Aripiprazole for the management of schizophrenia in the Japanese population: a systematic review and meta-analysis of randomized controlled trials

Taro Kishi
Yuki Matsuda
Shinji Matsunaga
Nakao Iwata

Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

Background: We conducted a systematic review and meta-analysis of randomized controlled trials comparing aripiprazole with pooled antipsychotics in Japanese patients with schizophrenia.

Methods: We performed a literature search of data published in PubMed®, the Cochrane Library database, the Japan Medical Abstracts Society, and PsycINFO® up to January 5, 2014. The odds ratio (OR), number-needed-to-harm (NNH), and standardized mean difference (SMD) based on a random effects model were calculated.

Results: We identified five relevant studies (seven comparisons, n=684; one comparison each for haloperidol [n=243], mosapramine [n=238], olanzapine [n=39], quetiapine [n=42], perospirone [n=100], and two comparisons for risperidone [n=66]). There were no significant differences in the Positive and Negative Syndrome Scale (PANSS) total, negative, and general scores (SMD=0.10, SMD=0.09, SMD=0.10, respectively); discontinuation rate associated with all causes (OR=1.35); or side effects (OR=1.03) between aripiprazole and the pooled antipsychotics. Aripiprazole was inferior to the pooled antipsychotics in PANSS positive subscale scores (SMD=0.17) and discontinuation because of inefficacy (OR=2.21, NNH=11). However, aripiprazole had fewer side effects compared with the pooled antipsychotics (OR=0.21, NNH=20 for one or more side effects), including fatigue (OR=0.22, NNH=8), hyperprolactinemia (OR=0.00, NNH=1), extrapyramidal symptoms (OR=0.46, NNH=6), and weight gain (OR=0.36, NNH=7). Moreover, aripiprazole was associated with lower total cholesterol (SMD=−0.20) and triglyceride (SMD=−0.17) levels and body weight (SMD=−0.20) compared with the pooled antipsychotics.

Conclusion: Although the discontinuation rate associated with inefficacy was higher with aripiprazole than with the pooled antipsychotics, aripiprazole was associated with a lower risk of hyperprolactinemia and metabolic and extrapyramidal symptoms compared with the pooled antipsychotics.

Keywords: Japan, antipsychotics, efficacy, safety, meta-analysis, systematic review

Introduction

“Aripiprazole”, a third-generation antipsychotic, is a potent, high-affinity, partial dopamine D_2 receptor and serotonin (5-HT)1A receptor agonist and a potent 5-HT_1B receptor antagonist, and has a lower risk of metabolic side effects compared with second-generation antipsychotics (SGAs). A recent network meta-analysis showed that aripiprazole caused more weight gain than placebo, with a small effect size (standardized mean difference [SMD]=0.17). Another meta-analysis showed that aripiprazole treatment was associated with a lower risk of both weight gain (vs olanzapine, SMD=−0.42) and an increase in total cholesterol (vs olanzapine, SMD=−0.34;
vs risperidone, SMD=−0.50) and blood glucose (vs olanzapine, SMD=−0.16). Aripiprazole also carries a low risk of hyperprolactinemia and sedation. Therefore, it has widely been recommended as a first-line treatment for schizophrenia.6

Recently, inter-racial and ethnic differences have been noted in the response to and side effects of antipsychotics. A recent meta-analysis demonstrated that the effect size for antipsychotic response was smaller in patients with schizophrenia in North America than in those in Europe and the rest of the world.7 Several articles have reported that the contribution of genetic variations may lead to differences in antipsychotic response and the development of side effects.8–10 The main therapeutic target of antipsychotics are dopamine D2 receptors, and the main metabolizer of antipsychotics is cytochrome P450. The minor allele frequencies rs2283265 (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2283265) and rs1076560 (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=1076560), which are associated with a reduced short-form lacking exon6 in the dopamine D2 receptor gene11 and cytochrome P450 (http://hapmap.ncbi.nlm.nih.gov/index.html_en), of each population were different. For example, the Asian population is reported to exhibit a higher rate of clozapine-induced agranulocytosis than other populations.12 Reports have also highlighted differences in the prevalence of metabolic syndrome and its predisposition factors, such as diabetes, between races of individuals without psychiatric disorders.13,14 A recent clinical study reported that olanzapine caused greater weight gain in black patients with schizophrenia than in white patients with schizophrenia.15 The Asian population is reported to exhibit a higher rate of clozapine-induced agranulocytosis than other populations.12 Although there is a lower prevalence of obesity in the Japanese population, there is a higher rate of diabetes than in other races.16 Metabolic syndrome is considered a complex disease in which the pathophysiology of insulin resistance probably includes a gene–environment interaction.17,18 A recent genome-wide association study showed that rs489693, which is located approximately 190 kb downstream from the melanocortin 4 receptor gene, was associated with antipsychotic-induced weight gain.19 The minor allele frequency rs489693 also differs between races of individuals (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=489693).

Thus, when data from various races or ethnicities are combined, the biological and environmental factors may notably influence the evidence regarding the efficacy and safety. The Japanese population reportedly has a higher homogeneity of genetic and cultural backgrounds than other populations.20 Moreover, the universal health care insurance system in Japan allows its population to obtain health services without suffering financial hardship. Therefore, to clarify the pharmacological profile of aripiprazole in the treatment of schizophrenia, we conducted a systematic review and meta-analysis of randomized controlled trials (RCTs) comparing aripiprazole with other pooled antipsychotics in the Japanese population.

Methods

This meta-analysis was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Supplementary appendix 1).21

Inclusion criteria and search strategy, data extraction, and outcomes

Randomized active-controlled trials of aripiprazole for schizophrenia in Japanese-only cohorts were selected. Open-label and crossover studies were included to increase the sample size for the meta-analysis. Relevant studies were identified through searches of PubMed®, and the databases of the Cochrane Library, the Japan Medical Abstracts Society, and PsycINFO® citations up to January 5, 2014. We used the following keywords: “Japan” or “Japanese” and “aripiprazole” and “schizophrenia” in English and Japanese. In addition, we used the drug package insert for each antipsychotic. We also inquired with Otsuka Pharmaceutical Co Ltd, who manufacture and sell aripiprazole, regarding its clinical Phase III trials. Three authors (TK, YM, and SM) checked the inclusion and exclusion criteria of the studies identified. When data required for the current meta-analysis were missing, the first/corresponding authors were contacted for additional information. Unpublished data were provided for the current study by Dr Yoshiteru Takekita, Dr Kotaro Hatta, Dr Goro Sato, Dr Jun Ishigooka, and by Otsuka Pharmaceutical Co Ltd. Discrepancies in the two coding forms were resolved by discussions between two authors (TK and NI). Three authors (TK, YM, and SM) independently extracted, checked, and entered the data into the Review Manager software (v 5.2, Cochrane Collaboration, Oxford, England; http://tech.cochrane.org/Revman).

Data synthesis and statistical analysis

We included the outcomes of at least two studies for each outcome. The primary outcome for measuring efficacy was the psychopathology of schizophrenia; that is, the Positive and Negative Syndrome Scale (PANSS) total scores (change of scores from baseline to endpoint from all studies22–25);
PANSS positive (endpoint scores from one study\textsuperscript{26} and change of scores from baseline to endpoint from other studies\textsuperscript{22–25}), negative (endpoint scores from one study\textsuperscript{26} and change of scores from baseline to endpoint from other studies\textsuperscript{22–25}), and general psychopathology (endpoint scores from one study\textsuperscript{26} and change of scores from baseline to endpoint from other studies\textsuperscript{22–25}) subscale scores;\textsuperscript{27} and discontinuation rate associated with inefficacy. The secondary outcomes were discontinuation rate associated with all causes, discontinuation rate associated with side effects, discontinuation rate associated with death, and discontinuation rate associated with death by suicide. We also evaluated individual side effects.

We used the Drug-Induced Extrapyramidal Symptom Scale (DIEPSS)\textsuperscript{28,29} to examine the severity of extrapyramidal symptoms. Two\textsuperscript{22–24} of three studies included change scores against the worst scores during the study, while the remaining study\textsuperscript{25} used the last observational scores.

We based our analyses on intention-to-treat (ITT) or modified ITT data (ie, at least one dose or at least one follow-up assessment); no observed case data were allowed. However, we included data to increase the sample size included in the meta-analysis. The meta-analysis was performed using Review Manager.

