Comparison between long-acting injectable aripiprazole versus paliperidone palmitate in the treatment of schizophrenia: systematic review and indirect treatment comparison

Chi-Un Pae[a,d], Sheng-Min Wang[a,b], Changsu Han[c], Won-Myong Bahk[a], Soo-Jung Lee[a], Ashwin A. Patkar[d], Prakash S. Masand[e,f], Alessandro Serretti[g] and Robin Emsley[h]

We investigated the relative efficacy and tolerability of aripiprazole once monthly (AOM) versus paliperidone palmitate (PP) for treating schizophrenia. Extensive databases searches on short-term, placebo-controlled, randomized studies of AOM and PP were performed. Indirect treatment comparisons were performed between the two long-acting injectable antipsychotics (LAIAs). The primary efficacy endpoint was the mean change in the Positive and Negative Syndrome Scale total score from baseline between each LAIA and placebo. The effect sizes were mean differences and odds ratio (ORs) with 95% confidence intervals (CIs) for the primary efficacy endpoint and safety/tolerability between two LAIAs, respectively. Mean difference in the primary efficacy endpoint was significantly different, favouring AOM over PP (OR: −6.4; 95% CI: −11.402 to −1.358); sensitivity analyses and noninferiority test (AOM vs. PP) confirmed the primary results. The overall early dropout rate was not significantly different between AOM and PP (OR: 1.223; 95% CI: 0.737–2.03). However, there was a significant difference in the early dropout rate in terms of lack of efficacy favouring AOM over PP (OR: 0.394; 95% CI: 0.185–0.841). Within the context of the inherent limitations of the current analysis, our results may suggest that there may be relative advantages for AOM over PP in the short-term treatment of schizophrenia. ***Int Clin Psychopharmacol*** **32:235–248**

International Clinical Psychopharmacology 2017, **32**:235–248

Keywords: aripiprazole once monthly, efficacy, indirect comparison, paliperidone palmitate, schizophrenia

[a]Department of Psychiatry, The Catholic University of Korea College of Medicine, [b]International Health Care Center, Seoul St Mary's Hospital, The Catholic University of Korea College of Medicine, [c]Department of Psychiatry, Korea University, College of Medicine, Seoul, Republic of Korea, [d]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina, [e]Global Medical Education, New York, New York, USA, [f]Department of Psychiatry, Duke-NUS Medical School, Singapore, [g]Department of Biomedical and NeuroMotor Sciences, Institute of Psychiatry, University of Bologna, Bologna, Italy and [h]Faculty of Medicine and Health Sciences, Psychiatry, Stellenbosch University, Cape Town, South Africa

Correspondence to 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-si, Gyeonggi-do 14647, Republic of Korea
Tel: +82 323 407 067; fax: +82 323 402 255; e-mail: pae@catholic.ac.kr

Received 10 January 2017 **Accepted** 16 March 2017

Introduction

Schizophrenia is a debilitating psychiatric illness with a lifetime prevalence of ~1%. Clinical manifestations include hallucinations, delusions and disordered thinking, and impairments of personal, social and occupational functioning, and it typically causes substantial and persistent disability in diverse aspects in their routine life (Saha *et al.*, 2005; McGrath *et al.*, 2008).

The course of illness is characterized by multiple relapses and recurrences, most often as a consequence of antipsychotic treatment discontinuation (Lieberman *et al.*, 2001; Emsley *et al.*, 2013a). Relapse and recurrences may occur very soon after treatment discontinuation and onset may be abrupt, with rapid return of severe symptoms (Lieberman *et al.*, 2001). In addition, a longer treatment period before abrupt discontinuation of antipsychotics does not reduce the risk of relapse and recurrence. Furthermore, successive relapses and recurrences may contribute towards emergent treatment refractoriness with longer treatment periods and higher antipsychotic doses necessary to regain symptom stabilization (Lieberman *et al.*, 1996; Emsley *et al.*, 2013a). In addition, frequent and multiple relapses may contribute towards family discord, increased social stigma, suicidal and homicidal behaviour, and increased medical costs and negatively affect personal, social, biological, functional and health-related quality-of-life (HRQoL) outcomes (Lieberman *et al.*, 2001). Multiple relapses may also have neurobiological consequences. A reduced threshold for psychotic decompensation, resurgence of a hyperdopaminergic state, mesolimbic dopaminergic supersensitivity,

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (*www.intclinpsychopharm.com*).

DOI: 10.1097/YIC.0000000000000177

increased excitatory glutamatergic activity and alterations in inflammatory system may contribute towards neuroprogression (Birchwood *et al*., 1989; Kapur *et al*., 2006; Leonard *et al*., 2012; Emsley *et al*., 2013a, 2013b). Finally, the increased economic burden associated with relapse is another caveat; relapse-related medical costs mainly account for more than 60% of all direct medical costs (Lindstrom *et al*., 2007). According to a recent 10-year follow-up population-based study investigating the readmission rate after first hospitalization (Chi *et al*., 2016), a quarter of patients were readmitted within 4 months of the first hospitalization and overall 71% patients were readmitted within 10 years. The median time between admissions was ~2 years, indicating that schizophrenia has a high rate of readmission in real-world treatment settings.

Antipsychotics, through dopamine receptor antagonism, provide the mainstay of treatment for schizophrenia. Psychosocial treatment modalities are also necessary to optimize functional capacity (Stahl *et al*., 2013). A wealth of evidence supports the use of antipsychotic maintenance treatment as a fundamental element of the management of schizophrenia (De Hert *et al*., 2015). This position is reflected in numerous treatment guidelines (Lehman *et al*., 2004; Barnes *et al*., 2011; Hasan *et al*., 2012, 2013; Galletly *et al*., 2016). However, the demonstrated efficacy of antipsychotic treatment needs to be considered in the context of several shortcomings. For example, in addition to their well-recognized side-effect burden, ~30% of patients with schizophrenia may relapse while receiving maintenance antipsychotic treatment (Hogarty and Ulrich, 1998). Furthermore, ~30–50% of patients with schizophrenia are poorly adherent to their antipsychotics at any given time (Valenstein *et al*., 2006; Cooper *et al*., 2007). Interestingly, a recent 3-year follow-up study (Caseiro *et al*., 2012) found that nonadherence increased proportionally over treatment and was the most important predictor of relapse. Multiple relapses also negatively affect patients' HRQoL including personal, social and occupational functioning. According to recent treatment guidelines, improvements in functioning and HRQoL are recognized ultimate treatment goals (Lehman *et al*., 2004; Barnes *et al*., 2011; Hasan *et al*., 2012, 2013; Stahl *et al*., 2013; Galletly *et al*., 2016). Therefore, given that relapse and poor HRQoL are strongly linked to antipsychotic nonadherence, methods of improving adherence need to be prioritized.

In this context, long-acting injectable antipsychotics (LAIAs) were developed to help patients adhere to antipsychotic maintenance treatment. Indeed, LAIAs were found to significantly reduce the risk of relapse compared with their oral formulations in real-world treatment settings (Kirson *et al*., 2013). LAIAs can also improve patient HRQoL, which is increasingly being recognized as an optimal treatment outcome (Stahl, 2014). Hence, LAIAs are considered a valuable treatment option for maintenance treatment for schizophrenia in most treatment guidelines (Lehman *et al*., 2004; Barnes *et al*., 2011; Hasan *et al*., 2012, 2013; Galletly *et al*., 2016). In the past several decades, LAIAs mainly originated from first-generation antipsychotics (FGAs) such as flupentixol, fluphenazine, haloperidol, perphenazine, pipotiazine and zuclopenthixol, after which LAIAs with second-generation antipsychotics (SGAs) were introduced, including risperidone, olanzapine and paliperidone palmitate (PP).

In March 2013, the intramuscular once-monthly formulation of aripiprazole (AOM) was approved by the US Food and Drug Administration for the treatment of schizophrenia (FDA. Abilify Maintena Kit, 2013) and was also approved in Europe for maintenance treatment of schizophrenia in adult patients stabilized with oral aripiprazole in November 2013 (Agency, 2013). Before the introduction of AOM, PP was the most recently introduced LAIA for the treatment of schizophrenia. Patients with schizophrenia may respond differently to individual antipsychotics in terms of symptom reduction and treatment emergent adverse events (TEAEs), in keeping with the differential pharmacological profiles of these agents (Leucht *et al*., 2013). Thus, there is a need for additional treatment options in terms of LAIAs as it is likely that LAIAs share the receptor-binding profiles of their oral formulations, although a dearth of evidence exists on differential effects among LAIAs (Correll *et al*., 2016). In this respect, recent studies suggest that AOM and PP may have different effects in terms of pharmacoeconomics (Sapin *et al*., 2016) and HRQoL (Naber *et al*., 2015), favouring AOM over PP. However, there has been a lack of short-term, head-to-head, randomized, direct comparison trials between AOM and PP and meta-analyses to date have focused on comparing multiple LAIAs with placebo (Fusar-Poli *et al*., 2013) or oral formulations (Kishimoto *et al*., 2013; Lafeuille *et al*., 2014). Only one meta-analysis used an indirect comparison to investigate relapse rates among multiple LAIAs (Haloperidol depot, AOM, risperidone LAIA, PP, olanzapine pamoate), oral formulations and placebo, and included a small number of studies ($N=6$).

This is the first systematic review and indirect meta-analysis of relative efficacy and tolerability of AOM versus PP in short-term randomized controlled trials (RCTs) in patients with schizophrenia. We adopted an indirect treatment comparison (ITC) approach to overcome the lack of studies directly comparing the two treatments.

