly encountered in routine practice. Hence, RCTs are also criticized in lacking external validity and it has been asserted that these efficacy studies do not provide sufficient information to clinicians in real-world settings.<sup>9</sup> In addition, they have limitations in terms of the huge study conduction cost, loss of professional autonomy, inflation of rating scale scores prior to randomization, patient expectations of improvement, unexpected influence by a sponsorship, lack of staff and training, observer biases among investigators, a longer time- frame for conduction and difficulties with randomisation or recruitment, e.t.c. Hence, the usual and easy way to go with such further step of researches is open-label and prospectively designed studies to investigations to explore new medication (or new acquisition of indication), to acquire supplemental information about a certain medication, and to facilitate exploratory and hypothesis-generating studies.

However, specific information on such open-trials for future RCTs or addressing certain clinical issues not to be explained or investigated in RCTs is still very lacking today. Hence, the objective of the present study is to provide whether open-label studies (OLS) may properly foresee the efficacy of RCTs using OLSs and RCTs for aripiprazole augmentation (AA) in the

treatment of MDD, with the use of meta-analysis based on published OLSs and RCTs.

## METHODS

### Source of data

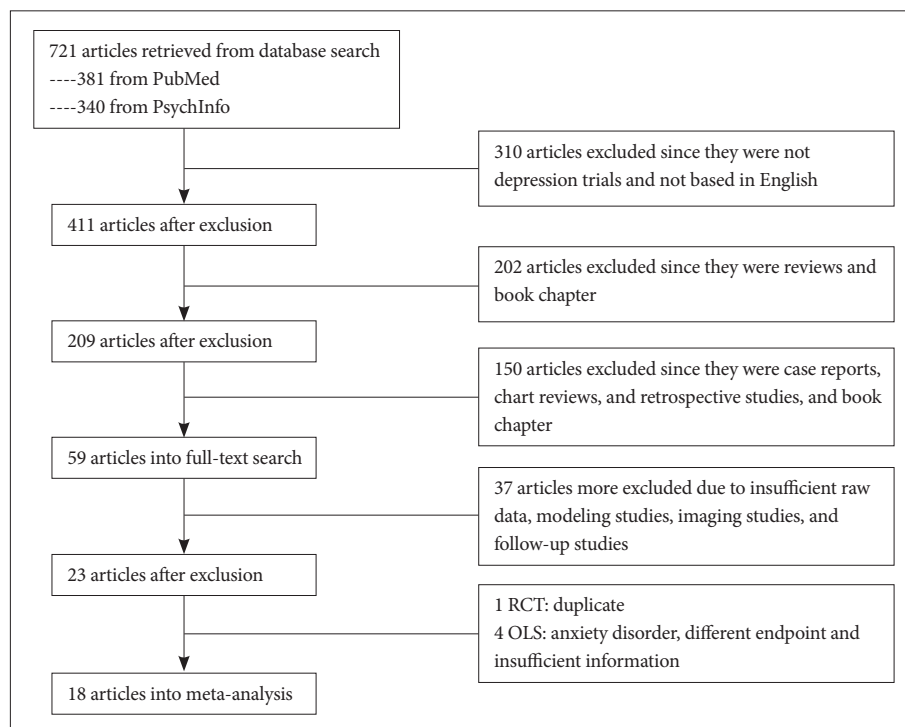
A search of the studies used the key terms "depression and aripiprazole" from the databases of PubMed/PsychInfo from Jan 2005 through July 2013. The data were verified for publication in English-based peer-reviewed journals. We also used reference lists from identified articles and reviews to find additional studies. Abstracts identified during literature search were screened by two review authors independently. Potentially eligible papers were read in full by two review authors to determine whether they met the eligibility criteria. Disagreements were discussed with a third review author until consensus was reached. Study selection was handled first by two of the authors, H.J. S and C. H. then independently reassessed by C.U. P.

### Inclusion criteria

We included studies of AA in the treatment of MDD. The studies to be included in the present meta-analysis should fulfill following criteria: 1) prospective design with commonly adopted in clinical trials as efficacy measures such as Hamilton Depression Rating scale (HAMD), e.t.c.; 2) at least one or more follow-up visit; 3) study duration at least for 4 weeks; 4) OLS or RCT for aripiprazole; and 5) published in English-based peer-reviewed journal. There were no requirements or restrictions in search of data for date. However, post-hoc analysis of RCT, redundant studies, case-report or letter-to-the-editor concerning interesting or rare case illustration in the use of aripiprazole for treating MDD were not included. If necessary, the study authors were contacted for additional information. Figure 1 summarizes the disposition of the included studies for the present meta-analysis.