We assessed the methodological qualities of the articles included in the meta-analysis on the basis of the Cochrane Risk of Bias Criteria.\textsuperscript{10} To combine studies, we used the conservative random effects model by DerSimonian and Laird\textsuperscript{11} because of the possibility that the underlying effect differed across studies and populations are usually heterogeneous. For continuous data, SMD was used by combining the effect size (Hedges’g) data. For dichotomous data, the odds ratio (OR) was estimated along with its 95% confidence interval (CI). In this study, when the random effects model showed significant differences between groups, the number-needed-to-harm (NNH) was calculated. Then, NNH values were derived from the risk differences (RDs) using the following formula: \( \text{NNH} = \frac{1}{\text{RD}} \), with the 95% CIs of NNH being the inverse of the upper and lower limits of the 95% CI of the RDs.

We explored study heterogeneity using the \( F \) statistic, considering values of \( \geq 50\% \) to reflect considerable heterogeneity.\textsuperscript{25} In cases of \( F \) values \( \geq 50\% \) for the primary outcomes, we planned to conduct sensitivity analyses to determine the reasons for the heterogeneity. However, we found no significant heterogeneity in the primary outcomes between aripiprazole and the pooled antipsychotics. On the other hand, because we merged the data of different clinical trial designs, we performed several subgroup analyses of primary outcomes. In addition, haloperidol is a high-potency first-generation antipsychotic (FGA) that is associated with a great safety risk.\textsuperscript{3,31} Because the haloperidol group consisted of a large proportion of the pooled cohort,\textsuperscript{24} which may cause bias, we also performed subgroup analyses excluding the haloperidol-controlled RCT to assess safety outcomes (ie, aripiprazole vs other pooled SGAs). Funnel plots were inspected visually to assess the possibility of publication bias.

**Results**

**Study characteristics**

The computerized search initially yielded 105 results after duplicates were removed. We excluded 65 studies following a review of the titles and abstracts. A further article was excluded after full-text review because this involved aripiprazole add-on and/or switching. Two additional studies\textsuperscript{22,24} were identified from the Japanese version of the aripiprazole package insert. In total, we identified five studies (seven comparisons, total \( n = 684 \); one comparison each for haloperidol \( n = 243 \),\textsuperscript{24} mosapramine \( n = 238 \),\textsuperscript{22} olanzapine \( n = 39 \),\textsuperscript{23} quetiapine \( n = 42 \),\textsuperscript{23} and perospirone \( n = 100 \),\textsuperscript{25} and two for risperidone \( n = 66 \)) (Figure 1).

Two studies\textsuperscript{22,24} were double-blind RCTs that mentioned the required study design details, and, of the remaining studies; one was rater-blinded;\textsuperscript{23} one was open-label;\textsuperscript{23} and one was a double-blind, randomized, crossover study\textsuperscript{26} (Supplementary appendix 2). We based the analyses on ITT or modified ITT data. However, data from the crossover study were not excluded to ensure that as much information as possible was obtained. Three of the studies\textsuperscript{23,25,26} were published in English and two\textsuperscript{22,24} were published in Japanese. The mean study duration was 8.8 weeks (8–12 weeks), with sample sizes with a range of 17–122 participants per treatment arm. All participants were adults diagnosed with schizophrenia. Only one study\textsuperscript{25} included emergency cases. The characteristics of the studies are shown in Table 1.

**Primary efficacy outcomes**

When aripiprazole was compared with the pooled antipsychotics, there were no significant differences in PANSS total scores (SMD=0.10, 95% CI=–0.10 to 0.31, \( P = 0.32, F = 36 \), six comparisons, \( n = 691 \); Figure 2). Although aripiprazole was inferior to the pooled antipsychotics with regard to the PANSS positive subscale scores (SMD=0.17, 95% CI=0.03 to 0.32, \( P = 0.02, F = 0 \), seven comparisons, \( n = 727 \); Figure 3), there were no significant differences in either the PANSS negative (SMD=–0.09, 95% CI=–0.24 to 0.06, \( P = 0.23, F = 2 \),
seven comparisons, n=727; Figure 4) or PANSS general psychopathology subscale scores (SMD=0.10, 95% CI=−0.10 to 0.29, P=0.33, F=31, seven comparisons, n=727; Figure 5) between both the treatment groups. In addition, aripiprazole resulted in a higher rate of discontinuation associated with inefficacy than the pooled antipsychotics (OR=2.21, 95% CI=1.09–4.47, P=0.03, I^2=40, NNH=11, P=0.03, seven comparisons, n=727; Figure 6). Visual inspection of the funnel plots for primary outcomes in both treatment groups did not suggest the presence of publication bias (Supplementary appendix 3).

**Subgroup analyses of primary outcomes**

With regard to the PANSS total scores, aripiprazole was inferior to the pooled antipsychotics according to subgroup analyses using data from RCTs with small samples (n<100) and only emergency cases (Table 2). Aripiprazole was inferior to the pooled antipsychotics in terms of the PANSS positive scores in subgroup analyses that included data from RCTs that were 8 weeks long, had a small sample size (n<100), were double-blind/rater-blind, involved SGA comparisons, were nonindustry-sponsored RCTs, were no crossover RCTs, and included only emergency cases (Table 2).

With regard to PANSS negative subscale scores, aripiprazole was superior to haloperidol only in subgroup analyses using data comparing haloperidol.

With regard to PANSS general subscale scores, aripiprazole was inferior to the pooled antipsychotics according to subgroup analyses using data from small-sample RCTs (n<100) or SGA comparator RCTs (Table 2).

With regard to discontinuation associated with inefficacy, subgroup analyses using data from SGA comparator RCTs or nonindustry-sponsored RCTs revealed that aripiprazole was inferior to the pooled antipsychotics (Table 2).

**Secondary outcomes**

No significant differences were detected in discontinuation rates associated with all causes, side effects, death, or death by suicide between aripiprazole and the pooled antipsychotics (Supplementary appendix 4).

**Individual side effects**

Aripiprazole was associated with a lower incidence of the following in comparison with the pooled antipsychotics (Table 3; Supplementary appendix 4): at least one side effect (OR=0.21, 95% CI=0.06–0.76, P=0.02, F=0, NNH=20, P=0.006, two comparisons, n=478), powerlessness (OR=0.24, 95% CI=0.09–0.60,
Table 1 Study, patient, and treatment characteristics of included randomized controlled trials

| Study                  | N   | Patient characteristics                                                                 | Diagnosis | Duration | Age, years (mean ± SD [range]) | Male (%) | Drug n | Dose, mg/day (mean ± SD) (HAL equivalent, mg/day) | PANSS score results |
|------------------------|-----|----------------------------------------------------------------------------------------|-----------|----------|--------------------------------|----------|--------|--------------------------------------------------|---------------------|
| Ishigooka et al[^4]    | 243 | Excluding patients who were pregnant or who wanted to become pregnant, had obvious complications, were in a coma, had an allergy to HAL, and/or had a neurological disorder | ICD-10    | 8 weeks  | ARI: 43.6±13.2 [16–65]          | ARI: 64, HAL: 57 | ARI    | 122 12.2±6.4 (6.1) [6–24], flexible              | Total: ARI = HAL, Positive: ARI = HAL, Negative: ARI > HAL, General: ARI = HAL |
|                        |     |                                                                                        |           |          | ARI: 43.7±13.6 [16–65]          |          | HAL    | 121 5.8±3.2 (5.8) [3–12], flexible              |                     |
| 031-95-003[^2,^3]      | 238 | Excluding patients who were pregnant or who wanted to become pregnant, had obvious complications, were in a coma, had an allergy to MOS, and/or had a neurological disorder | ICD-10    | 8 weeks  | ARI: 45.5±12.4 [16–65]          | ARI: 63, MOS: 71 | ARI    | 120 13 0±7.3 (6.5) [6–24], flexible              | Total: ARI = MOS, Positive: ARI < MOS, Negative: ARI = MOS, General: ARI = MOS |
|                        |     |                                                                                        |           |          | MOS: 45.2±12.7 [16–65]          |          | MOS    | 118 92.4±52.7 (5.6) [45–180], flexible          |                     |
| Hatta et al[^23]       | 80  | Excluding patients who were pregnant or who wanted to become pregnant, had obvious complications, and/or had brain organic disease | ICD-10    | 8 weeks  | ARI: 42.1±12.4 [18–66]          | ARI: 38, OLA: 71, QUE: 20, RIS: 45 | ARI    | 22 23.6±6.5 (11.8) [3–12], flexible              | Total: ARI = OLA = QUE = RIS, Positive: ARI = OLA = QUE = RIS, Negative: ARI = OLA = QUE = RIS, General: ARI = OLA = QUE = RIS |
|                        |     |                                                                                        |           |          | OLA: 39.8±10.8 [3–12]           |          | OLA    | 17 17.4±4.7 (13.9) [3–12], flexible              |                     |
|                        |     |                                                                                        |           |          | QUE: 39.8±1.2 [1–12]            |          | QUE    | 20 57.9±10.2 (17.5) [3–12], flexible            |                     |
|                        |     |                                                                                        |           |          | RIS: 41.1±8.8 [18–64]           |          | RIS    | 21 7.2±3.1 (7.2) [3–12], flexible               |                     |
| Takekita et al[^25]    | 100 | Excluding patients who were pregnant or who wanted to become pregnant, had obvious complications, and/or had refractory schizophrenia | DSM-IV-TR | 12 weeks| ARI: 41 [18–71] (PER: 47 [17–73]) | ARI: 47, PER: 39 | ARI    | 49 15.5±6.7 (7.6) [6–30], flexible              | Total: ARI = PER, Positive: ARI = PER, Negative: ARI = PER, General: ARI = PER |
|                        |     |                                                                                        |           |          | AR: 47 [18–71] (PER: 47 [17–73]) |          | PER    | 51 22.7±12.6 (5.7) [12–48], flexible            |                     |
| Sato et al[^26]        | 23  | Excluding patients who were suicidal, had a neurological disorder, had an acute or unstable medical condition, had a clinically significant laboratory test value, and/or had a substance dependence | DSM-IV    | 16 weeks| 38.5±14.8 (8 weeks each)        | ARI: 52 | AR: 23 | 17.5±8.0 (8.8) [NR], flexible                   | Total: ARI = RIS, Positive: ARI = RIS, Negative: ARI = RIS, General: ARI = RIS |
|                        |     |                                                                                        |           |          | RIS: 23 2.61±1.7 (2.6) [NR], flexible |          | RIS    | 23 2.61±1.7 (2.6) [NR], flexible               |                     |