Methods

Objective of the present study

A direct assessment of two treatments A and B should be valuable if direct comparative studies between A and B are performed. Unfortunately, many competitors have not been directly compared because of multiple reasons including lack of funding source and thereby direct

evidence between competitor medications is very scarce and insufficient. To overcome such limitations, ITC have been used considerably when direct evidence is not available and inadequate (Wells *et al.*, 2009). ITCs have been also utilized to compare certain antipsychotics in the treatment of schizophrenia, resulting in proving of consistency between the results from ITCs and direct comparison studies (Kunitomi *et al.*, 2014).

AOM and PP are relatively recently introduced LAIAs in the market and expected to have high competition in the treatment of schizophrenia; there has been no direct comparison study between AOM and PP to date. Therefore, the present study aimed to perform ITC between AOM and PP in the short-term treatment of schizophrenia.

Data search

PubMed, Embase, Medline, the Cochrane Central Register of Controlled Trials and the Cochrane Library were searched as of 31 May 2016. The search term was 'schizophrenia', 'abilify maintenna', 'aripiprazole', 'abilify', 'injection', 'bolus', 'paliperidone palmitate', 'paliperidone' and 'invega sustenna', which were properly combined as provided in Supplementary Appendix 1, Supplemental digital content 1, *http://links.lww.com/ICP/A24* and Supplementary Appendix 2, Supplemental digital content 2, *http://links.lww.com/ICP/A25*. A final research for studies was performed in the same search engines on 1 September 2016, in which we did not find more studies for inclusion in the present study. Reference lists from identified articles and reviews were also used to find additional studies. Abstracts identified by the literature search were evaluated independently by two authors (S.M.W. and C.U.P.); potentially eligible papers were then re-evaluated by two other authors (C.H. and S.J.L.) to determine whether they clearly fulfilled the selection criteria. If a disagreement occurred, the article in question was discussed and a consensus was reached by the second set of review authors on the basis of the check list of the present study as provided in Supplementary Appendix 3, Supplemental digital content 3, *http://links.lww.com/ICP/A26*.

Study selection and exclusion

A systematic literature review was performed to identify relevant RCTs using AOM and PP along with the use of placebo as a common comparator in the acute treatment of schizophrenia. Blinding of RCTs was restricted to double-blind designs. Our search strategy adopted the guideline on core interventions in the treatment and management of schizophrenia provided by the National Institute on Clinical and Health Excellence (NICE).

All short-term RCTs investigating the efficacy and safety of AOM versus placebo and PP versus placebo for schizophrenia were the primary inclusion criteria and the age of the patients had to be at least 18 years. Studies with three or more treatment arms were to be excluded as the objectives of the present study was to indirectly compare the efficacy and tolerability between AOM and PP in the treatment of schizophrenia and the estimates obtained from a three-arm trial are correlated (Bucher *et al.*, 1997); for instance, if certain RCT had three treatment arms with AOM, another active comparator and placebo, they were to be excluded as ITC analysis could be suitable for two arm RCTs (Bucher *et al.*, 1997; Wells *et al.*, 2009). If RCTs have two arms consisting of AOM and an active comparator that was also used in certain RCT with PP, such RCT with a common comparator could be included. Overall, the short-term RCTs of AOM and PP were mainly within 16 weeks. Patients had to fulfil the criteria for schizophrenia used in the individual trials. Studies were excluded if the primary efficacy endpoint outcome was the prevention of relapse or if the primary efficacy endpoint, change of the Positive and Negative Syndrome Scale (PANSS) total score from baseline, was not available. There were no requirements or restrictions in terms of the date, severity of illness, race, duration of illness, onset age, sex, minimum number of patients, study location or treatment basis (i.e. inpatient or outpatient).

Data extraction

Data on the treatment details, detailed study procedures/design, citation details, number of patients in each treatment arm, age, sex, duration of individual study, the change of PANSS total score from baseline to the end of treatment, doses of LAIAs, study location, race, onset age, duration of illness, early dropouts (EDs) and TEAEs were collected.

All data with all specified information extracted from individual RCT were carefully inspected and recorded in a separate data extraction file. The quality of individual RCT was evaluated in accordance with the recommendations from the Cochrane Review. The risk of bias associated with sequence generation, allocation concealment, the blinding of participants and investigators, the blinding of outcome assessments, incomplete outcome data, selective outcome reporting and other sources were assessed using specific and detailed criteria. If there was any disagreement in data extraction and assessment in the risk of bias in individual RCT, it was resolved by consensus between two authors who were independently in charge of data extraction and quality assessment of individual RCT (C.U.P. and C.H.).

Efficacy measures

The primary efficacy endpoint was the mean change in the PANSS total score from baseline to the end of treatment as defined by the individual RCT as it has been the most commonly and universally accepted for the evaluation of antipsychotic efficacy in the acute treatment of schizophrenia (Kay *et al.*, 1987; Hermes *et al.*, 2012).

Noninferiority test of aripiprazole once monthly over paliperidone palmitate

Noninferiority testing is usually utilized to evaluate whether a new treatment is not unacceptably less efficacious than the comparator treatment already in use. AOM has been approved relatively recently by authorities in some countries and it has not yet been approved in most Asian countries excluding Japan, Korea, Hongkong and Taiwan. Hence, we also attempted to test the noninferiority of AOM against PP in terms of the primary efficacy endpoint.

Safety and tolerability measures

Data on the number of EDs (for any reason) representing the usual tolerability of certain medications as they incorporate may reasons for early discontinuation of the study medication and TEAEs were also collected. Indeed, ED/TEAEs rates are well accepted for determining efficacy and safety/tolerability (Kunitomi *et al.*, 2014) as they include many reasons for EDs such as lack of efficacy, occurrence of safety/tolerability, loss to follow-up, patient's preference and so on; if the EDs rate are low, it indicates that study medication is tolerable. We attempted to use overall ED, ED because of TEAEs and ED because of lack of efficacy rates to compare the safety/tolerability between AOM and PP with the use of ITCs.

Data synthesis

In terms of continuous measures, data on the mean change of the PANSS total score from baseline to the end of treatment along with the SD and the number of patients in each treatment arm were extracted for the primary efficacy measure. Some studies do not report precise numerical data in the efficacy measures, but utilize only graphic demonstration, in which we could possibly consider the primary efficacy endpoint with visual measurement, which is vulnerable to any unexpected bias. Thus, such RCTs were not to be included in the present study. Missing SDs are also a common feature of meta-analyses of continuous outcome data and thereby missing SDs could also be taken from one or more other studies as imputation of SDs is usually known to produce approximately exact results. When several candidate SDs are available, their average could be used for the meta-analyses. SDs should also be calculated with standard error with 95% confidence intervals (95% CIs) as provided in individual studies. In terms of binary measures, events of ED and AE were collected for safety/tolerability evaluation.

Statistical analysis

Overall procedure

At first, separate meta-analyses for AOM versus placebo and PP versus placebo were carried out for direct comparison results between each LAI and placebo as placebo treatment was to be used as a common comparator for ITCs. Thereafter, ITCs were performed using the results from the meta-analyses. In addition, the noninferiority test of AOM against PP was also performed. All procedures of ITCs in the present study followed modern guidelines (Wells *et al.*, 2009; Hoaglin *et al.*, 2011).

The effect sizes for the primary efficacy endpoint in each RCT are presented as the mean difference (MD) with 95% CIs for both meta-analyses and ITCs. Odd ratio (OR) was used to assess the ED rates (overall/TEAEs/lack of efficacy) for both meta-analyses and ITCs.

The full analysis set was composed of all randomized patients who received at least one dose of study medication and had at least one valid postbaseline value for the primary efficacy assessment in each RCT. Full analysis set with a last-observation-carried-forward analysis was used to evaluate efficacy. The safety set included all randomized patients who received at least one dose of the study medication.

Random-effects models were applied for the analyses of primary efficacy endpoint and safety/tolerability measures in meta-analyses because of their better balancing (i.e. sampling and study size bias, etc.) than fixed-effects models.

Meta-analyses and indirect treatment comparisons

All efficacy and safety/tolerability data extracted from the individual RCTs were entered into the data box of the Comprehensive Meta-analysis, version 2.0 software for meta-analyses (CMA v2; Biostat Inc., Englewood, New Jersey, USA). The results from meta-analyses on efficacy and safety/tolerability data were eventually entered into the ITC software for ITCs between AOM and PP (Wells *et al.*, 2009). The ITC software based on the Bucher model (Bucher *et al.*, 1997) has been developed in visual basic to aid various calculations associated with ITCs; this can be easily downloaded in the public domain (*http://www.cadth.ca*) and is freely provided by the Canadian Agency for Drugs and Technologies in Health (CADTH) (Ottawa, Canada; v1.1., 2009). Briefly, the fundamental methodology of ITC was proposed and developed by Bucher *et al.* (1997) and it has been used widely for making indirect comparisons for various therapeutic agents in particular when direct comparisons are not available. The principal assumption of Bucher's ITC model (Bucher *et al.*, 1997) is that the relative efficacy of a treatment is the same across all trials included in the ITC. Upon no availability of a direct comparison between two treatments A and C, indirect estimate results obtained from trials A versus B and B versus C would be usable for their indirect comparison under the assumption that effects A and C are the same as in treatment B (Wells *et al.*, 2009).

Heterogeneity

Heterogeneity between studies was assessed using the I^2 statistic. This measure evaluates how much of the variance between studies can be attributed to the actual differences between the studies rather than to chance. A magnitude of considerable heterogeneity is usually $I^2 = 75–100\%$. The heterogeneity threshold was defined as 50% or more in the I^2 value and a P value less than 0.05. Sensitivity analyses were carried out to test the robustness of the impact of a single study on the overall results.