### Data extraction

Data on patients (e.g. age and gender), design (e.g. OL, randomization, allocation concealment), duration of treatment (weeks), name of antidepressant, range of aripiprazole doses, sample size, type of primary outcome measure for efficacy, baseline and endpoint mean and standard deviation (SD) of primary efficacy measures, remission and response rates based on the criteria as each study defined. The baseline and endpoint mean and SD were replaced by baseline ones or taken manually calculated from visual availability or taken from standard errors, confidence intervals (CIs) or t-values if such parameters were not available from the original study. The primary efficacy measure was HAMD-17, Montgomery-Åsberg



**Figure 1.** The disposition of selected studies for the meta-analysis. RCT: randomized placebo-controlled trial, OLS: open-label study.

Depression Rating Scale (MADRS), Quick Inventory of Depressive Symptomatology-16 items (QIDS-16) in the studies included in the present meta-analysis.

## Data analysis

### Primary efficacy measure

The primary efficacy measure was the mean change from baseline to endpoint in the total scores on HAD-17, MADRS, or QIDS which are the most frequently used rating scales in the OLS and RCTs included in the present meta-analysis. They have been founded to be highly and significantly correlated in a number of researches based on a direct comparison of HAMD-17 and MADRS scores in a rather homogeneous sample.<sup>16</sup>

### Safety and tolerability

Comparisons of safety and tolerability measures between OLSs and RCTs were not done due to very limited and inconsistency of reported measures.

### Effect size

The effect size for the primary efficacy measures in each study were presented as standardized mean differences (SMDs) with 95% CI as they were continuous parameters. Cohen's classification was used to evaluate the magnitude of the overall effect size with 1) SMD=0.2 to 0.5: small; 2) SMD=0.5 to 0.8: medium, and 3) SMD >0.8: large effect sizes. The calculation of SMDs were based on followings: 1) endpoint mean

primary efficacy score minus baseline primary efficacy score/pooled SD of the total treatment groups or 2) endpoint mean primary efficacy score of the active drug group minus baseline primary efficacy score of the placebo group/pooled SD.

### Statistical model

The random effects model of meta-analysis was applied for the analysis since it allows more balance than those under the fixed effects model because the smaller studies get more weight and the larger studies get less weight as well as allowing for sampling variability with and between studies under such model. In general, a random effects model is used to combine subgroups and yield the overall effect. The study-to-study variance (tau-squared) is assumed to be the same for all subgroups—this value is computed within subgroups and then pooled across subgroups.

### Heterogeneity and sensitivity analysis

The heterogeneity between studies was analyzed using the  $I^2$  statistics, a measure of how much variance between studies can be attributed to differences between studies rather than chance. The magnitude of considerable heterogeneity is usually  $I^2=75-100\%$ . To test the robustness of significant results, sensitivity analyses were conducted for studies with high versus low risk of bias. If statistical heterogeneity was present in the respective meta-analysis, subgroup and sensitivity analyses were also used to explore possible reasons for heterogeneity: judgment on whether one study has a huge impact on the

overall estimate or underlying influence attributable to the overall estimate. To do sensitivity analysis for each study in two groups (OLS and RCT), the pooled estimate was repeatedly calculated and analyzed with omission of one study at a time.

### Publication bias

The Egger test was also used for detection of publication bias to assess the bias associated with the greater likelihood of more publication of positive studies than negative studies. We adopted the method of Egger since the Egger's linear regression method quantifies the bias captured by the funnel plot using the actual values of the effect sizes and their precision, while Begg and Mazumdar's test uses ranks.

### Meta-regression

Meta-regression was performed to test the effect of study design (OLS vs. RCT) as an independent parameter on the mean change in the primary efficacy measure. In this method, usually a weighted logistic regression of the 2k cases per study is fit where k is the number of study arms, and the weight is the number of patients who have or do not have the outcome respectively.<sup>17</sup> Meta-regression is a sophisticated analytic approach method merging meta-analytic and linear regression principles. It aims to explore whether a linear relationship exists between an outcome measure and on or more covariates.

The associations found in a meta-regression should be considered hypothesis generating and not regarded as proof of causality.<sup>17</sup>

### Software program for meta-analysis

All directly extracted or computed data from the studies included were entered into Comprehensive Meta Analysis version 2.0 (CMA v2, Englewood, NJ, USA) to complete meta-analysis with data synthesis and then analyzed.