Abbreviations: ARI, aripiprazole; CO, crossover; DB, double-blind; DSM-IV-(TR), Diagnostic and Statistical Manual of Mental Disorders (Text Revision), fourth edition; HAL, haloperidol; ICD 10, International Statistical Classification of Diseases and Related Health Problems, tenth revision for mental and behavioral disorders; iP, inpatients; MOS, mosapramine; n, number of patients in subgroup; N, total number of patients; NR, not reported; OLA, olanzapine; PANSS, Positive and Negative Syndrome Scale; OP, outpatients; PER, perospirone; QUE, quetiapine; RB, rater-blind; RIS, risperidone; SD, standard deviation.
Fatigue (OR = 0.22, 95% CI = 0.11–0.45, P = 0.002, I² = 0, NNH = 8, P = 0.0003, F = 0, three comparisons, n = 481), body weight (SMD = −0.20, 95% CI = −0.35 to −0.05, P = 0.009, F = 0, seven comparisons, n = 704), total cholesterol levels (SMD = −0.20, 95% CI = −0.34 to −0.05, P = 0.009, F = 0, seven comparisons, n = 710), and triglyceride levels (SMD = −0.17, 95% CI = −0.32 to −0.02, P = 0.03, F = 0, seven comparisons, n = 696) in comparison with the pooled antipsychotics (Table 3; Supplementary appendix 4).

| Study or subgroup | Weight | Standardized mean difference IV, random (95% CI) | Standardized mean difference IV, random (95% CI) |
|-------------------|--------|---------------------------------|---------------------------------|
| 1.6.1 ARI vs HAL | Ishigooka et al²⁴ ARI vs HAL | 28.2% | −0.18 (–0.44, 0.07) | 28.2% | −0.18 (–0.44, 0.07) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=1.41 (P=0.16)** |
| 1.6.2 ARI vs MOS | 031-95-003²⁷ ARI vs MOS | 27.9% | 0.12 (–0.14, 0.37) | 27.9% | 0.12 (–0.14, 0.37) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=0.88 (P=0.38)** |
| 1.6.3 ARI vs OLA | Hatta et al²⁵ ARI vs OLA | 8.2% | 0.62 (–0.04, 1.27) | 8.2% | 0.62 (–0.04, 1.27) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=1.84 (P=0.07)** |
| 1.6.4 ARI vs PER | Takekita et al²⁶ ARI vs PER | 17.6% | 0.10 (–0.29, 0.49) | 17.6% | 0.10 (–0.29, 0.49) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=0.50 (P=0.62)** |
| 1.6.5 ARI vs QUE | Hatta et al²⁵ ARI vs QUE | 9.0% | 0.38 (–0.24, 1.00) | 9.0% | 0.38 (–0.24, 1.00) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=1.20 (P=0.23)** |
| 1.6.6 ARI vs RIS | Hatta et al²⁵ ARI vs RIS | 9.1% | 0.23 (–0.39, 0.84) | 9.1% | 0.23 (–0.39, 0.84) |
| **Subtotal (95% CI)** | **Heterogeneity: not applicable** | **Test for overall effect: Z=0.73 (P=0.46)** |

| **Total (95% CI)** | 100.0% | 0.10 (–0.10, 0.31) |
| **Heterogeneity: tau-squared=0.02; chi-square=7.79, df=5 (P=0.17); I²=36%** | **Test for overall effect: Z=0.99 (P=0.32)** |
| **Test for subgroup differences: chi-square=7.79, df=5 (P=0.17); I²=35.8%** | **Favors aripiprazole** | **Favors others** |

**Figure 2 Positive and Negative Syndrome Scale total scores.**

**Abbreviations:** ARI, aripiprazole; CI, confidence interval; df, degrees of freedom; HAL, haloperidol; MOS, mosapramine; OLA, olanzapine; PER, perospirone; QUE, quetiapine; RIS, risperidone.
Moreover, aripiprazole resulted in a greater incidence of weight loss than the pooled antipsychotics (OR = 2.82, 95% CI = 1.58–5.04, \( P = 0.0005 \), NNH = 9, \( P = 0.0003 \), two comparisons, n = 478; Table 3; Supplementary appendix 4).

The incidence of at least one extrapyramidal symptom (OR = 0.46, 95% CI = 0.27–0.79, \( P = 0.004 \), \( F = 48 \), NNH = 6, \( P = 0.008 \), five comparisons, n = 598) and the use of anticholinergic drugs (OR = 0.52, 95% CI = 0.30–0.91, \( P = 0.02 \), \( F = 47 \), NNH = not significant, seven comparisons, n = 744) were lower with aripiprazole than with the pooled antipsychotics (Table 3; Supplementary appendix 4).

Furthermore, aripiprazole was associated with lower DIEPSS overall severity (SMD = –0.40, 95% CI = –0.76 to –0.04, \( P = 0.03 \), \( F = 78 \), three comparisons, n = 573), DIEPSS dyskinesia (SMD = –0.25, 95% CI = –0.43 to –0.06, \( P = 0.008 \), \( F = 0 \), three comparisons, n = 573), DIEPSS dystonia (SMD = –0.24, 95% CI = –0.42 to –0.06, \( P = 0.009 \), \( F = 0 \), three comparisons, n = 573), DIEPSS bradykinesia

| Study or subgroup | Weight | Standardized mean difference IV, random (95% CI) | Standardized mean difference IV, random (95% CI) |
|------------------|--------|-----------------------------------------------|-----------------------------------------------|
| 1.7.1 ARI vs HAL |        |                                               |                                               |
| Ishigooka et al\(^{22}\) ARI vs HAL | 33.2%  | 0.00 (–0.25, 0.25)                             |                                               |
| Subtotal (95% CI) | 33.2%  | 0.00 (–0.25, 0.25)                             |                                               |
| Heterogeneity: not applicable | |                                               |                                               |
| Test for overall effect: \( Z = 2.3 (P = 0.02) \) | |                                               |                                               |

| 1.7.2 ARI vs MOS |        |                                               |                                               |
| 031-95-003\(^{22}\) ARI vs MOS | 31.9%  | 0.25 (–0.01, 0.51)                             |                                               |
| Subtotal (95% CI) | 31.9%  | 0.25 (–0.01, 0.51)                             |                                               |
| Heterogeneity: not applicable | |                                               |                                               |
| Test for overall effect: \( Z = 1.92 (P = 0.06) \) | |                                               |                                               |