Publication bias

Egger tests were used to evaluate publication bias. These methods were adopted because Egger's linear regression method quantifies the bias captured by a funnel plot using the actual values and precision of the effect sizes, whereas Begg and Mazumdar's test uses ranks.

Noninferiority of aripiprazole once monthly against paliperidone palmitate

Noninferiority testing is usually utilized to evaluate whether a new treatment is not unacceptably less efficacious than the comparator treatment already in use. AOM has been approved relatively recently by authorities in some countries and it has not yet been approved in most Asian countries excluding Japan, Korea, Hongkong and Taiwan. Hence, we also attempted to test the noninferiority of AOM against PP in terms of the primary efficacy endpoint.

The noninferiority margin delta (Δ), the maximum acceptable extent of the clinical noninferiority of a newer agent, must be defined in priori. Subjective specification of the noninferiority margin on the basis of clinical significance is also used; however, the usage of treatment differences for the control group (placebo or active comparators) on the basis of previously published RCTs would be one of the best approaches to consider the size of the statistical margin (Hahn, 2012). The noninferiority margin could be chosen to be a fraction of the historical comparator treatment effect (i.e. 25, 50% etc.). Hence, the noninferiority margin was set at 50% of the average obtained from entire placebo treatment effects in all studies included in the present study. The Δ was calculated was calculated using the following formula: [Δ (absolute value) = summation of all MDs from AOM and PP RCTs/number of RCTs] × 0.5. The upper limits of the 95% CIs of the MDs obtained from ITCs results were then compared with such noninferiority margins Δ.

Results

Description of studies included in the meta-analysis

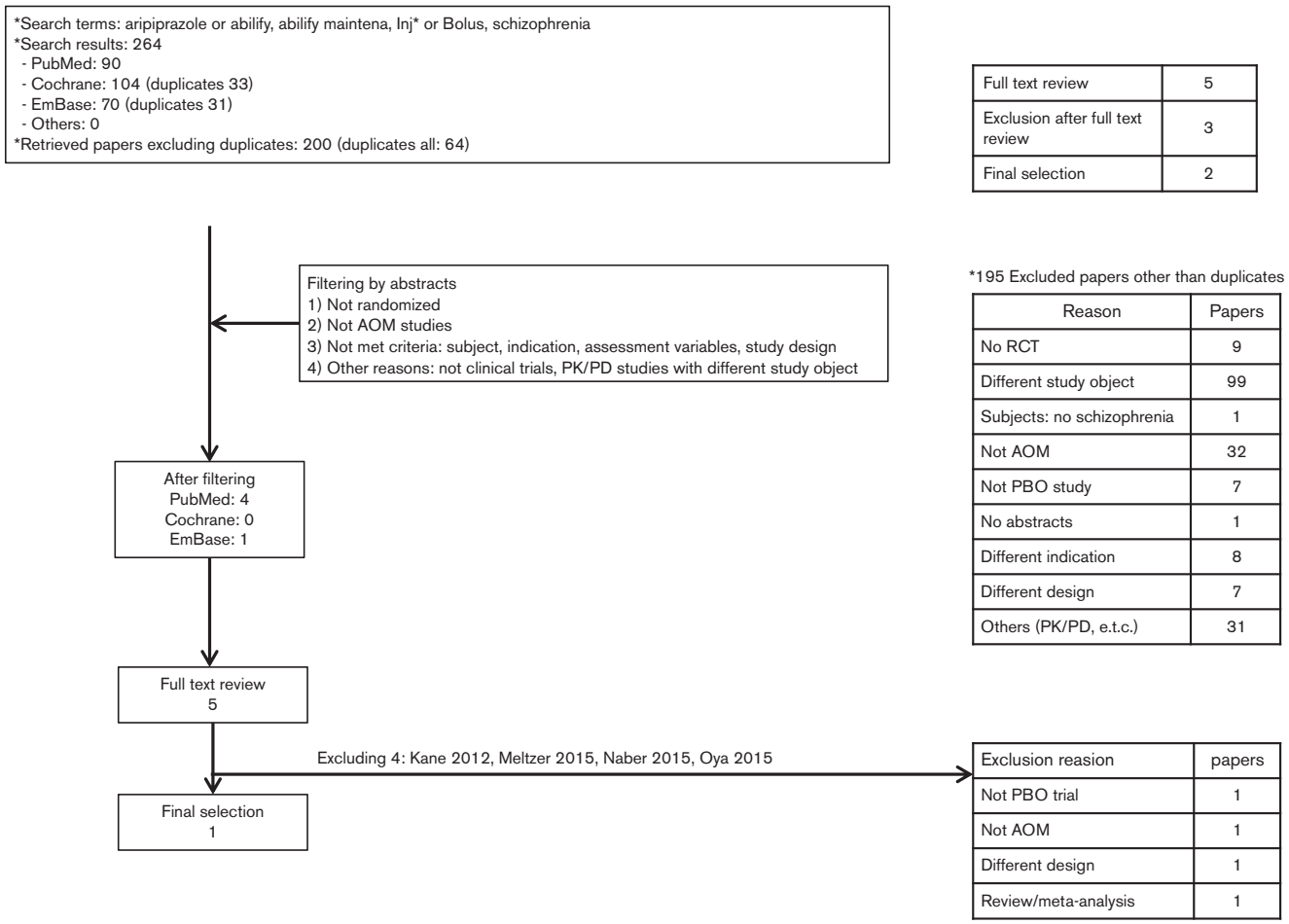
Of the 264 records on AOM trials identified by the search of the databases (Pubmed 90; Embase and Medline 70; Cochrane Library 104), 64 were excluded as they were duplicates. The remaining 200 studies were retrieved for more detailed evaluation as shown in Fig. 1. After a detailed evaluation, five studies were selected for a thorough review of the paper and thereafter only one RCT (Kane *et al.*, 2014) was included in the present study.

Of the 282 records on PP trials identified by the search of the databases (Pubmed 133; Embase and Medline 76; Cochrane Library 70, others 3), 90 were excluded as they were duplicates. The remaining 192 studies were retrieved for more detailed evaluation as shown in Fig. 2. After a detailed evaluation, nine studies were selected for a thorough review of the paper and thereafter four RCTs (Gopal *et al.*, 2010; Nasrallah *et al.*, 2010; Alphs *et al.*, 2011; Takahashi *et al.*, 2013) were included in the present study.

The main characteristics of these four short-term studies are presented in Table 1. All RCTs were multicentred and internationally conducted throughout the world; only one RCT was solely conducted in Asia including Japan, Korea and Taiwan. All study comparisons included 628 AOM/PP patients and 682 patients in the common comparator treatment arm. The entry total PANSS score was approximately more than or equal to 84 in whole treatment arms and the length of study was 12 weeks in an AOM RCT, whereas it was 13 weeks in four PP RCTs. All patients were diagnosed with a primary diagnosis of schizophrenia according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., text revision (DSM-IV-TR) criteria. Patients who had other comorbid psychiatric conditions or medically significant comorbid physical conditions were generally excluded.

Among 1310 patients in the studies included, 628 patients with schizophrenia were on AOM/PP and 682 were on placebo. The dose of AOM was 400 mg/month (Kane *et al.*, 2014), whereas PP used 156 mg/month (Gopal *et al.*, 2010; Nasrallah *et al.*, 2010; Alphs *et al.*, 2011) and 117 mg/month (Takahashi *et al.*, 2013). All studies had a tendency to show a preponderance of male patients, with proportions approximately ranging from 51 to 81% irrespective of treatment arms. All patients included in each treatment arm were considered to have at least moderate psychotic symptoms at baseline, represented by mean PANSS total scores approximately ranging from 84 to 104. All studies were financially supported by the pharmaceutical companies that manufactured the study drugs. In terms of the primary efficacy endpoint analysis, all RCTs included in the present study have shown clear superiority of AOM and PP over placebo group in the treatment of schizophrenia, showing more significant decreases in PANSS total scores from baseline at least twice than that of the placebo group. Most TEAEs were of mild to moderate intensity in both AOM and PP RCTs. The ED rate due to TEAEs was numerically lesser in AOM (7/168, 0.4%) than in placebo treatment (13/172, 0.8%). The ED rate due to TEAEs was also numerically lesser in PP (42/460, 9.1%) than in placebo treatment (73/510, 14.3%). There were two deaths in PP RCT (Nasrallah *et al.*, 2010), whereas this

Fig. 1



Search strategy and result of AOM clinical trials. Of the 264 records on AOM trials, five studies were eventually selected for a thorough review of the paper and thereafter only one RCT (Kane *et al.*, 2014) was included in the present study. AOM, aripiprazole once monthly; PBO, placebo; PD, pharmacodynamic; PK, pharmacokinetic; RCT, randomized clinical trial.

was not reported in AOM RCT. One patient in the PP group committed suicide after the third injection of the study medication and one patient in the placebo group died as a result of pancreatic carcinoma after receiving all four doses of the study medication (Nasrallah *et al.*, 2010).

Risk of bias

Figure 3 presents the overall risks of bias for all the studies included. The risk of bias was considered low or unclear in all studies on the basis of evaluations of all domains and no study was scored as presenting a high risk of bias in all domains. Overall, all the studies included were of high quality with respect to methodological considerations.