## RESULTS

### Demographics

With the search term and condition, we identified 721 articles in the PubMed and PsychInfo database. Seven hundred and three papers were excluded due to ineligibility, giving. The total number of subjects was 990, of which 627 was from RCTs and 363 from OLSs. Among the subject male was 317 (197 from RCTs and 137 from OLSs, 33.7%). The mean subject number of OLSs<sup>18-30</sup> and RCTs<sup>31-35</sup> were 27 and 125.4, respectively. The mean ages of OLSs and RCTs were 48.0 (11.5) and 43.6 (11.0) years, respectively. The mean duration of trials was 9.5 and 7.2 weeks in OLSs and RCTs, respectively. Table 1 depicts a summary of the studies included in the present meta-analysis.

**Table 1.** The characteristics of the open-label studies and randomized-controlled clinical trials

	Duration (weeks)	Number of subjects (n)	Male (n)	Age (years, mean and SD)	Primary efficacy rating scale	ARP dose (mg/d)
Chen et al. <sup>18</sup>	4	9	5	38.3 (12.2)	HAMD-17	4.2
Fabrazzo et al. <sup>19</sup>	24	35	14	38.8 (11.5)	HAMD-21	5
Sheffrin et al. <sup>20</sup>	12	24	10	73.9 (6.6)	HAMD-17	9.0
Matthews et al. <sup>21</sup>	7	16	10	41.8 (12.9)	HAMD-17	14.4
Hellerstein et al. <sup>22</sup>	12	15	5	46.1 (13.0)	HAMD-24	5–30
Schule et al. <sup>23</sup>	4	40	13	44.7 (12.6)	HAMD-21	15
Rutherford et al. <sup>24</sup>	6	20	5	63.0 (9.6)	HAMD-24	5–15
Patkar et al. <sup>25</sup>	6	10	4	44.9 (12.2)	HAMD-17	13.2
Papakostas et al. <sup>26</sup>	8	12	4	46.6 (11.3)	HAMD-17	≤30 (mean maximum dose=22.5)
Simon et al. <sup>27</sup>	8	15	6	44.1 (NA)	HAMD-17	2.5–10 (66.7%=10)
Nierenberg et al. <sup>28</sup>	12	50	19	43.0 (11.0)	QIDS-C/S-16	11 at exit
Pae et al. 2013 <sup>29</sup>	12	104	37	50.2 (13.1)	MADRS	6.6 at exit
Pae et al. 2007 <sup>30</sup>	8	13	5	43.3 (11.9)	HAMD-17	10.8
Berman et al. 2007 <sup>31</sup>	6	182	70	45.6 (11.3)	MADRS	11.8 at exit
Berman et al. 2009 <sup>32</sup>	6	177	39	46.5 (10.6)	MADRS	10.7 at exit
Marcus et al. <sup>33</sup>	6	191	65	44.6 (11.0)	MADRS	11.0 at exit
Fava et al. <sup>34</sup>	8	56	19	45.4 (10.4)	MADRS	2 or 5
Lin et al. <sup>35</sup>	10	21	4	35.9 (10.4)	HAMD-17	2.5

HAMD: Hamilton Depression Rating Scale, MADRS: Montgomery-Åsberg Depression Rating Scale, QIDS-C/S-16: 16-item Quick Inventory of Depressive Symptomatology Clinician-Rated/Self Rated, ARP: aripiprazole, SD: standard deviation

Briefly, the clinical benefit of AA for treating patients with MDD have been proven in a number of early phase small scale OLSs with a use of various primary efficacy rating scales such as HAMD, MADRS and QIDS-16.<sup>18-24</sup> In the small OLSs, the primary endpoint improvement was variable across the studies due to multiple factors (e.g., patient characteristics, AA dose, duration of treatment, e.t.c.); for instance, the cumulative remission and response rates showed that, approximately 60% met criteria for remission and 80% met criteria for response at the end of treatment in the recent 12-week OLS,<sup>28</sup> while similar 12-week OLS showed that at the endpoint, the remission rate was 41.3% and the response rate was 55.2%. There have been three identically designed initial phase RCTs<sup>31-33</sup> and two subsequent RCTs.<sup>34,36</sup> As for the three RCTs,<sup>31-33</sup> patients with 1-3 historical failures in adequate antidepressant trials (total score  $\geq 18$  on the HAM-D17) were screened and then entered an 8-week prospective treatment phase. Incomplete responders were then randomized for treatment with either aripiprazole or placebo for 6 weeks. The primary efficacy endpoint was the mean change from baseline for the MADRS total score. In total 1,092 prospectively identified partial responders were randomized and 940 (86.4%) patients completed the 6-week three RCTs. In these three RCTs, significant improvements in the mean change of the MADRS total score (range=-8.5 to -10.1) with AA over placebo (-5.8 to -6.4) were observed.<sup>31-33</sup>