| 1.7.3 ARI vs OLA |        |                                               |                                               |
| Hatta et al\(^{23}\) ARI vs OLA | 4.9%   | 0.65 (–0.01, 1.31)                             |                                               |
| Subtotal (95% CI) | 4.9%   | 0.65 (–0.01, 1.31)                             |                                               |
| Heterogeneity: not applicable | |                                               |                                               |
| Test for overall effect: \( Z = 1.94 (P = 0.05) \) | |                                               |                                               |

| 1.7.4 ARI vs PER |        |                                               |                                               |
| Takekita et al\(^{24}\) ARI vs PER | 13.8%  | 0.12 (–0.27, 0.51)                             |                                               |
| Subtotal (95% CI) | 13.8%  | 0.12 (–0.27, 0.51)                             |                                               |
| Heterogeneity: not applicable | |                                               |                                               |
| Test for overall effect: \( Z = 2.6 (P = 0.008) \) | |                                               |                                               |

| 1.7.5 ARI vs QUE |        |                                               |                                               |
| Hatta et al\(^{23}\) ARI vs QUE | 5.6%   | 0.32 (–0.30, 0.94)                             |                                               |
| Subtotal (95% CI) | 5.6%   | 0.32 (–0.30, 0.94)                             |                                               |
| Heterogeneity: not applicable | |                                               |                                               |
| Test for overall effect: \( Z = 1.02 (P = 0.31) \) | |                                               |                                               |

| 1.7.6 ARI vs RIS |        |                                               |                                               |
| Hatta et al\(^{23}\) ARI vs RIS | 5.5%   | 0.42 (–0.20, 1.04)                             |                                               |
| Sato et al\(^{25}\) ARI vs RIS | 5.0%   | 0.06 (–0.59, 0.71)                             |                                               |
| Subtotal (95% CI) | 10.5%  | 0.25 (–0.20, 0.70)                             |                                               |
| Heterogeneity: tau-squared=0.00; chi-square=0.62, \( df = 1 \) (\( P = 0.43 \)); \( P = 0 \% \) | |                                               |                                               |
| Test for overall effect: \( Z = 1.09 (P = 0.28) \) | |                                               |                                               |

| Total (95% CI) |        |                                               |                                               |
|               | 100.0% | 0.17 (0.03, 0.32)                             |                                               |
| Heterogeneity: tau-squared=0.00; chi-square=5.19, \( df = 6 \) (\( P = 0.52 \)); \( P = 0 \% \) | |                                               |                                               |
| Test for overall effect: \( Z = 2.33 (P = 0.02) \) | |                                               |                                               |
| Test for subgroup differences: chi-square=4.58, \( df = 5 \) (\( P = 0.47 \)); \( I^2 = 0 \% \) | |                                               |                                               |

Figure 3 Positive and Negative Syndrome Scale subscale positive scores.

Abbreviations: ARI, aripiprazole; CI, confidence interval; \( df \), degrees of freedom; HAL, haloperidol; MOS, mosapramine; OLA, olanzapine; PER, perospirone; QUE, quetiapine; RIS, risperidone.
(SMD=−0.39, 95% CI=−0.62 to −0.17, \( P=0.0006, F=42 \), three comparisons, \( n=573 \)), DIEPSS rigidity (SMD=−0.29, 95% CI=−0.46 to −0.13, \( P=0.0005, F=0 \), three comparisons, \( n=573 \)), DIEPSS gait (SMD=−0.34, 95% CI=−0.56 to −0.12, \( P=0.002, F=40 \), three comparisons, \( n=573 \)), and DIEPSS salivation subscale scores (SMD=−0.38, 95% CI=−0.55 to −0.22, \( P<0.0001, F=0 \), three comparisons, \( n=573 \)) than the pooled antipsychotics (Table 3; Supplementary appendix 4).

Although there were no differences in akathisia, dystonia, dry mouth, and increased salivation between aripiprazole and the pooled antipsychotics, aripiprazole was associated with less dyskinesia (OR=0.21, 95% CI=0.08–0.56, \( P=0.002, F=0 \), \( NNH=25, P<0.001 \), six comparisons, \( n=698 \)), bradykinesia (OR=0.41, 95% CI=0.26–0.67, \( P=0.0003, F=11, NNH=8, P<0.0001 \), six comparisons, \( n=698 \)), tremor (OR=0.58, 95% CI=0.40–0.85, \( P=0.005, F=0 \), \( NNH=14, P=0.01 \), six comparisons, \( n=727 \)), rigidity (OR=0.47, 95% CI=0.27–0.83, \( P=0.008, F=14, NNH=20, P=0.01 \), six comparisons, \( n=727 \)).
When excluding haloperidol-controlled RCTs (ie, aripiprazole versus other pooled SGAs), aripiprazole was superior to the other pooled SGAs in terms of dyskinesia (OR=0.23, 95% CI=0.08–0.65, P=0.006, I²=0, NNH=17, P=0.005, five comparisons, n=458), dystonia (OR=0.21, 95% CI=0.06–0.75, P=0.02, F=0, NNH=20, P=0.03, four comparisons, n=358), bradykinesia (OR=0.49, 95% CI=0.25–0.93, P=0.03, F=20, NNH=10, P=0.005, five comparisons, n=458), tremor (OR=0.57, 95% CI=0.35–0.94, P=0.03,

| Study or subgroup | Weight | Standardized mean difference IV, random (95% CI) | Standardized mean difference IV, random (95% CI) |
|-------------------|--------|-----------------------------------------------|-----------------------------------------------|
| 1.9.1 ARI vs HAL |        |                                               |                                               |
| Ishigooka et al23 ARI vs HAL | 26.4% | -0.22 (−0.47, 0.04)                           | -0.22 (−0.47, 0.04)                           |
| Subtotal (95% CI) | 26.4% |                                               |                                               |
| Heterogeneity: not applicable |     |                                               |                                               |
| Test for overall effect: Z=1.67 (P=0.09) |     |                                               |                                               |
| 1.9.2 ARI vs MOS |        |                                               |                                               |
| 031-95-0032 ARI vs MOS | 26.1% | 0.12 (−0.14, 0.37)                            | 0.12 (−0.14, 0.37)                            |
| Subtotal (95% CI) | 26.1% |                                               |                                               |
| Heterogeneity: not applicable |     |                                               |                                               |
| Test for overall effect: Z=0.88 (P=0.38) |     |                                               |                                               |
| 1.9.3 ARI vs OLA |        |                                               |                                               |
| Hatta et al23 ARI vs OLA | 7.5%  | 0.54 (−0.11, 1.19)                            | 0.54 (−0.11, 1.19)                            |
| Subtotal (95% CI) | 7.5%  |                                               |                                               |
| Heterogeneity: not applicable |     |                                               |                                               |
| Test for overall effect: Z=1.62 (P=0.11) |     |                                               |                                               |
| 1.9.4 ARI vs PER |        |                                               |                                               |
| Takekita et al24 ARI vs PER | 16.2% | 0.07 (−0.32, 0.46)                            | 0.07 (−0.32, 0.46)                            |
| Subtotal (95% CI) | 16.2% |                                               |                                               |
| Heterogeneity: not applicable |     |                                               |                                               |
| Test for overall effect: Z=0.36 (P=0.72) |     |                                               |                                               |
| 1.9.5 ARI vs QUE |        |                                               |                                               |
| Hatta et al23 ARI vs QUE | 8.2%  | 0.37 (−0.25, 0.99)                            | 0.37 (−0.25, 0.99)                            |
| Subtotal (95% CI) | 8.2%  |                                               |                                               |
| Heterogeneity: not applicable |     |                                               |                                               |
| Test for overall effect: Z=1.18 (P=0.24) |     |                                               |                                               |
| 1.9.6 ARI vs RIS |        |                                               |                                               |
| Hatta et al23 ARI vs RIS | 8.3%  | 0.18 (−0.44, 0.79)                            | 0.18 (−0.44, 0.79)                            |
| Sato et al25 ARI vs RIS | 7.3%  | 0.37 (−0.29, 1.03)                            | 0.37 (−0.29, 1.03)                            |
| Subtotal (95% CI) | 15.6% |                                               |                                               |
| Heterogeneity: tau-squared=0.00; chi-square=0.18, df=1 (P=0.67); I²=0% |     |                                               |                                               |
| Test for overall effect: Z=1.16 (P=0.24) |     |                                               |                                               |
| Total (95% CI) | 100.0% | 0.10 (−0.10, 0.29)                           | 0.10 (−0.10, 0.29)                           |
| Heterogeneity: tau-squared=0.02; chi-square=8.74, df=6 (P=0.19); I²=31% |     |                                               |                                               |
| Test for overall effect: Z=0.98 (P=0.33) |     |                                               |                                               |
| Test for subgroup differences: chi-square=8.57, df=5 (P=0.13); I²=41.6% |     |                                               |                                               |

Figure 5 Positive and Negative Syndrome Scale subscale general scores.
Abbreviations: ARI, aripiprazole; CI, confidence interval; df, degrees of freedom; HAL, haloperidol; MOS, mosapramine; OLA, olanzapine; PER, perospirone; QUE, quetiapine; RIS, risperidone.