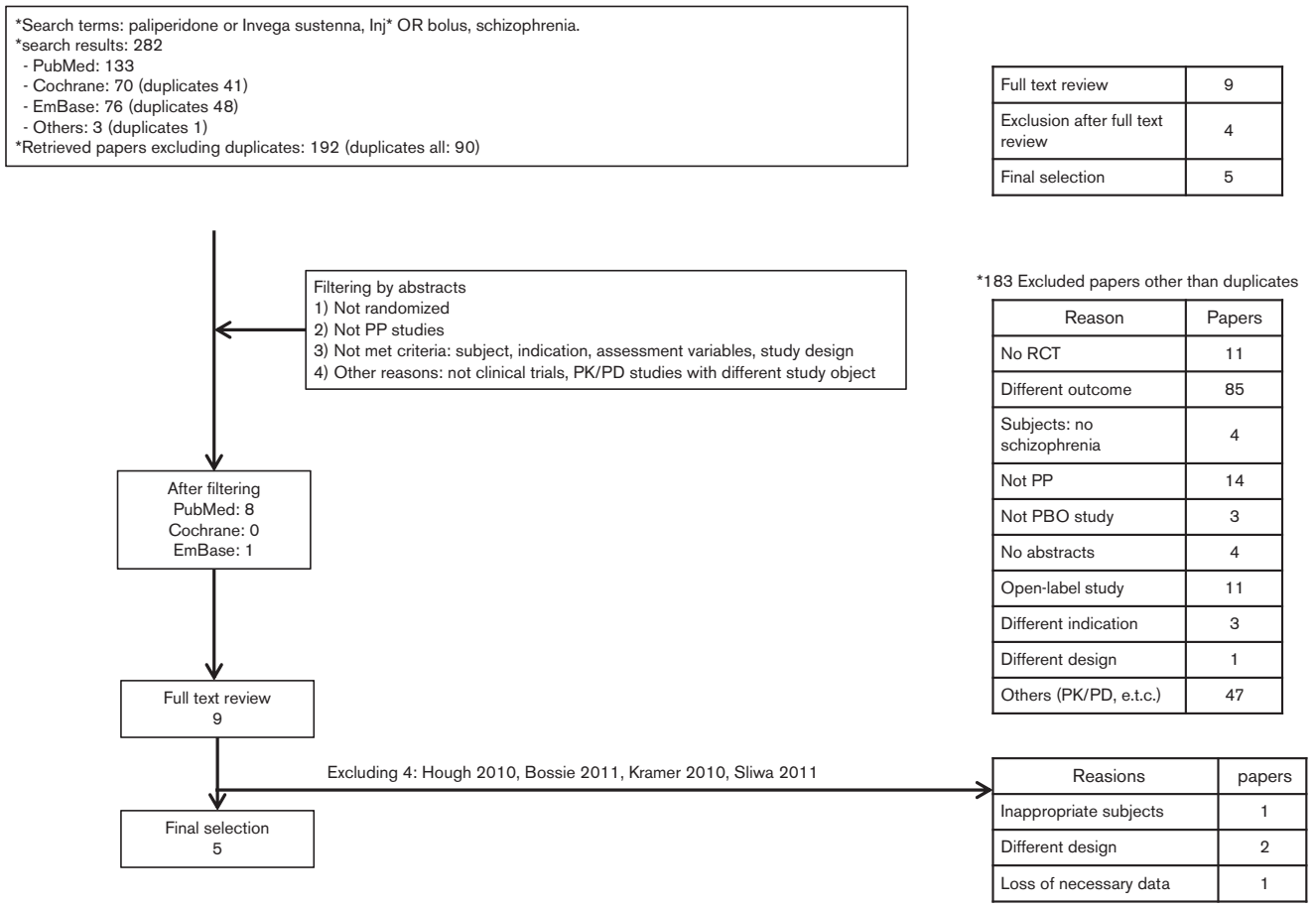
Efficacy

Meta-analyses (aripiprazole once monthly and/or paliperidone palmitate vs. placebo)

Overall: The result of the meta-analysis in terms of the primary efficacy endpoint is presented in Table 2. The MD (−15.1) in the primary efficacy endpoint was significantly different between AOM and placebo treatments, favouring AOM over placebo (95% CIs: −19.399 to −10.801). The MD (−8.7) in the primary efficacy endpoint was significantly different between PP and placebo treatments, favouring PP over placebo (95% CIs: −11.319 to −6.125). The MD (−10.4) in the primary efficacy endpoint was significantly different between AOM/PP as a whole treatment and placebo treatment, favouring AOM/PP over placebo (95% CIs: −12.650 to −8.204).

Sensitivity analysis, heterogeneity and publication bias: The heterogeneity among all RCTs was not significant (Q: 6.69, $d.f.=4$; $I^2=40.02\%$; $P=0.154$). Sensitivity analyses also indicated that there was no study strongly affected the primary endpoint results as a whole of AOM/PP RCTs or among PP RCTs only (MD = −11.1 to −8.7; all 95% CIs were also in significant ranges), indicating a clearly even superiority of AOM and/or PP over placebo treatment. The

Fig. 2



Search strategy and result of PP clinical trials. Of the 282 records on PP trials, nine studies were selected for a thorough review of the paper and thereafter four RCTs (Gopal *et al.*, 2010; Nasrallah *et al.*, 2010; Alphs *et al.*, 2011; Takahashi *et al.*, 2013) were included in the present study. PBO, placebo; PD, pharmacodynamic; PK, pharmacokinetic; PP, paliperidone palmitate; RCT, randomized clinical trial.

Egger test showed no significant difference ($P = 0.583$) among all studies, indicating no publication bias.

Indirect treatment comparisons (aripiprazole once monthly vs. paliperidone palmitate)

The MD (−6.4) in the primary efficacy endpoint was significantly different between AOM and PP in the primary efficacy endpoint, favouring AOM over PP (95% CIs: −11.402 to −1.358). The MD (−6.9) in the primary efficacy endpoint was significantly different between AOM and PP excluding one Asian RCT, favouring AOM over PP (95% CIs: −12.333 to −1.547). The MD (−5.4) in the primary efficacy endpoint was also significantly different between AOM and Asian PP RCT only, favouring AOM over PP (95% CIs: −11.481 to 0.681; Table 2).

Noninferiority test (aripiprazole once monthly against paliperidone palmitate)

As for the noninferiority test of AOM against PP, the Δs were 2.7 and 4.1 for the primary efficacy endpoint between AOM and all four PP RCTs and between AOM and three PP RCTs excluding one Asian RCT, respectively, as calculated using the Δ formula. The upper limits of 95% CIs were −1.358 from the result of ITC between AOM and all four PP RCTs and −1.547 from the result of ITC between AOM and three PP RCTs excluding Asian RCT, respectively. Therefore, the Δ (2.7) was larger than 95% CI upper limit (−1.358) of the result from ITC between AOM and all four PP RCTs and the Δ (4.1) was also larger than 95% CI upper limit (−1.547) of the result from ITC between AOM and three PP RCTs excluding one Asian RCT, respectively. In addition, the Δ was 1.4 for the primary efficacy endpoint between AOM and Asian PP RCT only as calculated using the Δ formula. The upper limit of 95% CIs was 0.681 from ITC between AOM and Asian PP RCT only; therefore, the upper limit of 95% CI 0.681 was lower than Δ (1.4). Given these results, AOM was not inferior to PP, but it was superior to PP in the primary efficacy endpoint in all comparisons (AOM vs. four PP

Table 1 Characteristics of randomized, placebo-controlled clinical trials included in the present study

References	Funding	LAI	Race	DS (months)	BMI (LAI/PBO) (kg/m^2)	Dose	Primary efficacy endpoint	Number (LAI/PBO)[a]	Sex (male) (LAI/PBO)	Age (LAI/PBO)	Onset age (LAI/PBO) (years)	DUS	PANSS total score (B) (LAI/PBO)	ED rates (LAI : PBO)
Kane *et al.* (2014)	OPDC/LB	AOM	W (2 As only)	12	28.4 (5.6)/28.5 (5.2)	400 mg or PBO	Change of PANSS total score[b]	168/172	130 (77.4)/139 (80.8)	42.1 (11.0)/42.7 (10.9)	24.4 (8.3)/24.0 (9.3)	12	103.0 (11.3)/104.0 (11.0)	60/168 (35.7) : 87/172 (50.6)
Alphs *et al.* (2011)	OJSA	PP	W > A	12	NR/NR	156 mg (paliperidone 100 mg eq)	Change of PANSS total score[b]	72/83	47 (65.3)/56 (67.5)	38.4 (10.6)/40.3 (11.2)	25.8 (8.7)/24.9 (8.1)	13	94.5 (7.9)/92.6 (9.2)	36/72 (50.0) : 51/83 (61.4)
Gopal *et al.*, 2010	JJ	PP	W > > A	12	29 (8.1)/28 (7.9)	156 mg (paliperidone 100 mg eq)	Change of PANSS total score[b]	97/136	61 (64.9)/94 (71.2)	39 (10.7)/41 (11.0)	25.0 (8.1)/26.0 (8.7)	13	90 (11.7)/92 (12.6)	44/97 (45.4) : 85/136 (62.5)
Nasrallah *et al.* (2010)	JJ	PP	W (2 As only)	12	27.7 (6.4)/27.5 (6.6)	156 mg (paliperidone 100 mg eq)	Change of PANSS total score[b]	131/127	85 (64.8)/78 (62.4)	42.3 (10.7)/41.1 (11.8)	27.4 (9.6)/26.9 (9.6)	13	90.8 (11.7)/90.7 (12.2)	56/131 (42.7) : 79/127 (62.2)
Takahashi *et al.* (2013)	JP	PP	A only	12	23.5 (4.37)/23.9 (3.92)	117 mg (paliperidone 75 mg eq)	Change of PANSS total score[b]	160/164	101 (63.5)/83 (50.6)	46 (13.6)/44 (12.4)	27.4 (13)[c]/27.5 (12)[c]	13	85.7 (14.57)/83.5 (15.18)	64/160 (40.0): 109/164 (66.5)

A, Asian; AE, adverse event; AOM, aripiprazole once monthly; DS, duration of stabilization before study entry; DUS, duration of study (weeks); ED, early dropout; eq, equivalent; JJ, Johnson and Johnson Pharmaceutical Research and Development; LLC.; JP, Janssen Pharmaceutical KK, Japan; LAI, long-acting injection; LB, H. Lundbeck A/S.; OJSA, Ortho-McNeil Janssen Scientific Affairs; OPDC, Otsuka Pharmaceutical Development and Commercialization Inc.; PANSS, Positive and Negative Syndrome Scale; PBO, placebo; PP, paliperidone palmitate; W, western (White, Black and others).