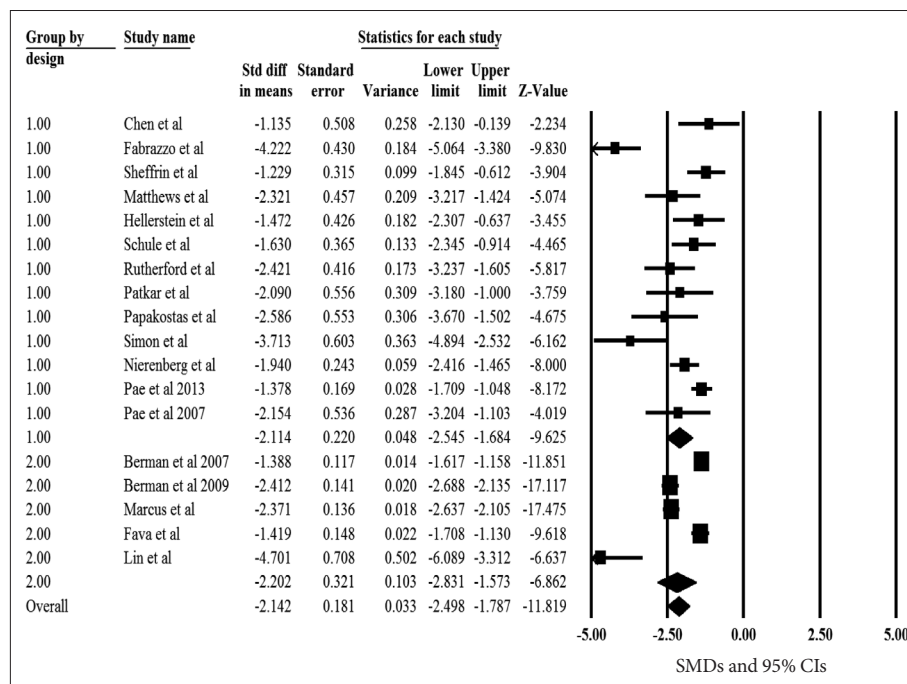
**OLSs**

The pooled SMDs for the primary efficacy measure was statistically significant, pointing out the significant reduction of

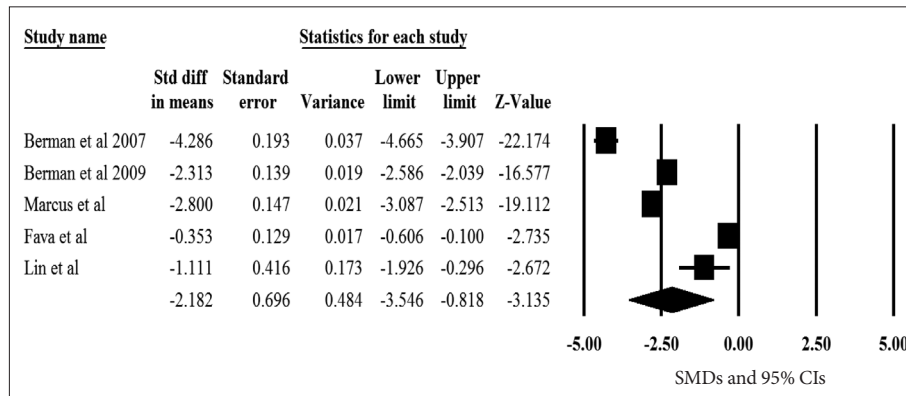
depressive symptoms after aripiprazole augmentation to current antidepressant treatment (pooled SMD=-2.114, z=-9.625, p<0.001) (Figure 2). The heterogeneity between OLSs was significant, pointing out the substantial variability of in the magnitude of treatment difference and underlying variance influencing on the outcome ( $I^2=80.1\%$ , Q-value=60.3, p<0.001). The pooled SMD was repeatedly calculated and analyzed with omission of one study at a time to perform a sensitivity analysis; the pooled SMD of the primary efficacy measure ranged from -2.210 to -1.898 (all 95% CIs indicated the statistical significance: range from -2.710 to -1.552), proposing that no one study has strongly impacted the pooled SMD. The Egger test was not statistically significant (t=2.114, p=0.055), indicating no publication bias.