CI=0.30–0.74, P=0.001, F=0, NNH= not significant, six comparisons, n=698), and gait disturbance (OR=0.37, 95% CI=0.22–0.62, P=0.0002, F=0, NNH=11, P=0.0006, six comparisons, n=698) than the pooled antipsychotics (Table 3; Supplementary appendix 4).

However, there were no significant differences in the outcomes related to individual side effects between aripiprazole and the pooled antipsychotics (Supplementary appendix 4).
Figure 6 Discontinuation associated with inefficacy.

Abbreviations: ARI, aripiprazole; CI, confidence interval; df, degrees of freedom; HAL, haloperidol; M–H, Mantel-Haenszel method; MOS, mosapramine; OLA, olanzapine; PER, perospirone; QUE, quetiapine; RIS, risperidone.

| Study or subgroup | Weight | Odds ratio M–H, random (95% CI) | Odds ratio M–H, random (95% CI) |
|-------------------|--------|---------------------------------|---------------------------------|
| 1.2.1 ARI vs HAL  |        |                                 |                                 |
| Ishigooka et alP | 17.7%  | 0.82 (0.24, 2.76)               | 0.82 (0.24, 2.76)               |
| Subtotal (95% CI) | 17.7%  |                                 |                                 |
| Total events     |        |                                 |                                 |
| Heterogeneity: not applicable | |
| Test for overall effect: Z=0.32 (P=0.75) | |
| 1.2.2 ARI vs MOS  |        |                                 |                                 |
| 031-95-00322 ARI vs MOS | 20.2%  | 2.75 (0.95, 7.96)               | 2.75 (0.95, 7.96)               |
| Subtotal (95% CI) | 20.2%  |                                 |                                 |
| Total events     |        |                                 |                                 |
| Heterogeneity: not applicable | |
| Test for overall effect: Z=1.86 (P=0.06) | |
| 1.2.3 ARI vs OLA  |        |                                 |                                 |
| Hatta et alP ARI vs OLA | 11.8%  | 6.25 (1.14, 34.12)              | 6.25 (1.14, 34.12)              |
| Subtotal (95% CI) | 11.8%  |                                 |                                 |
| Total events     |        |                                 |                                 |
| Heterogeneity: not applicable | |
| Test for overall effect: Z=2.12 (P=0.03) | |
| 1.2.4 ARI vs PER  |        |                                 |                                 |
| Takekita et alP ARI vs PER | 12.2%  | 3.42 (0.66, 17.83)              | 3.42 (0.66, 17.83)              |
| Subtotal (95% CI) | 12.2%  |                                 |                                 |
| Total events     |        |                                 |                                 |
| Heterogeneity: not applicable | |
| Test for overall effect: Z=1.46 (P=0.14) | |
| 1.2.5 ARI vs QUE  |        |                                 |                                 |
| Hatta et alP ARI vs QUE | 17.7%  | 0.68 (0.20, 2.30)               | 0.68 (0.20, 2.30)               |
| Subtotal (95% CI) | 17.7%  |                                 |                                 |
| Total events     |        |                                 |                                 |
| Heterogeneity: not applicable | |
| Test for overall effect: Z=0.62 (P=0.54) | |
| 1.2.6 ARI vs RIS  |        |                                 |                                 |
| Hatta et alP ARI vs RIS | 15.4%  | 3.54 (0.90, 14.01)              | 3.54 (0.90, 14.01)              |
| Sato et alP ARI vs RIS | 4.9%   | 13.97 (0.73, 269.23)            | 13.97 (0.73, 269.23)            |
| Subtotal (95% CI) | 20.4%  | 4.52 (1.30, 15.73)              | 4.52 (1.30, 15.73)              |
| Total events     |        |                                 |                                 |
| Heterogeneity: tau-squared=0.00; chi-square=0.71, df=1 (P=0.40); I²=0% | |
| Test for overall effect: Z=2.37 (P=0.02) | |
| Total (95% CI) | 100.0% | 2.21 (1.09, 4.47)               | 2.21 (1.09, 4.47)               |
| Total events     |        |                                 |                                 |
| Heterogeneity: tau-squared=0.35; chi-square=9.94, df=6 (P=0.13); I²=40% | |
| Test for overall effect: Z=2.80 (P=0.003) | |
| Test for subgroup differences: chi-square=9.19, df=5 (P=0.10); I²=45.6% | |
### Table 2: Sensitivity/subgroup analyses of primary efficacy outcomes

| Outcome                        | Variable                      | Subgroup            | Comparisons, N | Patients, N | SMD*       | 95% CI          | P       | I² |
|--------------------------------|-------------------------------|----------------------|----------------|-------------|------------|----------------|---------|----|
| **PANSS total scores**         | Trial duration                | 8 weeks             | 5              | 591         | 0.13       | -0.13 to 0.39  | 0.34    | 48 |
|                                |                               | 12 weeks            | 1              | 100         | 0.10       | -0.29 to 0.49  | 0.62    | na |
|                                | Participants, N               | <100                 | 3              | 120         | 0.40       | 0.04 to 0.76   | 0.03    | 0  |
|                                |                               | ≥100                 | 3              | 571         | -0.01      | -0.21 to 0.20  | 0.96    | 33 |
|                                | Blinding                      | Double-blind/rater-blind | 5             | 591         | 0.13       | -0.13 to 0.39  | 0.34    | 48 |
|                                |                               | Open                 | 1              | 100         | 0.10       | -0.29 to 0.49  | 0.62    | na |
|                                | Comparators                   | Second-generation antipsychotic | 5             | 452         | 0.19       | 0.00 to 0.37   | 0.05    | 0  |
|                                |                               | First-generation antipsychotic | 1             | 239         | -0.18      | -0.44 to 0.07  | 0.16    | na |
|                                | Sponsorship                   | Industry             | 2              | 471         | -0.03      | -0.33 to 0.26  | 0.82    | 62 |
|                                |                               | Nonindustry          | 4              | 220         | 0.26       | -0.01 to 0.53  | 0.06    | 0  |
|                                | Trial design                   | No crossover         | 6              | 691         | 0.10       | -0.10 to 0.31  | 0.32    | 36 |
|                                |                               | Crossover            | 0              | 571         | -0.01      | -0.21 to 0.20  | 0.96    | 33 |
|                                | Patients                       | Emergency cases      | 3              | 120         | 0.40       | 0.04 to 0.76   | 0.03    | 0  |
|                                |                               | Others               | 3              | 571         | -0.01      | -0.21 to 0.20  | 0.96    | 33 |
| **PANSS positive subscale scores** | Trial duration                | 8 weeks             | 6              | 627         | 0.18       | 0.02 to 0.35   | 0.02    | 2  |
|                                |                               | 12 weeks            | 1              | 100         | 0.12       | -0.27 to 0.51  | 0.55    | na |
|                                | Participants, N               | <100                 | 4              | 156         | 0.36       | 0.04 to 0.68   | 0.03    | 0  |
|                                |                               | ≥100                 | 3              | 571         | 0.12       | -0.04 to 0.29  | 0.14    | 0  |
|                                | Blinding                      | Double-blind/rater-blind | 6             | 627         | 0.18       | 0.02 to 0.35   | 0.02    | 2  |
|                                |                               | Open                 | 1              | 100         | 0.12       | -0.27 to 0.51  | 0.55    | na |
|                                | Comparators                   | Second-generation antipsychotic | 6             | 488         | 0.26       | 0.08 to 0.44   | 0.004   | 0  |
|                                |                               | First-generation antipsychotic | 1             | 239         | 0.00       | -0.25 to 0.25  | 1.00    | na |
|                                | Sponsorship                   | Industry             | 2              | 471         | 0.13       | -0.12 to 0.37  | 0.32    | 47 |
|                                |                               | Nonindustry          | 5              | 256         | 0.27       | 0.02 to 0.51   | 0.03    | 0  |
|                                | Trial design                   | No crossover         | 6              | 691         | 0.18       | 0.03 to 0.33   | 0.02    | 1  |
|                                |                               | Crossover            | 1              | 56         | 0.06       | -0.59 to 0.71  | 0.86    | na |
|                                | Patients                       | Emergency cases      | 3              | 120         | 0.46       | 0.09 to 0.82   | 0.01    | 0  |
|                                |                               | Others               | 4              | 607         | 0.12       | -0.04 to 0.28  | 0.14    | 0  |
| **PANSS negative subscale scores** | Trial duration                | 8 weeks             | 6              | 627         | -0.11      | -0.28 to 0.06  | 0.20    | 7  |
|                                |                               | 12 weeks            | 1              | 100         | 0.07       | -0.33 to 0.46  | 0.74    | na |
|                                | Participants, N               | <100                 | 4              | 156         | 0.06       | -0.26 to 0.38  | 0.71    | 1  |
|                                |                               | ≥100                 | 3              | 571         | -0.13      | -0.30 to 0.03  | 0.11    | 0  |
|                                | Blinding                      | Double-blind/rater-blind | 6             | 627         | -0.11      | -0.28 to 0.06  | 0.20    | 7  |
|                                |                               | Open                 | 1              | 100         | 0.07       | -0.33 to 0.46  | 0.74    | na |
|                                | Comparators                   | Second-generation antipsychotic | 7             | 488         | -0.01      | -0.19 to 0.16  | 0.88    | 0  |
|                                |                               | First-generation antipsychotic | 1             | 239         | -0.26      | -0.51 to -0.00 | 0.05    | na |
|                                | Sponsorship                   | Industry             | 2              | 471         | -0.18      | -0.36 to 0.00  | 0.05    | 0  |
|                                |                               | Nonindustry          | 5              | 256         | 0.06       | -0.18 to 0.31  | 0.62    | 0  |