[a]Randomized number.

[b]From baseline to the end of treatment.

[c]Not reported in the study and thus calculated on the basis of onset age and duration of illness.

RCTs or vs. three PP without Asian study or vs. Asian PP RCT only).

Safety and tolerability

Overall

The ED rates were numerically lesser in the AOM/PP group than in placebo treatment. Across all the studies, frequently reported TEAEs were weight gain, headache, akathisia, constipation, cough, dyspepsia, nervousness, pain on injection site, sedation and toothache in AOM treatment, whereas they were insomnia, pain on injection site, nasopharyngitis, extrapyramidal symptoms (EPS), anxiety, constipation, headache, akathisia, nausea, weight gain, tension and toothache in PP treatment. Among these TEAEs, weight gain, akathisia and pain on injection site were significantly more developed in AOM than placebo, whereas headache, insomnia, pain on injection site, EPS and nausea were significantly more developed in PP than placebo.

Meta-analyses (aripiprazole once monthly and/or paliperidone palmitate vs. placebo)

Table 3 presents the results of meta-analyses in detail. The overall ED rate was significantly different between AOM and placebo (OR: 0.543; 95% CIs: 0.351–0.838), favouring AOM over placebo. It was also significantly different between PP and placebo, favouring PP over placebo (OR: 0.444; 95% CIs: 0.342–0.575). This trend was also replicated in the meta-analysis between AOM/PP as a whole and placebo (OR: 0.468; 95% CIs: 0.342–0.575), indicating the superiority of LAIAs over placebo.

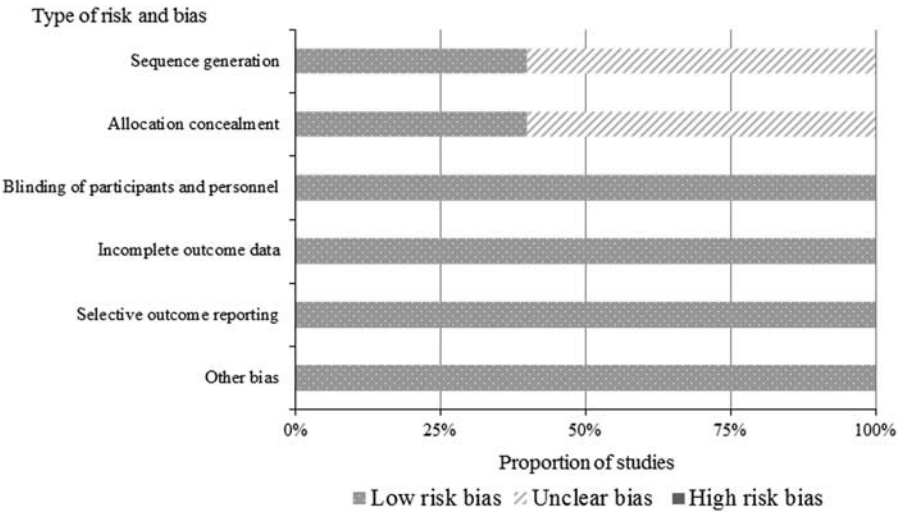
As for the ED rate because of TEAEs, there was a small but statistical difference between PP and placebo (OR: 0.590; 95% CIs 0.352–0.987), whereas there was no significant difference between AOM and placebo (OR: 0.532; 95% CIs: 0.207–1.368). It was significantly different between AOM/PP as a whole and placebo (OR: 0.578; 95% CIs: 0.364–0.919), indicating the superiority of LAIAs over placebo (Table 3).

As for the ED rate because of lack of efficacy, there was a significant difference between AOM and placebo (OR: 0.188; 95% CIs: 0.094–0.374), favouring AOM over placebo. It was also significantly different between PP and placebo (OR: 0.477; 95% CIs: 0.349–0.653), favouring PP over placebo. It was significantly different between AOM/PP as a whole and placebo (OR: 0.315; 95% CIs: 0.127–0.782), indicating the superiority of LAIAs over placebo (Table 3).

Indirect treatment comparisons (aripiprazole once monthly vs. paliperidone palmitate)

The overall EDs rate was not significantly different between AOM and PP (OR: 1.223; 95% CIs: 0.737–2.03), indicating similar safety/tolerability profile between two LAIAs. As for the ED rate because of TEAEs, it was also

Fig. 3



The overall risks of bias for all included studies. The risk of bias was considered low or unclear in all studies on the basis of evaluations of all domains, and no study was scored as presenting a high risk of bias in all domains.

Table 2 Meta-analysis (vs. placebo) and indirect treatment comparisons (aripiprazole once monthly vs. paliperidone palmitate) on the mean difference of the changes in the Positive and Negative Syndrome Scale total score from baseline to the end of treatment

		Statistics	
	References	MD	95% CIs
Meta-analyses[a]			
LAI type			
AOM	Kane *et al.* (2014)	−15.100	−19.399 to −10.801
AOM subtotal (vs. placebo)		−15.100	−19.399 to −10.801
PP			
	Alphs *et al.* (2011)	−8.900	−16.701 to −1.099
	Gopal *et al.* (2010)	−6.800	−11.995 to −1.605
	Nasrallah *et al.* (2010)	−9.100	−14.055 to −4.145
	Takahashi *et al.* (2013)	−9.700	−14.000 to −5.400
PP subtotal (vs. placebo)		−8.722	−11.319 to −6.125
LAIs total (AOM/PP vs. placebo)		−10.427	−12.650 to −8.204
ITCs[b]			
AOM vs. all PP		−6.38	−11.402 to −1.358
AOM vs. PP excluding Asian RCT		−6.94	−12.333 to −1.547
AOM vs. Asian PP only		−5.4	−11.481–0.681

The MD in the primary efficacy endpoint was significantly higher in AOM and PP than in the placebo treatment; The MD in the primary efficacy endpoint was significantly higher in AOM than in the PP treatment.
ITC, indirect treatment comparison; AOM, aripiprazole once monthly; PP, paliperidone palmitate; RCT, randomized double-blind, placebo-controlled clinical trial; MD, mean difference.
[a]Random-effect model meta-analysis.
[b]Results of the indirect comparison using ITC software [ITC2, Canadian Agency for Drugs and Technologies in Health (CADTH), Ottawa, Canada].

not significantly different between AOM and PP (OR: 0.902; 95% CIs: 0.308–2.644) (Table 3). However, there was a significant difference in the ED rate because of lack of efficacy between AOM and PP (OR: 0.394; 95% CIs: 0.185–0.841), favouring AOM over PP (Table 3). Even on excluding an Asian PP study (Takahashi *et al.*, 2013) because of its geographic preponderance, comparison of EDs rates between AOM and PP, EDs rates between AOM and PP were similar as presented in the primary analyses (overall ED, OR: 1.069; 95% CIs: 0.624–1.830; TEAEs ED, OR: 0.857; 95% CIs: 0.192–3.821. Lack of efficacy ED, OR: 0.367; 95% CIs: 0.153–0.879).

Discussion

ITCs may be a valid methodological approach to assessing relative efficacy and safety/tolerability when there is a lack of direct comparative studies. Indeed, numerous ITCs have been conducted in various therapeutic areas including psychiatry (Fisher *et al.*, 2001; Kunitomi *et al.*, 2014). Ideally, adjusted ITC would supplement the results from head-to-head RCTs and would aid

Table 3 Meta-analysis (vs. placebo) and indirect treatment comparisons (aripiprazole once monthly vs. paliperidone palmitate) on the early dropouts because of any reasons, treatment emergent adverse events and lack of efficacy

		Statistics	
	References	OR	95% CIs
Overall ED rates			
Meta-analyses[a]			
LAI type			
AOM	Kane *et al.* (2014)	0.543	0.351–0.838
AOM subtotal (vs. placebo)			
PP	Alphs *et al.* (2011)	0.627	0.331– 1.189
	Gopal *et al.* (2010)	0.498	0.293– 0.846
	Nasrallah *et al.* (2010)	0.454	0.276–0.747
	Takahashi *et al.* (2013)	0.336	0.214–0.529
PP subtotal (vs. placebo)		0.444	0.342–0.575
AOM/PP total (vs. placebo)		0.468	0.342–0.575
ITCs (AOM vs. PP)[b]		1.223	0.737–2.03
ED rates d/t TEAEs			
Meta-analyses[a]			
LAI type			
AOM	Kane *et al.* (2014)	0.532	0.207–1.368
AOM subtotal (vs. placebo)		0.532	0.207–1.368
PP	Alphs *et al.* (2011)	1.169	0.390–3.508
	Gopal *et al.* (2010)	0.199	0.044–0.904
	Nasrallah *et al.* (2010)	0.714	0.241–2.119
	Takahashi *et al.* (2013)	0.537	0.314–0.919
PP subtotal (vs. placebo)		0.590	0.352–0.987
AOM/PP total (vs. placebo)		0.578	0.364–0.919
ITCs (AOM vs. PP)[b]		0.902	0.308–2.644
ED rates d/t lack of efficacy			
Meta-analyses[a]			
LAI type			
AOM	Kane *et al.*, 2014	0.188	0.094–0.374
AOM subtotal (vs. placebo)		0.188	0.094–0.374
PP	Alphs *et al.*, 2011	0.593	0.280–1.259
	Gopal *et al.*, 2010	0.671	0.379–1.188
	Nasrallah *et al.*, 2010	0.348	0.193–0.629
	Takahashi *et al.*, 2013	0.400	0.225–0.709
PP subtotal (vs. placebo)		0.477	0.349–0.653
AOM/PP total (vs. placebo)		0.315	0.127–0.782
ITCs (AOM vs. PP)[b]		0.394	0.185–0.841

The overall and TEAEs EDs rate were not significantly different between AOM and PP treatment, indicating a similar safety/tolerability profile between two LAIAs, whereas the ED rate because of lack of efficacy was significantly different between AOM and PP treatment, favouring AOM over PP.