**RCTs**

The pooled SMDs for the primary efficacy measure was statistically significant, pointing out the significant reduction of depressive symptoms after aripiprazole augmentation to current antidepressant treatment (pooled SMD=-2.202, z=-6.862, p<0.001) (Figure 2). The heterogeneity between RCTs was significant, pointing out the substantial variability of in the magnitude of treatment difference and underlying variance influencing on the outcome ( $I^2=94.3\%$ , Q-value=70.6, p<0.001). The pooled SMD was repeatedly calculated and analyzed with omission of one study at a time to perform a sensitivity analysis; the pooled SMD of the primary efficacy measure ranged from -2.430 to -1.896 (all 95% CIs indicated the statistical significance: range from -3.154 to -1.333), proposing that no one study has strongly impacted the pooled SMD. The Egger test



**Figure 2.** A meta-analysis of the primary efficacy measure of open-label studies (OLSs) and randomized, placebo-controlled trials (RCTs). The numbers of 1 and 2 indicate study design, OLS and RCT, respectively. SMD: standardized mean difference, CIs: confidence intervals.



**Figure 3.** A meta-analysis of the primary efficacy measure of aripiprazole versus placebo in randomized, placebo-controlled trials. SMD: standardized mean difference, CIs: confidence intervals.

was not statistically significant ( $t=1.117$ ,  $p=0.345$ ), indicating no publication bias.

### Aripiprazole vs. placebo SMDs for RCTs

The pooled SMDs for the primary efficacy measure was statistically significant, pointing out the significant reduction of depressive symptoms after aripiprazole augmentation vs. placebo to current antidepressant treatment (pooled SMD=-2.182,  $z=-3.135$ ,  $p=0.002$ ) (Figure 3). The heterogeneity was also significant, pointing out the substantial variability of in the magnitude of treatment difference and underlying variance influencing on the outcome among study ( $I^2=98.8\%$ ,  $Q$ -value=341.7,  $p<0.001$ ). The pooled SMD was repeatedly calculated and analyzed with omission of one study at a time to perform a sensitivity analysis; the pooled SMD of the primary efficacy measure ranged from -2.673 to -1.655 (all 95% CIs indicated the statistical significance: range from -4.025 to -0.255), proposing that no one study has strongly impacted the pooled SMD. The Egger test was not statistically significant ( $t=0.416$ ,  $p=0.705$ ), indicating no publication bias.

### Meta-regression

Meta-regression was performed to test the effect of study design (OLs vs. RCTs) as an independent parameter on the mean change in the primary efficacy measure. There was no evidence of statistical difference in the pooled SMDs for the primary efficacy measure between OLs and RCTs ( $t=0.119$ ,  $p=0.737$ ), suggesting no substantial influence of design on the primary treatment outcome. However, the pooled SMD was numerically higher in RCTs (-2.202, 95% CIs=-2.831, -1.573) than in OLs (-2.114, 95% CIs=-2.545, -1.684). Given aforementioned results, the overall pooled SMDs between the two designs were sufficiently resembling each other and the correlation between the two design of pooled SMDs were 0.714, indicating a similarity of the results between such two designs.

## DISCUSSION

We tried to find any useful and informative data between the two study designs for AA, OLS and RCT, in the treatment of MDD: OLS may have a potential utility to guide the RCT for proving the effect of AA in the treatment of MDD. According to our results, the pooled SMDs for the primary efficacy measure was statistically significant in both study design, showing a significant reduction of depressive symptoms after AA treatment to current antidepressant treatment in OLs and in RCTs. The effect sizes measured by SMDs between OLs and RCT design was quite similar and adequately correlated, indicating a practical utility of OLS design to move to RCT conduction in the treatment of MDD. When sensitivity analyses show that the overall result and conclusions are not affected by the different decisions that could be made during the review process, the results of the review can be regarded with a higher degree of certainty. Where sensitivity analyses identify particular decisions or missing information that greatly influence the findings of the review, greater resources can be deployed to try and resolve uncertainties and obtain extra information, possibly through contacting trial authors and obtained individual patient data.<sup>37</sup> If this cannot be achieved, the results must be interpreted with an appropriate degree of caution. Such findings may generate proposals for further investigations and future research.<sup>37</sup> In addition, no publication biases were found in the present meta-analysis, indicating the validity of the results of a meta-analysis results; if not so, no matter how systematic and thorough in other respects in meta-analysis, the results are not confident.<sup>37</sup> Our meta-regression analysis clearly revealed no influence of the study design on for treatment outcome, proving that study design would not have any role as a predictor to the primary treatment outcome.<sup>17</sup>