(Continued)
Table 2 Continued

| Outcome Variable | Subgroup | Comparisons, N | Patients, N | SMD* | 95% CI       | P   | I² |
|------------------|----------|----------------|-------------|-------|--------------|-----|----|
| PANSS general subscale scores | Trial duration | 8 weeks | 6 | 651 | 2.12 | 0.95–4.73 | 0.07 | 48 | na | na |
|                  |          | 12 weeks | 1 | 100 | 3.42 | 0.66–17.8 | 0.14 | 57 | na | na |
|                  | Participants, N | <100 | 4 | 170 | 2.88 | 0.81–10.3 | 0.10 | 57 | na | na |
|                  |          | ≥100 | 3 | 581 | 1.88 | 0.79–4.47 | 0.16 | 28 | na | na |
|                  | Blinding | Double-blind/rater-blind | 6 | 651 | 2.12 | 0.95–4.73 | 0.07 | 48 | na | na |
|                  |          | Open | 1 | 100 | 3.42 | 0.66–17.8 | 0.14 | 28 | na | na |
|                  | Comparators | Second-generation antipsychotic | 6 | 508 | 2.69 | 1.29–5.60 | 0.008 | 31 | 8 | 0.006 |
|                  |          | First-generation antipsychotic | 1 | 243 | 0.82 | 0.24–2.76 | 0.75 | na | na | na |
|                  | Sponsorship | Industry | 2 | 481 | 1.56 | 0.48–5.08 | 0.46 | 54 | na | na |
|                  |          | Nonindustry | 5 | 270 | 2.84 | 1.06–7.61 | 0.04 | 45 | 6 | 0.01 |
|                  | Trial design | No crossover | 6 | 705 | 2.00 | 0.99–4.02 | 0.05 | 39 | na | na |
|                  |          | Crossover | 1 | 46 | 13.97 | 0.73–269.2 | 0.08 | na | na | na |
|                  | Patients | Emergency cases | 3 | 124 | 2.27 | 0.59–8.70 | 0.23 | 63 | na | na |
|                  |          | Others | 4 | 627 | 2.24 | 0.89–5.59 | 0.09 | 34 | na | na |

Notes: *Negative SMD values favor aripiprazole; positive SMD values favor other pooled antipsychotics. †OR < 1 favors aripiprazole; OR > 1 favors other pooled antipsychotics. Bold indicates significance.

Abbreviations: CI, 95% confidence interval; PANSS, Positive and Negative Syndrome Scale; na, not applicable; NNH, numbers-needed-to-harm; OR, odds ratio; SMD, standardized mean difference.
| Outcome                                           | Comparisons, N (comparators) | Patients, N | P    | OR or SMD     | 95% CI            | P  | NNH | P     |
|--------------------------------------------------|-------------------------------|-------------|------|---------------|-------------------|----|-----|-------|
| At least one side effect                         | 2 (HAL, MOS)                  | 478         | 0    | 0.21^†        | 0.06–0.76         | 0.02 | 20  | 0.006 |
| Powerlessness                                     | 2 (HAL, MOS)                  | 478         | 0    | 0.24^†        | 0.09–0.60         | 0.002 | 13  | 0.0005 |
| Fatigue                                          | 2 (HAL, MOS)                  | 478         | 0    | 0.22^†        | 0.11–0.45         | <0.00001 | 8  | <0.00001 |
| Hyperprolactinemia                               | 2 (HAL, MOS)                  | 388         | 0    | 0.00^†        | 0.00–0.01         | <0.00001 | 1  | 0.003 |
| Blood prolactin level                            | 3 (HAL, MOS, RIS)             | 424         | 90   | –1.01^*       | –1.73 to –0.30    | 0.005 | na  | na    |
| QTc interval                                      | 3 (HAL, MOS, PER)             | 481         | 0    | –0.34^*       | –0.53 to –0.16    | 0.0003 | na  | na    |
| Weight gain                                      | 5 (HAL, MOS, OLA, QUE, RIS)   | 602         | 46   | 0.36^†        | 0.16–0.81         | 0.01 | 7   | 0.03  |
| Weight loss                                       | 2 (HAL, MOS)                  | 478         | 0    | 2.82^†        | 1.58–5.04         | 0.0005 | 9  | 0.0003 |
| Body weight                                       | 7 (HAL, MOS, OLA, PER, QUE, RIS) | 704     | 0    | –0.20^*       | –0.35 to –0.05    | 0.009 | na  | na    |
| Total cholesterol level                           | 7 (HAL, MOS, OLA, PER, QUE, RIS) | 710     | 0    | –0.20^*       | –0.34 to –0.05    | 0.009 | na  | na    |
| Triglyceride level                                | 7 (HAL, MOS, OLA, PER, QUE, RIS) | 696     | 0    | –0.17^*       | –0.32 to –0.02    | 0.03 | na  | na    |
| At least one extrapyramidal symptom              | 5 (HAL, MOS, OLA, QUE, RIS)   | 598         | 48   | 0.46^†        | 0.27–0.79         | 0.004 | 6   | 0.008 |
| Parkinsonism                                      | 4 (HAL, OLA, QUE, RIS)        | 360         | 16   | 0.89^†        | 0.38–2.05         | 0.78 | na  | na    |
| Use of anticholinergic drugs                      | 7 (HAL, MOS, OLA, PER, QUE, RIS) | 744     | 47   | 0.52^†        | 0.30–0.91         | 0.02 | NS  |       |
| DIEPSS overall severity subscale scores           | 4 (HAL, MOS, PER, RIS)        | 608         | 79   | –0.31^*       | –0.69 to 0.08     | 0.12 | na  | na    |
| DIEPSS akathisia subscale scores                 | 3 (HAL, MOS, PER)             | 573         | 78   | –0.40^*       | –0.76 to –0.04    | 0.03 | na  | na    |
| Akathisia                                        | 6 (HAL, MOS, OLA, QUE, RIS)   | 698         | 34   | 0.77^†        | 0.47–1.28         | 0.32 | na  | na    |
| DIEPSS akathisia subscale scores                 | 3 (HAL, MOS, PER)             | 573         | 52   | –0.16^*       | –0.41 to 0.08     | 0.2  | na  | na    |
| Dyskinesia                                       | 6 (HAL, MOS, OLA, PER, QUE, RIS) | 698     | 0    | 0.21^†        | 0.08–0.56         | 0.002 | 25  | 0.001 |
| DIEPSS dystonia subscale scores                  | 3 (HAL, MOS, PER)             | 573         | 0    | –0.25^*       | –0.43 to –0.06    | 0.008 | na  | na    |
| Dystonia                                         | 5 (HAL, MOS, OLA, QUE, RIS)   | 598         | 0    | 0.49^†        | 0.22–1.05         | 0.07 | na  | na    |
| DIEPSS dystonia subscale scores                  | 3 (HAL, MOS, PER)             | 573         | 0    | –0.24^*       | –0.42 to –0.06    | 0.009 | na  | na    |
| Bradykinesia                                      | 6 (HAL, MOS, OLA, PER, QUE, RIS) | 698     | 11   | 0.41^†        | 0.26–0.67         | 0.0003 | 8  | <0.0001 |
| DIEPSS bradykinesia subscale scores              | 3 (HAL, MOS, PER)             | 573         | 42   | –0.39^*       | –0.62 to –0.17    | 0.0006 | na  | na    |
| Tremor                                           | 6 (HAL, MOS, OLA, PER, QUE, RIS) | 698     | 0    | 0.58^†        | 0.40–0.85         | 0.005 | 14  | 0.01  |
| DIEPSS tremor subscale scores                    | 3 (HAL, MOS, PER)             | 573         | 63   | –0.27^*       | –0.55 to 0.01     | 0.06 | na  | na    |
| Rigidity                                         | 6 (HAL, MOS, OLA, PER, QUE, RIS) | 698     | 0    | 0.47^†        | 0.30–0.74         | 0.001 | NS  |       |
| DIEPSS rigidity subscale scores                  | 3 (HAL, MOS, PER)             | 573         | 0    | –0.29^*       | –0.46 to –0.13    | 0.0005 | na  | na    |
| Gait disturbance                                 | 6 (HAL, MOS, OLA, QUE, RIS)   | 698         | 0    | 0.37^†        | 0.22–0.62         | 0.0002 | 11  | 0.0006 |
| DIEPSS gait subscale scores                      | 3 (HAL, MOS, PER)             | 573         | 40   | –0.34^*       | –0.56 to –0.12    | 0.002 | na  | na    |
| Increased salivation                             | 5 (HAL, MOS, OLA, QUE, RIS)   | 598         | 59   | 0.61^†        | 0.24–1.53         | 0.29 | na  | na    |
| Dry mouth                                        | 2 (HAL, MOS)                  | 478         | 0    | 0.75^†        | 0.35–1.58         | 0.45 | na  | na    |
| DIEPSS salivation scores                         | 3 (HAL, MOS, PER)             | 573         | 0    | –0.38^*       | –0.55 to –0.22    | <0.00001 | na  | na    |