AOM, aripiprazole once monthly; CIs, confidence intervals; ED, early dropout; ES, effect size; ITC, indirect treatment comparison; OR, odd ratio; PP, paliperidone palmitate; RD, risk difference; RR, relative risk; TEAE, treatment emergent adverse event.

[a]Random-effect model meta-analysis.

[b]Results of the indirect comparison using ITC software [ITC2, Canadian Agency for Drugs and Technologies in Health (CADTH), Ottawa, Canada].

cross-examinations on the validity and applicability of such trials (Song *et al.*, 2008). Indeed, a number of ITCs have also been conducted to compare the relative efficacy and safety among oral antipsychotics in the treatment of schizophrenia. For instance, such studies have compared aripiprazole versus olanzapine (Kunitomi *et al.*, 2014), olanzapine versus risperidone (Sauriol *et al.*, 2001) and FGAs versus. SGAs (Rabinowitz *et al.*, 2009). The relative efficacy and safety of oral antipsychotics observed through previous ITCs appeared to be similar to that found in direct comparative RCTs among these agents (Sauriol *et al.*, 2001), thereby contributing towards the establishment of the validity, usefulness and informativeness of ITCs. Our ITC is the first work to investigate the relative efficacy between AOM and PP in the short-term treatment of schizophrenia in which the superiority of AOM over PP in terms of MD in the change of the PANSS total score from baseline was found.

In the meta-analyses, the present study showed statistically superior efficacy of both AOM and PP both combined and as individual LAIAs as evidenced in numerous RCTs of each LAIA over placebo for treating schizophrenia. In the present ITC, the primary efficacy endpoint (change of the PANSS total score from baseline as presented by MD with 95% CIs), was superior in AOM compared with PP, and this remained the case with the exclusion of some PP RCTs in the sensitivity analyses, suggesting consistent superiority of AOM over PP in the primary efficacy endpoint. The MD was −6.4 in the primary efficacy endpoint between AOM and PP, favouring AOM over PP. This is consistent with a recent 28-week, randomized, noninferiority, open-label, rater-blinded, head-to-head RCT (Qualify study) between AOM and PP (Naber *et al.*, 2015) that reported the superiority of AOM over PP on the HRQoL measured by Quality-of-Life Scale (QLS). According to the Qualify study, the MD of the QLS total score from baseline (4.7; 95% CI: 0.32–9.02) was significantly greater in the AOM group (7.5) than the PP group (2.8), which was approximately in line with the suggested minimal clinically important difference (MCID) of 5.3 points (Falissard *et al.*, 2016). In the Qualify study, there were also significantly more improvements in the Clinical Global Impression-Severity scale (CGI-S) in AOM than in PP (MD = −0.28; 95% CI: −0.48 to −0.09]. The predefined subgroup analyses not only showed a consistent pattern of significance but also a more profound difference favouring AOM over PP in patients up to 35years in both measures of QLS (MD = 10.7; 95% CIs: 0.70–20.7) and CGI-S (MD = −0.44; 95% CIs: −0.83 to −0.06). When comparing the change in the PANSS total score with the change in the CGI-S total score, it has been estimated that an absolute reduction of the PANSS total score by ~15 points corresponds with a reduction in the CGI severity score by one point (Leucht *et al.*, 2006). Applying this to the results of the Qualify study (Naber *et al.*, 2015), the difference in the change of the CGI-S score of −0.3 between AOM and PP is equivalent to a difference in the changes in the PANSS total score of −4.5 (Naber *et al.*, 2015).

This is not dissimilar to the MD of the change in the PANSS total score (−6.4) between AOM and PP in the present ITC. In other words, our finding of a greater reduction in the PANSS total score with AOM in the present ITC is consistent with the findings from the head-to-head study (Qualify study) between AOM and PP. In addition, the ED because of lack of efficacy was also in favour of AOM versus PP.

There is some suggestion that AOM and PP have different efficacy profiles, with paliperidone more effective in patients with predominantly positive symptoms (Heres *et al.*, 2014) and aripiprazole more effective in patients with a short duration of illness (Takaesu *et al.*, 2016). Given the above, it is possible that the small imbalance in baseline symptom severity may have favoured AOM over PP in the present ITC. In addition, this may also raise a need for a more detailed profiling in the use of antipsychotics to achieve the targets of precision medicine on the basis of clinical/biological predictors (DeLisi and Fleischhacker, 2016).

Meanwhile, the baseline severity of psychopathology measured by the PANSS total score has been known to influence the treatment response and this could again have favoured AOM. According to a recent meta-analysis investigating the relationship between baseline severity of schizophrenia using RCTs, it was found that greater baseline severity was significantly correlated with a greater magnitude of the differences between active treatment and placebo (Furukawa *et al.*, 2015). However, the mean baseline PANSS total scores of 103 and 90 for AOM and PP, respectively, both fell within the range of 'marked psychopathology' (Leucht *et al.*, 2005). Furthermore, according to the recent meta-analysis (Furukawa *et al.*, 2015), the current practice of setting the threshold to a PANSS total score of 75 may be sufficient and justifiable to strike a balance between patient recruitment and signal detection; hence, the recruitment of the samples in both AOM and PP was in line with this methodological suggestion for rigorous RCT. Finally, despite the difference in the baseline PANSS total score between AOM and PP, the scores fell within the range of SDs (7.9–15.6 in PP group).

ITC has the strength of partially retaining of randomization and is therefore less likely to be influenced by patients' characteristics unrelated to the treatment (Bucher *et al.*, 1997; Wells *et al.*, 2009). Thus, it may be conjectured that the difference in terms of baseline PANSS total scores between patients included in AOM and PP RCTs was not considerably large. However, despite no heterogeneities being found among RCTs of AOM and PP, it is possible that an undetected and hidden subpopulation might exist in the AOM and PP RCTs included in the present ITC. Interestingly, the MD (−15.1) between AOM and placebo was higher than those (−5.2 to −12.7) from registry RCTs of oral aripiprazole (10–30 mg/day) in the treatment of schizophrenia (4–6 weeks), whereas the MD between PP and placebo of −8.7 was lower than those (−11.0 to −16.6) from registry RCTs of oral paliperidone ER (3–12 mg/day) (Meltzer *et al.*, 2008) in the treatment of schizophrenia (6 weeks). Given the above, we could not completely exclude the possibility that the difference in efficacy in our ITC might be at least partially influenced by factors not identified in the AOM and PP RCTs (e.g. anxiety, negative symptoms, cognitive impairment, etc.). As an example, in the AOM study, patients with a good response were recruited (Kane *et al.*, 2014), whereas in some PP studies, only resistant patients were excluded, thus suggesting that the inclusion of partially resistant patients in PP studies may have reduced the overall observed efficacy. Another point is that the population may be more chronic and relatively stable compared with those with true acute phase schizophrenia as LAIAs are usually used for maintenance and relapse prevention for schizophrenia. In addition, previous head-to-head studies of LAIAs have consistently reported no differential efficacy between individual LAIAs (Pandina *et al.*, 2011; Fleischhacker *et al.*, 2012). Finally, the relative dose relationship among LAIAs has yet to be elucidated. The AOM dose was 400 mg 4-weekly, whereas the mean dose for PP was 156 mg 4-weekly. We could not exactly estimate oral equivalents of AOM and PP doses. Hence, potential efficacy differences between AOM and PP need to be investigated in future research using more adequately powered and rigorously designed randomized studies.

The concept of the MCID has been useful in assessing the clinical relevance of changes in standardized instrument scores (Hermes *et al.*, 2012). According to the previous study utilizing the Clinical Antipsychotic Trials of Intervention Effectiveness schizophrenia trial data (Hermes *et al.*, 2012), the MCID was found to be ~15 points (34% reduction) in the PANSS total score from baseline in the treatment of schizophrenia. The MD of the primary efficacy endpoint was −6.4 point, favouring AOM (25% reduction) over PP (13.7% reduction) in the present study, indicating that the clinical importance of such an MD of −6.4 points (11.3% difference in PANSS total score reduction from baseline) between AOM and PP might be questionable for its utility in routine clinical practice.