Our results are in line with the previous meta-analysis found similarities in the treatment effects between OLs and RCTs in youth with bipolar disorder indicating that studies with open-label design are useful predictors of the potential safety

and efficacy of a given compound in the treatment of pediatric bipolar disorder, which was the first meta-analysis investigating such design issue in psychiatry.<sup>38</sup> In the study, the pooled effect size was statistically significant in both OLSs ( $z=8.88$ ,  $p<0.001$ ) and RCTs ( $z=13.75$ ,  $p<0.001$ ), indicating a significant reduction in the Young Mania Rating Scale (YMRS) from baseline to the end of treatment in both study designs. The meta-regression also confirmed that study design was not a significant predictor of mean change in the YMRS. Therefore, our meta-analysis clearly replicated the previous findings by Biederman et al that study design would not affect treatment outcome and OLS may be a substantial indicator to lead a subsequent RCT to fully address a certain medication's efficacy.

Recently, there have been a number of evolving meta-analytic approaches to investigate clinically critical and very informative issues in terms of study design implicated in clinical practice as well. For instance, placebo-response rates that is challenging obstacles for new treatment development in MDD (according to the results, relative efficacy of the active drug compared to placebo in clinical trials for MDD is highly heterogeneous across studies with different placebo response rates; the more placebo response the less performance of active drug),<sup>39</sup> the application of a prospective lead-in trial phase to assess antidepressant nonresponse (historical data only to define treatment resistance prior to patient enrollment),<sup>40</sup> impact of number of follow-up assessment (increasing the number of follow-up visits, specifically after the third week rather than within the first 3 week of the trial, may be an effective approach to improve the likelihood of trial success),<sup>40</sup> the impact of study duration on treatment outcome (4 weeks is the minimum adequate length of a trial), and starting dose issue of SSRIs (Higher starting dose higher response), e.t.c. These results deliver useful and valuable information to clinicians about the trial design as well as doing clinical practice.

The most important clinical implication in terms of OLS and RCT design should be placed on the generalizability and external validity, since patients encountered in clinical practice often do not mirror populations of patients enrolled in well-controlled and adequately powered industry-sponsored or independent clinical trials.<sup>41</sup> This emphasis has yielded large effectiveness trials such as the Sequenced Treatment Alternatives to Relieve Depression (STAR\*D) and the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) trials designed to inform clinicians about the relative strengths of already existing but not fully investigated treatment approaches in the management of major psychiatric disorders. These studies aim to enroll typical community patients by having relatively lenient inclusion/exclusion criteria and concomitant medication restrictions, thus maximizing external validity. Hence, innovations in clinical trials methodology may

stem from a handful of case reports and small-scale OLSs, which may be a useful transition process to RCT or further advanced, controlled clinical trials.<sup>41</sup> Accordingly, OLS and RCT may have their own merits and limitations as a research methodology, by which we have to consider their complementary role to another one to have better and clear understanding for achievement of advanced and innovative treatment tactics and strategies.

The limitation of our results include: 1) the data base searched in the present meta-analysis were confined to PubMed and PsychInfo, and published journals, so that we could not collect all the available clinical trial data, although our results did not show any evidence of publication bias and skew in sensitivity analyses 2) the clinical samples were adult population and main portion was female 3) the mean duration was less than 10 weeks in both trial design, thus not be able to ensure any different results in longer-term clinical trials 4) the inclusion of different primary efficacy measure such as HAMD, MADRS and QIDS, although such rating scales are found to be highly correlated to each other, and finally 5) no inclusion of safety and tolerability measure due to their high variability in the measurement methods, by which we could not apply our results in such clinical issues.

In conclusion, we found that OLS for AA may be a useful indicator to conduct such time-consuming, complex and very expensive RCT in the treatment of MDD. Our results should be proved in other studies with different atypical antipsychotics for the treatment of MDD as well. Furthermore, the value of OLS should be re-evaluated as one of crucial steps for development of new drugs or acquisition for new indications.

#### Acknowledgments

This work was supported by a grant of the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (HI12C0003).