Notes: *Negative SMD values favor aripiprazole; positive SMD values favor other pooled antipsychotics. †OR < 1 favors aripiprazole; OR > 1 favors other pooled antipsychotics. Bold indicates significance.

Abbreviations: CI, confidence interval; DIEPSS, Drug-Induced Extrapyramidal Symptoms Scale; HAL, haloperidol; MOS, mosapramine; N, number; na, not applicable; NNH, numbers-needed-to-harm; NS, not significant; OLA, olanzapine; OR, odds ratio; PER, perospirone; QTc, corrected QT interval; QUL, quetiapine; RIS, risperidone; SMD, standardized mean difference.
Discussion

As far as we are aware, this is the first comprehensive meta-analysis to have been undertaken of the efficacy, effectiveness, and side effects of aripiprazole for the treatment of Japanese patients with schizophrenia, who are considered a homogeneous genetic and cultural population. Aripiprazole resulted in a higher rate of discontinuation associated with inefficacy compared with the pooled antipsychotics and seemed to be inferior to the pooled antipsychotics in the treatment of positive symptoms. In particular, when we conducted subgroup analyses using only SGA comparator studies, aripiprazole seemed to be inferior to the pooled antipsychotics with regard to all the primary efficacy outcomes, with the exception of PANSS negative subscale scores.

A recent meta-analysis showed that aripiprazole resulted in a higher rate of discontinuation associated with inefficacy compared with olanzapine.\(^34\) It suggests that the discontinuation rate is influenced by study design (blinded or open-label). After excluding an open-label study from our meta-analysis, aripiprazole was marginally inferior to the pooled antipsychotics (OR=2.12, \(P=0.07\)) in terms of the discontinuation rate associated with inefficacy. The mean dose of aripiprazole in the remaining studies was <18 mg/day (12.2–17.5 mg/day) compared with that in Hatta et al’s study, which averaged 23.6 mg/day.\(^23\) According to another recent meta-analysis, aripiprazole also has few sedative effects;\(^3\) therefore, patients who discontinued their treatment because of inefficacy may have done so because aripiprazole was not escalated to the optimum dose for those patients. Park et al determined that, when optimizing the dose of aripiprazole for schizophrenia, at least 3 weeks are needed to predict response or nonresponse at 6 weeks.\(^35\) Given this evidence, clinicians need to increase the aripiprazole dose to an adequate therapeutic range for at least 3 weeks.

Aripiprazole was generally well tolerated in comparison with the pooled SGAs in previous meta-analyses, with lower rates of overall side effects, cardiometabolic risk, and hyperprolactinemia.\(^34\) However, we found that aripiprazole was not associated with lower blood glucose levels, which conflicted with the results of previous meta-analysis\(^4\) and a clinical trial.\(^36\) Moreover, a meta-analysis of studies using only olanzapine and quetiapine as comparators revealed that aripiprazole was associated with higher blood glucose levels than either comparator drug (SMD=0.63, \(P=0.007\)). Although we could not identify a robust explanation for the discrepancy, it is possible that 1) statistical error was caused by the number of included studies, 2) the inclusion of all emergency cases in this meta-analysis led to bias,\(^23\) and 3) there was a possible alteration in glucose metabolism due to psychomotor excitement.

Recently, studies reported that aripiprazole might be associated with an increased risk of neuroleptic malignant syndrome (NMS).\(^37,38\) Despite performing a meta-analysis of NMS, we found no difference in this outcome between aripiprazole and the pooled antipsychotics. However, two patients receiving aripiprazole developed NMS during one study,\(^22\) suggesting that clinicians should be vigilant for the development of NMS during aripiprazole therapy.

Hatta et al reported that while the striatal dopamine activity in patients with acute state increased, that in patients during the chronic state decreased.\(^23\) This difference may reflect a therapeutic response to a dopamine stabilizer such as aripiprazole or to dopamine D\(_2\) receptor antagonists such as haloperidol. The results of our meta-analysis revealed that aripiprazole was inferior to other antipsychotics in treating PANSS positive subscale scores without significant heterogeneities. Moreover, when including only emergency cases, aripiprazole was inferior to other antipsychotics with regard to the PANSS total scores and the positive subscale scores and was marginally inferior to the other antipsychotics with regard to the PANSS general subscale scores.
Limitations
There are several limitations to the present analysis. First, there were few studies, which resulted in a small sample size. Second, our meta-analysis had several differences in patient populations (i.e., chronic/acute patients, inpatients/outpatients), which could have generated heterogeneity despite sensitivity analyses. Third, all studies included in this meta-analysis had short trial durations of 8–12 weeks. Further research is therefore required to elucidate the long-term efficacy and tolerability of aripiprazole treatment in Japanese patients with schizophrenia. Fourth, we compared aripiprazole with the pooled antipsychotics to reveal the pharmacological characteristics of aripiprazole against all antipsychotics. Therefore, differences in comparator drug profiles may have influenced the results of our meta-analysis. A recent network meta-analysis showed that although there were no notable differences in the effect size for efficacy between individual antipsychotics (haloperidol, olanzapine, quetiapine, and risperidone) and placebo, there were notable differences in the incidence of individual side effects. In our meta-analysis, we did not find significant heterogeneity in the primary outcomes, and safety outcomes were uncommon where there was significant heterogeneity despite comparing aripiprazole with pooled antipsychotics. Fifth, we did not compare the results of our meta-analysis with those of previous meta-analyses. To our knowledge, there is only one meta-analysis of SGA versus pooled FGAs in schizophrenia. However, only one FGA-controlled study was included in our meta-analysis; therefore, we did not compare those results focusing on the differences in ethnicities. To do this, a network meta-analysis of the efficacy and tolerability of antipsychotics in Japanese-only patients with schizophrenia will be required. Sixth, we did not control for multiple testing because this is not usually done in meta-analyses and the number of studies included was small. Finally, although two studies were clinical Phase III trials of aripiprazole in Japanese patients with schizophrenia, approximately 8% of participants received additional antipsychotic treatment and comparators (haloperidol or mosapramine). We could not exclude the effects of additional antipsychotic therapy on the outcomes.