In terms of treatment acceptability, ED because of any reason was not significantly different between AOM and PP, indicating that both LAIAs are similarly safe and tolerable in the treatment of schizophrenia. Overall incidences of TEAEs were similar between the two LAIAs. However, unique TEAS to individual LAIA were also noted. Weight gain, akathisia and pain on injection site were significantly more frequent in AOM than placebo, whereas headache, insomnia, pain on injection site, EPS and nausea were significantly more frequent in PP than placebo. In contrast, ED because of TEAEs was not

significantly different between AOM and PP, whereas discontinuation because of lack of efficacy was significantly higher in PP than in AOM. Once again, although this finding needs to be interpreted with caution, it is in line with the primary efficacy endpoint analyses of the present ITCs.

It has been proposed that LAIs may be usable even in the early stage of schizophrenia, rather than being specifically reserved for non-adherent or chronic schizophrenia patients, suggesting a possible treatment shift of LAIs as the first-line treatment agents for early schizophrenia patients as well (Stahl, 2014). Hence, the present result is also in line with the currently proposed strategy of an early intervention with LAIs and should provide at least some preliminary evidence of the relative efficacy among LAIs for the treatment of schizophrenia.

Our study has the following clear limitations. First, there are inherent pitfalls with ITCs in terms of methodology and data interpretation (Jansen *et al.*, 2011). Transparent establishment of homogeneity, elaborated assessment of study quality, firm internal validity, consistency for data acquisition, proper statistical approach, prudent interpretation of results, complete literature search, confirmation of study design similarity and accurate assumptions for adjusted ITCs are difficult tasks as observed commonly in meta-analyses (Song *et al.*, 2009; Donegan *et al.*, 2010). Indeed, previous reviews have explored some methodological flaws and weaknesses in previous ITCs (Song *et al.*, 2009; Donegan *et al.*, 2010). In particular, with respect to clinical implications and interpretations, findings from ITCs cannot supersede those from direct comparisons and they are not the same evidence; this well-defined difference should be taken into consideration to avoid misinterpretation. Hence, direct evidence from good-quality RCTs is usually recommended ahead of ITCs if available data resources are sufficient and adequate. A proper combination of evidences from both direct comparisons and ITCs along with empirical data would maximize the probability of achieving clear and precise estimates for better clinical decisions in routine practice (Song *et al.*, 2003, 2009; Glenny *et al.*, 2005). Second, the usefulness of subgroup analysis and meta-regression may be limited, given that the number of trials investigating the relative efficacy and safety between new therapeutic agents is particularly small. Third, methodologically inappropriate use of the placebo arm within the same trial is vulnerable to underestimate direct effects of certain study drug as well as reduces precision of comparison (Song *et al.*, 2009; Donegan *et al.*, 2010). Fourth, we included only short-term RCTs of AOM and PP and thus our findings cannot be generalized to maintenance treatment of schizophrenia. Treatment of schizophrenia is not short term based, but usually requires continuous and maintenance therapy; thus, long-term RCTs investigating potential differences between AOM and PP should be included in future ITCs. Fifth, although the study design and methodology were similar across all PP RCTs, some differences were found among PP RCTs. One notable finding was that the Asian PP study showed a smaller reduction in the PANSS total score from baseline (PP = −3.5 vs. placebo = 6), although exclusion of this study from the ITC analyses did not alter the primary results showing AOM superiority over PP (sensitivity analysis). Indeed, Asian psychiatrists tend to assess psychotic symptoms significantly lower than western psychiatrists as observed in the measurements of PANSS total scores as well as all three subscales and most individual items (Aggarwal *et al.*, 2011). In addition, the Asian PP study had the lowest baseline PANSS total scores and the longest duration of illness compared with those from the other PP RCTs. Such subtle geographic differences were also reported in previous meta-analyses in other therapeutic areas (Pae *et al.*, 2015a, 2015b). Therefore, international, multicentre trials should pay more attention to the design, recruitment strategy of samples and conduct of studies. Sixth, the number of RCTs and participants was too small to draw meaningful conclusions for clinical practice and therefore the present results should be interpreted with caution. In particular, only one study for AOM was included. Seventh, we only assessed score psychopathology changes as measured by one specific rating scale. As the primary goal of treatment for schizophrenia includes not only reduction of psychotic symptoms but also recovery of functional impairment and health-related quality of life, these outcome measures need to be investigated in future studies. Finally, it is possible that clinical heterogeneity (e.g. study location, baseline parameters, etc.) including unidentified differences in study and population characteristics could confound the results.

Conclusion

Because of a dearth of direct comparative RCTs between AOM and PP, it is premature to conclude that AOM may be superior to PP in terms of efficacy. However, within the context of the inherent limitations of the current analysis, our results may suggest that there may be relative advantages for AOM over PP in the short-term treatment of schizophrenia. Adequately powered, longer-term trials directly comparing AOM and PP are evidently necessary aid clinicians in making the most appropriate treatment choices for individual patients.

Acknowledgements

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

Conflicts of interest

There are no conflicts of interest.

References

Aggarwal NK, Tao H, Xu K, Stefanovics E, Zhening L, Rosenheck RA (2011). Comparing the PANSS in Chinese and American inpatients: cross-cultural psychiatric analyses of instrument translation and implementation. *Schizophr Res* **132**:146–152.

Alphs L, Bossie CA, Sliwa JK, Ma YW, Turner N (2011). Onset of efficacy with acute long-acting injectable paliperidone palmitate treatment in markedly to severely ill patients with schizophrenia: post hoc analysis of a randomized, double-blind clinical trial. *Ann Gen Psychiatry* **10**:12.

Barnes TR, Schizophrenia Consensus Group of British Association for Psychopharmacology (2011). Evidence-based guidelines for the pharmacological treatment of schizophrenia: recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* **25**:567–620.

Birchwood M, Smith J, Macmillan F, Hogg B, Prasad R, Harvey C, *et al.* (1989). Predicting relapse in schizophrenia: the development and implementation of an early signs monitoring system using patients and families as observers, a preliminary investigation. *Psychol Med* **19**:649–656.

Bucher HC, Guyatt GH, Griffith LE, Walter SD (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* **6**:683–691.

Caseiro O, Perez-Iglesias R, Mata I, Martinez-Garcia O, Pelayo-Teran JM, Tabares-Seisdedos R, *et al.* (2012). Predicting relapse after a first episode of non-affective psychosis: a three-year follow-up study. *J Psychiatr Res* **46**: 1099–1105.

Chi MH, Hsiao CY, Chen KC, Lee LT, Tsai HC, Hui Lee I, *et al.* (2016). The readmission rate and medical cost of patients with schizophrenia after first hospitalization – a 10-year follow-up population-based study. *Schizophr Res* **170**:184–190.

Cooper D, Moisan J, Gregoire JP (2007). Adherence to atypical antipsychotic treatment among newly treated patients: a population-based study in schizophrenia. *J Clin Psychiatry* **68**:818–825.

Correll CU, Citrome L, Haddad PM, Lauriello J, Olfson M, Calloway SM, *et al.* (2016). The use of long-acting injectable antipsychotics in schizophrenia: evaluating the evidence. *J Clin Psychiatry* **77** (Suppl 3):1–24.

De Hert M, Sermon J, Geerts P, Vansteelandt K, Peuskens J, Detraux J (2015). The use of continuous treatment versus placebo or intermittent treatment strategies in stabilized patients with schizophrenia: a systematic review and meta-analysis of randomized controlled trials with first- and second-generation antipsychotics. *CNS drugs* **29**:637–658.

DeLisi LE, Fleischhacker WW (2016). How precise is precision medicine for schizophrenia? *Curr Opin Psychiatry* **29**:187–189.

Donegan S, Williamson P, Gamble C, Tudur-Smith C (2010). Indirect comparisons: a review of reporting and methodological quality. *PloS one* **5**:e11054.

Emsley R, Chiliza B, Asmal L, Harvey BH (2013a). The nature of relapse in schizophrenia. *BMC Psychiatry* **13**:50.

Emsley R, Oosthuizen P, Koen L, Niehaus D, Martinez L (2013b). Comparison of treatment response in second-episode versus first-episode schizophrenia. *J Clin Psychopharmacol* **33**:80–83.

European Medicines Agency (2013). AbilifyMaintena (EMA/489804/2013). London, UK.

Falissard B, Sapin C, Loze JY, Landsberg W, Hansen K (2016). Defining the minimal clinically important difference (MCID) of the Heinrichs-carpenter quality of life scale (QLS). *Int J Methods Psychiatr Res* **25**:101–111.

FDA. Abilify Maintena Kit (2013). (NDA 202971). Silver Spring, MD.

Fisher LD, Gent M, Buller HR (2001). Active-control trials: how would a new agent compare with placebo? a method illustrated with clopidogrel, aspirin, and placebo. *Am Heart J* **141**:26–32.

Fleischhacker WW, Gopal S, Lane R, Gassmann-Mayer C, Lim P, Hough D, *et al.* (2012). A randomized trial of paliperidone palmitate and risperidone long-acting injectable in schizophrenia. *Int J Neuropsychopharmacol* **15**:107–118.

Furukawa TA, Levine SZ, Tanaka S, Goldberg Y, Samara M, Davis JM, *et al.* (2015). Initial severity of schizophrenia and efficacy of antipsychotics: participant-level meta-analysis of 6 placebo-controlled studies. *JAMA Psychiatry* **72**:14–21.

Fusar-Poli P, Kempton MJ, Rosenheck RA (2013). Efficacy and safety of second-generation long-acting injections in schizophrenia: a meta-analysis of randomized-controlled trials. *Int Clin Psychopharmacol* **28**:57–66.