#### REFERENCES

1. Egede LE. Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *Gen Hosp Psychiatry* 2007;29:409-416.
2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003;289:3095-3105.
3. Practice guideline for the treatment of patients with major depressive disorder (revision). American Psychiatric Association. *Am J Psychiatry* 2000;157(4 Suppl):1-45.
4. Thase ME. Are SNRIs more effective than SSRIs? A review of the current state of the controversy. *Psychopharmacol Bull* 2008;41:58-85.
5. Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC. Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? A meta-analysis of studies of newer agents. *Biol Psychiatry* 2007;62:1217-1227.
6. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill

- R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746-758.
7. Fava M, Rush AJ, Trivedi MH, Nierenberg AA, Thase ME, Sackeim HA, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR\*D) study. *Psychiatr Clin North Am* 2003;26:457-494.
  8. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR\*D: implications for clinical practice. *Am J Psychiatry* 2006;163:28-40.
  9. Nierenberg AA, Husain MM, Trivedi MH, Fava M, Warden D, Wisniewski SR, et al. Residual symptoms after remission of major depressive disorder with citalopram and risk of relapse: a STAR\*D report. *Psychol Med* 2010;40:41-50.
  10. Paykel ES, Ramana R, Cooper Z, Hayhurst H, Kerr J, Barocka A. Residual symptoms after partial remission: an important outcome in depression. *Psychol Med* 1995;25:1171-1180.
  11. Thase ME. Evaluating antidepressant therapies: remission as the optimal outcome. *J Clin Psychiatry* 2003;64 Suppl 13:18-25.
  12. Bauer M, Whybrow PC, Angst J, Versiani M, Moller HJ. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders, Part 1: Acute and continuation treatment of major depressive disorder. *World J Biol Psychiatry* 2002;3:5-43.
  13. Anderson IM, Ferrier IN, Baldwin RC, Cowen PJ, Howard L, Lewis G, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2008;22:343-396.
  14. Kennedy SH, Lam RW, Cohen NL, Ravindran AV; CANMAT Depression Work Group. Clinical guidelines for the treatment of depressive disorders. IV. Medications and other biological treatments. *Can J Psychiatry* 2001;46(Suppl 1):38S-58S.
  15. Philip NS, Carpenter LL, Tyrka AR, Price LH. Augmentation of antidepressants with atypical antipsychotics: a review of the current literature. *J Psychiatr Pract* 2008;14:34-44.
  16. Muller MJ, Szegei A, Wetzl H, Benkert O. Moderate and severe depression. Gradients for the Montgomery-Asberg Depression Rating Scale. *J Affect Disord* 2000;60:137-140.
  17. Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI; Health Outcomes, Policy, and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. *Int J Clin Pract* 2009;63:1426-1434.
  18. Chen SJ, Hsiao YL, Shen TW, Chen ST. The effectiveness and safety of adjunctive aripiprazole in Taiwanese patients with antidepressant-refractory major depressive disorder: a prospective, open-label trial. *J Clin Psychopharmacol* 2012;32:56-60.
  19. Fabrazzo M, Perris F, Monteleone P, Esposito G, Catapano F, Maj M. Aripiprazole augmentation strategy in clomipramine-resistant depressive patients: an open preliminary study. *Eur Neuropsychopharmacol* 2012;22:132-136.
  20. Sheffrin M, Driscoll HC, Lenze EJ, Mulsant BH, Pollock BG, Miller MD, et al. Pilot study of augmentation with aripiprazole for incomplete response in late-life depression: getting to remission. *J Clin Psychiatry* 2009;70:208-213.
  21. Matthews JD, Siefert C, Dording C, Denninger JW, Park L, van Nieuwenhuizen AO, et al. An open study of aripiprazole and escitalopram for psychotic major depressive disorder. *J Clin Psychopharmacol* 2009;29:73-76.
  22. Hellerstein DJ, Batchelder S, Hylar S, Arnaout B, Corpuz V, Coram L, et al. Aripiprazole as an adjunctive treatment for refractory unipolar depression. *Prog Neuropsychopharmacol Biol Psychiatry* 2008;32:744-750.
  23. Schule C, Baghai TC, Eser D, Hecht S, Hermisson I, Born C, et al. Mirtazapine monotherapy versus combination therapy with mirtazapine and aripiprazole in depressed patients without psychotic features: a 4-week open-label parallel-group study. *World J Biol Psychiatry* 2007;8:112-122.
