

# Simultaneous Comparison of Efficacy and Tolerability of Second-Generation Antipsychotics in Schizophrenia: Mixed-Treatment Comparison Analysis Based on Head-to-Head Trial Data

Gyu Han Oh<sup>1</sup>, Je-Chun Yu<sup>1</sup>, Kyeong-Sook Choi<sup>1</sup>, Eun-Jeong Joo<sup>2</sup>, and Seong-Hoon Jeong<sup>1</sup> ✉

<sup>1</sup>Department of Psychiatry, Eulji University School of Medicine, Eulji University Hospital, Daejeon, Republic of Korea

<sup>2</sup>Department of Psychiatry, Eulji University School of Medicine, Eulji General Hospital, Seoul, Republic of Korea

**Objective** Second-generation antipsychotics have been repeatedly shown to be superior to placebo. However, the comparative efficacy among these drugs has not been systematically evaluated. In this study, we used Mixed Treatment Comparison (MTC) procedures to elucidate the comparative efficacy and tolerability of second-generation antipsychotics.

**Methods** Seven antipsychotics were selected based on the availability of the relevant data. Data were gathered from a series of review article published by the Cochrane Collaboration. Six outcome measures were analyzed: 1) percentage of no clinically important response as defined by the original authors, 2) PANSS total score change from baseline to endpoint, 3) percentage of akathisia, 4) percentage of antiparkinson medication use, 5) percentage of total body weight increase more than 7%, and 6) percentage of drop-out due to any reasons.

**Results** All the second-generation antipsychotics included in this study showed fairly similar efficacy but widely different tolerability. In terms of efficacy, amisulpride, clozapine and olanzapine were ranked higher than aripiprazole, quetiapine and ziprasidone. Clozapine and olanzapine were superior in terms of akathisia and extrapyramidal symptom risk, but, far more prone to induce clinically important weight gain.

**Conclusion** Using MTC methodology, we could line up the second generation antipsychotics according to their hierarchical superiority in terms of efficacy and tolerability. Though the wide overlap among the confidence intervals and the inconsistency between the direct and indirect comparison results may limit the validity of these results, it may still allow the important insights into the relative merits of the available drugs.

**Psychiatry Investig 2015;12(1):46-54**

**Key Words** Second-generation antipsychotics, Schizophrenia, Clinical trial, Meta-analysis, Mixed treatment comparison procedure.

## INTRODUCTION

Second-generation antipsychotics or atypical antipsychotics are generally recommended as first-line agents for the treatment and management of patients suffering from schizophrenia. However, the discrepancies among them have not been

adequately appreciated, though, considerable variations in efficacy and tolerability are to be expected.<sup>1,2</sup> This lack of evidence partly came from the scarcity of head-to-head clinical trials conducted on second-generation antipsychotics.<sup>3,4</sup>

The principles of evidence-based medicine requires physicians to search and appraise contemporary research findings and base their clinical decision on both valid and relevant evidences.<sup>5</sup> For this purpose, clinicians often rely upon systemic reviews or meta-analysis reports. However, traditional meta-analysis only permits summarizing the comparison data of the same two treatments, but not the others. Recently, the so-called “umbrella review” or “review of reviews” format has been recommended by the Cochrane Collaboration.<sup>6,7</sup> For a given condition, which has multiple competing treatments, the um-

**Received:** October 22, 2013 **Revised:** January 13, 2014

**Accepted:** February 15, 2014 **Available online:** October 1, 2014

✉ **Correspondence:** Seong-Hoon Jeong, MD, PhD

Department of Psychiatry, Eulji University School of Medicine, Eulji University Hospital, 95 Dunsanseo-ro, Seo-gu, Daejeon 302-799, Republic of Korea

**Tel:** +82-42-611-3443, **Fax:** +82-42-611-3444

**E-mail:** [anselmjeong@gmail.com](mailto:anselmjeong@gmail.com)

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

umbrella reviews summarize the results of existing systemic reviews. Even though it was helpful to some degree, there was still difficulty in forming a coherent hierarchical view regarding the relative rank of available treatments.