Conclusion
Our results suggest that, although aripiprazole results in a higher rate of discontinuation associated with inefficacy compared with the pooled antipsychotics, it carries a lower risk of hyperprolactinemia and metabolic and extrapyramidal symptoms.

Acknowledgments
We thank Dr Yoshiiteru Takekita, Dr Kotaro Hatta, Dr Goro Sato, Dr Jun Ishigooka, and the Otsuka Pharmaceutical Co Ltd for providing information necessary for the study.

Disclosure
No funding sources were received for this study. The authors declare the following competing interests: Taro Kishi has received speaker’s honoraria from Abbott, Astellas, Daiichi Sankyo, Dainippon Sumitomo, Eisai, Eli Lilly, GlaxoSmithKline, Janssen, Yoshitomi, Otsuka, Meiji, Shionogi, Tanabe-Mitsubishi, Tsumura, Novartis, and Pfizer; Yuki Matsuda has received speaker’s honoraria from Dainippon Sumitomo Eli Lilly and Otsuka; Shinji Matsunaga has received speaker’s honoraria from Eisai, Janssen, Novartis, Daiichi Sankyo, Ono, Eli Lilly, Takeda, and Otsuka; and Nakao Iwata has received speaker’s honoraria from Astellas, Dainippon Sumitomo, Eli Lilly, GlaxoSmithKline, Janssen, Yoshitomi, Otsuka, Meiji, Shionogi, Novartis, and Pfizer.

References
1. Kishi T, Matsuda Y, Nakamura H, Iwata N. Blonanserin for schizophrenia: systematic review and meta-analysis of double-blind, randomized, controlled trials. J Psychiatr Res. 2013;47(2):149–154.
2. Kishi T, Iwata N. Efficacy and tolerability of perospirone in schizophrenia: a systematic review and meta-analysis of randomized controlled trials. CNS Drugs. 2013;27(9):731–741.
3. Leucht S, Cipriani A, Spinelli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet. 2013;382(9896):951–962.
4. Rummel-Kluge C, Komossa K, Schwarz S, et al. Head-to-head comparisons of metabolic side effects of second generation antipsychotics in the treatment of schizophrenia: a systematic review and meta-analysis. Schizophr Res. 2010;123(2–3):225–233.
5. Buchanan RW, Kreyenbuhl J, Kelly DL, et al; Schizophrenia Patient Outcomes Research Team (PORT). The 2009 schizophrenia PORT psychopharmacological treatment recommendations and summary statements. Schizophr Bull. 2010;36(1):71–93.
6. Kuipers E, Yesufu-Udechuku A, Taylor C, Kendall T. Management of psychosis and schizophrenia in adults: summary of updated NICE guidance. BMJ. 2013;448:1173.
7. Mattila T, Wohlfarth T, Koeter M, et al. Geographic variation in efficacy of atypical antipsychotics for the acute treatment of schizophrenia— an individual patient data meta-analysis. Eur Neuropsychopharmacol. 2014;24(7):1067–1077.
8. Zandi PP, Judy JT. The promise and reality of pharmacogenetics in psychiatry. Clin Lab Med. 2010;30(4):931–974.
9. Reynolds GP. The pharmacogenetics of symptom response to antipsychotic drugs. Psychiatr Invest. 2012;9(1):1–7.
10. Mulligan MJ, Vasu R, Grossi CE, et al. Neoplastic meningitis with eosinophilic pleocytosis in Hodgkin’s disease: a case with cerebellar dysfunction and a review of the literature. Am J Med Sci. 1988;296(5):322–326.
11. Zhang Y, Bertolino A, Fazio L, et al. Polymorphisms in human dopamine D2 receptor gene affect gene expression, splicing, and neuronal activity during working memory. Proc Natl Acad Sci U S A. 2007;104(51):20552–20557.
12. Munro J, O’Sullivan D, Andrews C, Arana A, Mortimer A, Kerwin R. Active monitoring of 12,760 clozapine recipients in the UK and Ireland. Beyond pharmacovigilance. Br J Psychiatry. 1999;175:576–580.
13. Danaei G, Finucane MM, Lu Y, et al; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet*. 2011;378(9783):31-40.

14. Yoon KH, Lee JH, Kim JW, et al. Epidemic obesity and type 2 diabetes in Asia. *Lancet*. 2006;368(9548):1681–1688.

15. Stauffer VL, Sniaidecky IL, Piezer KW, et al. Impact of race on efficacy and safety during treatment with olanzapine, schizophrenia, schizophreniform or schizoaffective disorder. *BMC Psychiatry*. 2010;10:89.

16. Sone H, Ito H, Ohashi Y, Akanuma Y, Yamada N; Japan Diabetes Complication Study Group. Obesity and type 2 diabetes in Japanese patients. *Lancet*. 2003;361(9351):85.

17. Slomko H, Heo HJ, Einstein FH. Minireview: Epigenetics of obesity and diabetes in humans. *Endocrinology*. 2012;153(3):1025–1030.

18. Choi JH, Yoo HW. Control of puberty: genetics, endocrinology, and environment. *Curr Opin Endocrinol Diabetes Obes*. 2013;20(1):62–68.

19. Malhotra AK, Correll CU, Chowdhury NI, et al. Association between common variants near the melanocortin 4 receptor gene and severe antipsychotic drug-induced weight gain. *Arch Gen Psychiatry*. 2012;69(9):904–912.

20. Yamaguchi-Kabata Y, Nakazone K, Takahashi A, et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet*. 2008;83(4):445–456.

21. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.

22. 031-95-003 (study name). Clinical evaluation of aripiprazole in schizophrenic patients: A comparative double-blind study with mosapramine. Available from: http://www.info.pmda.go.jp/shinyaku/P200600001/18007800_218000AMZ10011_G100_1.pdf?search=‘03195003’. Accessed November 12, 2014.

23. Hatta K, Sato K, Hamakawa H, et al. Effectiveness of second-generation antipsychotic drugs for schizophrenia. *Hiroshima J Med Sci*. 2012;61(4):75–83.

24. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261–276.

25. Takekita Y, Kato M, Wakeno M, et al. A 12-week randomized, open-label study of perospirone versus aripiprazole in the treatment of Japanese schizophrenia patients. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;40:110–114.

26. Sato G, Yoshimura S, Yamashita H, Okamoto Y, Yamawaki S. The neurocognitive effects of aripiprazole compared with risperidone in the treatment of schizophrenia. *Hiroshima J Med Sci*. 2012;61(4):75–83.

27. Khanna P, Komossa K, Rummel-Kluge C, et al. Aripiprazole versus olanzapine in the treatment of schizophrenia. *Schizophr Res*. 2002;57(2–3):227–238.

28. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

29. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188.

30. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.

31. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188.

32. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

33. Leucht S, Corves C, Arbiter D, Engel RR, Li C, Davis JM. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet*. 2009;373(9657):31–41.

34. Khanna P, Komossa K, Rummel-Kluge C, et al. Aripiprazole versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev.* 2013;2:CD006569.

35. The advantage of using 3-week data to predict response to aripiprazole at week 6 in first-episode psychosis. Park JI, Cho DH, Hahn SW, Jeong B, Kim JH, Kim SW, Koo MS, Lee SH, Lee SJ, Lee YH, Park JI, Rho SH, Chung YC. *Int Clin Psychopharmacol*. 2014;29(2):77-85.

36. Correll CU, Manu P, Olshanskiy V, Napolitano B, Kane JM, Malhotra AK. Cardiometabolic risk of second-generation antipsychotic medications during first-time use in children and adolescents. *JAMA*. 2009;302(16):1765–1773.

37. Su YP, Chang CK, Hayes RD, et al. Retrospective chart review on exposure to psychotropic medications associated with neuroleptic malignant syndrome. *Acta Psychiatr Scand*. 2014;130(1):52–60.

38. Kimura G, Kadoyama K, Brown JB, et al. Antipsychotics-Associated Serious Adverse Events in Children: An Analysis of the FAERS Database. *Int J Med Sci*. 2015; Jan 5;12(2):135–140.