Galletly C, Castle D, Dark F, Humberstone V, Jablensky A, Killackey E, *et al.* (2016). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. *Aust N Z J Psychiatry* **50**:410–472.

Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, *et al.*, International Stroke Trial Collaborative Group (2005). Indirect comparisons of competing interventions. *Health Technol Assess* **9**:1–134.

Gopal S, Hough DW, Xu H, Lull JM, Gassmann-Mayer C, Remmerie BM, *et al.* (2010). Efficacy and safety of paliperidone palmitate in adult patients with acutely symptomatic schizophrenia: a randomized, double-blind, placebo-controlled, dose-response study. *Int Clin Psychopharmacol* **25**:247–256.

Hahn S (2012). Understanding noninferiority trials. *Korean J Pediatr* **55**:403–407.

Hasan A, Falkai P, Wobrock T, Lieberman J, Glenthoj B, Gattaz WF, *et al.* (2012). World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Schizophrenia, part 1: update 2012 on the acute treatment of schizophrenia and the management of treatment resistance. *World J Biol Psychiatry* **13**:318–378.

Hasan A, Falkai P, Wobrock T, Lieberman J, Glenthoj B, Gattaz WF, *et al.* (2013). World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of schizophrenia, part 2: update 2012 on the long-term treatment of schizophrenia and management of antipsychotic-induced side effects. *World J Biol Psychiatry* **14**:2–44.

Heres S, Don L, Herceg M, Bidzan L, Blanc M, Siracusano A, *et al.* (2014). Treatment of acute schizophrenia with paliperidone ER: predictors for treatment response and benzodiazepine use. *Prog Neuropsychopharmacol Biol Psychiatry* **48**:207–212.

Hermes ED, Sokoloff D, Stroup TS, Rosenheck RA (2012). Minimum clinically important difference in the Positive and Negative Syndrome Scale with data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). *J Clin Psychiatry* **73**:526–532.

Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, *et al.* (2011). Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value Health* **14**:429–437.

Hogarty GE, Ulrich RF (1998). The limitations of antipsychotic medication on schizophrenia relapse and adjustment and the contributions of psychosocial treatment. *J Psychiatr Res* **32**:243–250.

Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, *et al.* (2011). Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* **14**:417–428.

Kane JM, Peters-Strickland T, Baker RA, Hertel P, Eramo A, Jin N, *et al.* (2014). Aripiprazole once-monthly in the acute treatment of schizophrenia: findings from a 12-week, randomized, double-blind, placebo-controlled study. *J Clin Psychiatry* **75**:1254–1260.

Kapur S, Agid O, Mizrahi R, Li M (2006). How antipsychotics work – from receptors to reality. *NeuroRx* **3**:10–21.

Kay SR, Fiszbein A, Opler LA (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* **13**:261–276.

Kirson NY, Weiden PJ, Yermakov S, Huang W, Samuelson T, Offord SJ, *et al.* (2013). Efficacy and effectiveness of depot versus oral antipsychotics in schizophrenia: synthesizing results across different research designs. *J Clin Psychiatry* **74**:568–575.

Kishimoto T, Nitta M, Borenstein M, Kane JM, Correll CU (2013). Long-acting injectable versus oral antipsychotics in schizophrenia: a systematic review and meta-analysis of mirror-image studies. *J Clin Psychiatry* **74**:957–965.

Kunitomi T, Hashiguchi M, Mochizuki M (2014). Indirect comparison analysis of efficacy and safety between olanzapine and aripiprazole for schizophrenia. *Br J Clin Pharmacol* **77**:767–776.

Lafeuille MH, Dean J, Carter V, Duh MS, Fastenau J, Dirani R, *et al.* (2014). Systematic review of long-acting injectables versus oral atypical antipsychotics on hospitalization in schizophrenia. *Curr Med Res Opin* **30**:1643–1655.

Lehman AF, Lieberman JA, Dixon LB, McGlashan TH, Miller AL, Perkins DO, *et al.* (2004). Practice guideline for the treatment of patients with schizophrenia, second edition. *Am J Psychiatry* **161** (2 Suppl):1–56.

Leonard BE, Schwarz M, Myint AM (2012). The metabolic syndrome in schizophrenia: is inflammation a contributing cause? *J Psychopharmacol* **26** (5 Suppl):33–41.

Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR (2005). What does the PANSS mean? *Schizophr Res* **79**:231–238.

Leucht S, Kane JM, Etschel E, Kissling W, Hamann J, Engel RR (2006). Linking the PANSS, BPRS, and CGI: clinical implications. *Neuropsychopharmacology* **31**:2318–2325.

Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, *et al.* (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* **382**:951–962.

Lieberman JA, Alvir JM, Koreen A, Geisler S, Chakos M, Sheitman B, *et al.* (1996). Psychobiologic correlates of treatment response in schizophrenia. *Neuropsychopharmacology* **14** (3 Suppl):13S–21S.

Lieberman JA, Perkins D, Belger A, Chakos M, Jarskog F, Boteva K, *et al.* (2001). The early stages of schizophrenia: speculations on pathogenesis, pathophysiology, and therapeutic approaches. *Biol Psychiatry* **50**:884–897.

Lindstrom E, Eberhard J, Neovius M, Levander S (2007). Costs of schizophrenia during 5 years. *Acta Psychiatr Scand Suppl* **435**:33–40.

McGrath J, Saha S, Chant D, Welham J (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev* **30**:67–76.

Meltzer HY, Bobo WV, Nuamah IF, Lane R, Hough D, Kramer M, *et al.* (2008). Efficacy and tolerability of oral paliperidone extended-release tablets in the treatment of acute schizophrenia: pooled data from three 6-week, placebo-controlled studies. *J Clin Psychiatry* **69**:817–829.

Naber D, Hansen K, Forray C, Baker RA, Sapin C, Beillat M, *et al.* (2015). Qualify: a randomized head-to-head study of aripiprazole once-monthly and paliperidone palmitate in the treatment of schizophrenia. *Schizophr Res* **168**: 498–504.

Nasrallah HA, Gopal S, Gassmann-Mayer C, Quiroz JA, Lim P, Eerdekens M, *et al.* (2010). A controlled, evidence-based trial of paliperidone palmitate, a long-acting injectable antipsychotic, in schizophrenia. *Neuropsychopharmacology* **35**:2072–2082.

Pae CU, Wang SM, Han C, Lee SJ, Patkar AA, Masand PS, *et al.* (2015a). Vortioxetine, a multimodal antidepressant for generalized anxiety disorder: a systematic review and meta-analysis. *J Psychiatr Res* **64**:88–98.

Pae CU, Wang SM, Han C, Lee SJ, Patkar AA, Masand PS, *et al.* (2015b). Vortioxetine: a meta-analysis of 12 short-term, randomized, placebo-controlled clinical trials for the treatment of major depressive disorder. *J Psychiatry Neurosci* **40**:174–186.

Pandina G, Lane R, Gopal S, Gassmann-Mayer C, Hough D, Remmerie B, *et al.* (2011). A double-blind study of paliperidone palmitate and risperidone long-acting injectable in adults with schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* **35**:218–226.

Rabinowitz J, Levine SZ, Barkai O, Davidov O (2009). Dropout rates in randomized clinical trials of antipsychotics: a meta-analysis comparing first- and second-generation drugs and an examination of the role of trial design features. *Schizophr Bull* **35**:775–788.

Saha S, Chant D, Welham J, McGrath J (2005). A systematic review of the prevalence of schizophrenia. *PLoS Med* **2**:e141.

Sapin C, Hartry A, Kamat SA, Beillat M, Baker RA, Eramo A (2016). Pharmacoeconomic comparison of aripiprazole once-monthly and paliperidone palmitate from a head-to-head clinical trial in schizophrenia: a US analysis. *Drugs Context* **5**:212301.

Sauriol L, Laporta M, Edwardes MD, Deslandes M, Ricard N, Suissa S (2001). Meta-analysis comparing newer antipsychotic drugs for the treatment of schizophrenia: evaluating the indirect approach. *Clin Ther* **23**:942–956.

Song F, Altman DG, Glenny AM, Deeks JJ (2003). Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* **326**:472.

Song F, Harvey I, Lilford R (2008). Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* **61**:455–463.

Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG (2009). Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* **338**: b1147.

Stahl SM (2014). Long-acting injectable antipsychotics: shall the last be first? *CNS Spectr* **19**:3–5.

Stahl SM, Morrissette DA, Citrome L, Saklad SR, Cummings MA, Meyer JM, *et al.* (2013). 'Meta-guidelines' for the management of patients with schizophrenia. *CNS Spectr* **18**:150–162.

Takaesu Y, Kishimoto T, Murakoshi A, Takahashi N, Inoue Y (2016). Factors associated with discontinuation of aripiprazole treatment after switching from other antipsychotics in patients with chronic schizophrenia: a prospective observational study. *Psychiatry Res* **236**:71–74.

Takahashi N, Takahashi M, Saito T, Iizumi M, Saito Y, Shimizu H, *et al.* (2013). Randomized, placebo-controlled, double-blind study assessing the efficacy and safety of paliperidone palmitate in Asian patients with schizophrenia. *Neuropsychiatr Dis Treat* **9**:1889–1898.

Valenstein M, Ganoczy D, McCarthy JF, Myra Kim H, Lee TA, Blow FC (2006). Antipsychotic adherence over time among patients receiving treatment for schizophrenia: a retrospective review. *J Clin Psychiatry* **67**:1542–1550.

Wells GA, Sultan SA, Chen L, Khan M, Coyle D (2009). *Indirect evidence: indirect treatment comparisons in meta-analysis*: Canadian Agency for Drugs and Technologies in Health.