The latest statistical methodology called Mixed Treatment Comparison (MTC) has been introduced in order to bridge this increasing gap between the practical needs of the clinicians and the scarcity of the necessary data.<sup>8,9</sup> This method is currently under development so that different authors use different names such as ‘mixed treatment comparison’, ‘multiple-treatment meta-analysis’, ‘indirect-treatment comparison’, or ‘network meta-analysis’.<sup>9,10</sup> It offers a quantitative approach of integrating all the data from available comparisons, whether they are direct or indirect comparisons.<sup>11-13</sup>

MTC has been widely adopted in all medical fields and successfully applied to various complex problems.<sup>14</sup> Psychiatric literature is no exception. After Cipriani’s pioneering work, which reported the comparative efficacy and acceptability of the latest antidepressants, more works have adapted the similar methodology.<sup>15</sup> Recently, two independent groups of researchers reported the MTC results of efficacy and safety of many antipsychotics.<sup>16,17</sup> In both studies, the authors included the available single-drug placebo-controlled trials and used the placebo group response as a common comparator throughout the evaluation. Unlike the active drug comparator used in the head-to-head studies, the term ‘placebo control’ means heterogeneous treatment settings in different trials and, accordingly, the placebo response has been shown to vary widely among the trials. Therefore, it is questionable to use placebo control group as a common comparator.<sup>18-20</sup>

Since the initial publication of “Aripiprazole versus Other Atypical Antipsychotics for Schizophrenia” in 2009, a group of authors had been publishing a series of umbrella reviews about the relative efficacy and tolerability of a second-generation antipsychotic compared to other second-generation antipsychotics.<sup>21-31</sup> These reviews are unique in such a way that only the head-to-head comparison data are analyzed. These reviews provided not only the list of trials and the aggregated results, but also the summaries of raw data which would hardly be accessible even from the original article. These open data were in ideal format for a MTC procedure.

As a result, we collected and prepared data from these reviews to compile the head-to-head comparison trial data among second-generation antipsychotics. Synthesizing all the gathered evidence via direct or indirect comparisons using MTC procedure, we ranked the second-generation antipsychotics in terms of their relative efficacy, acceptability and tolerability.

## METHODS

### Data sources and selection criteria

The required data was obtained second-handedly from a series of Cochrane reviews. From 2009 to the present, the Cochrane Collaboration has been publishing a series of umbrella reviews about the relative efficacy and tolerability of second-generation antipsychotics.<sup>21-31</sup> As a result, 11 reviews were published covering 9 antipsychotics (amisulpride, aripiprazole, clozapine, olanzapine, quetiapine, risperidone, sertindole, ziprasidone, and zotepine). These Cochrane reviews aimed to be as comprehensive as possible in the collection of available data. They included the randomized controlled trials (RCTs) comparing two or more second-generation antipsychotics directly with each other. Some of the trials were head-to-head studies, which directly compared two antipsychotics, and the others were multi-arm trials that included two or more antipsychotics with a placebo group as the common control.

The only qualified participants of the study were people with schizophrenia and other types of schizophrenia-like psychosis. The applied diagnostic criteria were inconsistent; however, all of the included studies have applied reasonable screening procedures. The durations of the trials ranged from several weeks to several months. In order to deal with this inevitable heterogeneity, the authors of the Cochrane reviews had grouped the trials into short-term (up to 12 weeks), medium-term (13 to 26 weeks), and long-term (over 26 weeks) groups. Nevertheless, the medium and long-term studies were relatively rare and the outcome measures used in these studies were not as diversified as the short-term studies. In this study, the present authors confined the scope only to the short-term studies so as to maximize the homogeneity among the data.

In addition, among the 9 second-generation antipsychotics mentioned in the Cochrane reviews, the number of trials including sertindole and zotepine was so limited (2 for sertindole and 3 for zotepine) that proper analysis seemed impractical. For this reason, only seven antipsychotics (amisulpride, aripiprazole, clozapine, olanzapine, quetiapine, risperidone, and ziprasidone in alphabetical order) were compared in this study.

### Compared outcome measures

The original Cochrane reviews tried to include as many outcome variables as possible and catalogued every outcome measures used in the original studies. However, each study had utilized a unique set of outcome measures so that it was nearly impossible to extract a common set of outcome measures. For instance, certain studies had used the “presence of extrapyramidal symptoms”, while other studies had used “use of antiparkinsonism medication” as a tolerability measure. In

order to perform a proper analysis, an adequate number of trials must be obtained. Therefore, the outcome measures used in this study had to be quite limited. We had selected two efficacy measures (“no clinically important response”, “change in PANSS total score”) and four safety measures (“akathisia”, “use of antiparkinsonism medication”, “clinically important weight gain”, “drop-out due to any reasons”), which have been widely used in the trials.

In the Cochrane reviews, the outcome measure “no clinically important response as defined by the individual studies” was selected as the primary outcome measure. Even though all of the clinical trials included somehow defined “clinically important responses”, these definitions were far from standardized. Since each study applied a different study duration, selection criteria or dosage schedule, we only had to resort to the original authors’ definition of “clinically important response”. In order to compensate for this limitation, the change of the PANSS total score during the trial period was included as an auxiliary outcome measure.