  24. Rutherford B, Sneed J, Miyazaki M, Eisenstadt R, Devanand D, Sackeim H, et al. An open trial of aripiprazole augmentation for SSRI non-remitters with late-life depression. *Int J Geriatr Psychiatry* 2007;22:986-991.
  25. Patkar AA, Peindl K, Mago R, Mannelli P, Masand PS. An open-label, rater-blinded, augmentation study of aripiprazole in treatment-resistant depression. *Prim Care Companion J Clin Psychiatry* 2006;8:82-87.
  26. Papakostas GI, Petersen TJ, Kinrys G, Burns AM, Worthington JJ, Alpert JE, et al. Aripiprazole augmentation of selective serotonin reuptake inhibitors for treatment-resistant major depressive disorder. *J Clin Psychiatry* 2005;66:1326-1330.
  27. Simon JS, Nemeroff CB. Aripiprazole augmentation of antidepressants for the treatment of partially responding and nonresponding patients with major depressive disorder. *J Clin Psychiatry* 2005;66:1216-1220.
  28. Nierenberg AA, Trivedi MH, Gaynes BN, Mitchell J, Davis LL, Husain MM, et al. Effectiveness Study of Venlafaxine-XR Combined with Aripiprazole for Chronic or Recurrent Major Depressive Disorder. *Australian and New Zealand Journal of Psychiatry* 2009;43:956-967.
  29. Pae CU, Jeon HJ, Lee BC, Seo HJ, Kim SG, Park EJ, et al. Aripiprazole augmentation for treatment of patients with chronic or recurrent major depressive disorder: a 12-week prospective open-label multicentre study. *Int Clin Psychopharmacol* 2013;28:322-329.
  30. Pae CU, Patkar AA, Jun TY, Lee C, Masand PS, Paik IH. Aripiprazole augmentation for treatment of patients with inadequate antidepressant response. *Depress Anxiety* 2007;24:522-526.
  31. Berman RM, Marcus RN, Swanink R, McQuade RD, Carson WH, Corey-Lisle PK, et al. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a multicenter, randomized, double-blind, placebo-controlled study. *J Clin Psychiatry* 2007;68:843-853.
  32. Berman RM, Fava M, Thase ME, Trivedi MH, Swanink R, McQuade RD, et al. Aripiprazole augmentation in major depressive disorder: a double-blind, placebo-controlled study in patients with inadequate response to antidepressants. *CNS Spectr* 2009;14:197-206.
  33. Marcus RN, McQuade RD, Carson WH, Hennicken D, Fava M, Simon JS, et al. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a second multicenter, randomized, double-blind, placebo-controlled study. *J Clin Psychopharmacol* 2008;28:156-165.
  34. Fava M, Mischoulon D, Iosifescu D, Witte J, Pencina M, Flynn M, et al. A double-blind, placebo-controlled study of aripiprazole adjunctive to antidepressant therapy among depressed outpatients with inadequate response to prior antidepressant therapy (ADAPT-A Study). *Psychother Psychosom* 2012;81:87-97.
  35. Lin CH, Lin SH, Jang FL. Adjunctive low-dose aripiprazole with standard-dose sertraline in treating fresh major depressive disorder: a randomized, double-blind, controlled study. *J Clin Psychopharmacol* 2011;31:563-568.
  36. Mischoulon D, Witte J, Levy M, Papakostas GI, Pet LR, Hsieh WH, et al. Efficacy of dose increase among nonresponders to low-dose aripiprazole augmentation in patients with inadequate response to antidepressant treatment: a randomized, double-blind, placebo-controlled, efficacy trial. *J Clin Psychiatry* 2012;73:353-357.
  37. Collaboration TC. 9.7 Sensitivity analyses In *Cochrane Handbook for Systematic Reviews of Interventions*. Ed by Higgins JPT and Green S accessed on August 30, 2013. Available at <http://handbook.cochrane.org>.
  38. Biederman J, Petty CR, Woodworth KY, Lomedico A, O'Connor KB, Wozniak J, et al. How informative are open-label studies for youth with bipolar disorder? A meta-analysis comparing open-label versus randomized, placebo-controlled clinical trials. *J Clin Psychiatry* 2012;73:358-365.
  39. Iovieno N, Papakostas GI. Correlation between different levels of placebo response rate and clinical trial outcome in major depressive disorder: a meta-analysis. *J Clin Psychiatry* 2012;73:1300-1306.



40. Iovieno N, Papakostas GI. Does the presence of an open-label antidepressant treatment period influence study outcome in clinical trials examining augmentation/composition strategies in treatment partial responders/nonresponders with major depressive disorder? *J Clin Psychiatry* 2012;73:676-683.
41. Marks DM, J T, Pae CU. Innovations in clinical research design and conduct in psychiatry: shifting to pragmatic approaches. *Psychiatry Investig* 2009;6:1-6.