Other outcome measures such as the “use of antiparkinsonism medication” or “presence of akathisia” also suffered from similar problems. Some studies had applied standardized rating scales, while others had not. Even among the studies which had used rating scales, differences abounded in the applied scales and the cut-off points. Therefore, rather than using the “presence of extrapyramidal symptoms” or “presence of parkinsonism”, we decided on the “use of antiparkinsonism medication”. In case of the “presence of akathisia”, we followed the original authors’ decisions to define the presence or absence of akathisia. Fortunately, for weight gain, most trials had used “more than 7% increase in total body weight”; therefore, we made use of this outcome measure.

### Mixed treatment comparison procedure

The computational procedure of MTC is quite complicated and requires a special type of software. It was traditionally carried out by a general-purpose Markov Chain Monte Carlo (MCMC) software such as WinBUGS<sup>32</sup> or JAGS.<sup>33</sup> However, it was a very tedious process and required a detailed knowledge of the Bayesian data analysis. Recently, specialized software called GeMTC has been developed. This software automates the data preparation, model building, script generation, and report display.<sup>34</sup> The data preparation and analysis of this study was done with GeMTC.

We ran four independent chains with different prior distributions each with 50,000 iterations. The chains were checked for convergence by the Brooks-Gelman-Rubin diagnostic tool. The Brooks-Gelman-Rubin score had to be below 1.01 to ensure adequate convergence.<sup>35</sup> The final results were produced from the aggregation of all four chains (200,000 iterations in

total). For the binary variables such as “no clinically important response”, logit values produced from the MTC procedures were converted into odds ratios (ORs) to aid the interpretation. For continuous variables, the weighted mean difference (WMD) values were reported. Through the simple consistency model, we obtained mean ORs and their 95% credible intervals for the pairwise comparison of antipsychotics. This procedure was repeated for all of the pre-determined outcome measures.

In the iteration of each MCMC sampling process, the hierarchies of the compared antipsychotics (superiority or inferiority) can be determined. At the end of the analysis, the probabilities that a certain antipsychotic had ranked the first, second or *n*th position during the whole iteration process (=rank probabilities) could be obtained.<sup>36</sup> From these ranked probabilities, the Surface under the Cumulative Ranking Curve (SUCRA) was calculated as suggested by Salanti et al.<sup>36</sup> The reported rank of the second-generation antipsychotics in terms of their superiority in efficacy and tolerability was determined based on these SUCRA values.

We did not take any measures to incorporate a study-level covariate into the analysis. Furthermore, we did not weigh each study according to their study qualities. These extra measures might mitigate a possible bias due to the heterogeneity or inconsistency. However, implementing these measures into the MTC procedures would be highly sophisticated.

## RESULTS

### The summary of the originally included studies

Overall, a total of 73 independent randomized clinical trials (RCTs) were included. The total number of involved participants in these studies was 15,480. Meanwhile, for each measure variable, only a subset of the RCTs had to be used because all of the RCTs had utilized different sets of measure variables. The number of trials included in each analysis was shown in Table 1.

### Percentage of “no clinically important response”

Forty independent trials had been included in the analysis. All of the second-generation antipsychotics revealed a fairly similar percentage of “no clinically important response”. Comparison pairs for which the 95% confidence interval did not include OR 1.0 were 1) clozapine and ziprasidone (OR=0.62, 95% CI=0.4–0.99), 2) olanzapine and ziprasidone (OR=0.65, 95% CI=0.47–0.93), and 3) amisulpride and ziprasidone (OR=0.62, 95% CI=0.41–0.95). The low OR indicates a lower percentage of “no clinically important response”, thereby the former drug being more effective than the latter drug. In all three comparison pairs, the ziprasidone was the least effective. Although statistically insignificant, the amisulpride, clo-

zapine and olanzapine seemed to be more effective than the aripiprazole, quetiapine and ziprasidone (Table 2). The expected ranks also revealed that the amisulpride, clozapine and olanzapine were the most effective antipsychotics (Table 7).

### Change in total PANSS score

When the actual changes in PANSS total score during the trial period were compared, the differences among the antipsychotics have become more indistinct. The only comparison pair with statistically significant mean difference was between olanzapine and aripiprazole (WMD=-3.15, 95% CI=-5.97-0.06) with olanzapine being more effective. The expected ranks seemed similar to those obtained from comparing the response percentages. Olanzapine, risperidone and clozapine were ranked highest in terms of the change in PANSS total score.

### Percentage of akathisia

Quetiapine was found to be significantly less prone to inducing akathisia compared to aripiprazole (OR=0.46, 95% CI=0.23-0.91) and to risperidone (OR=0.50, 95% CI=0.30-

0.78). Other antipsychotics except quetiapine were fairly similar in their propensity to producing akathisia (Table 3). Although statistically insignificant, risperidone and aripiprazole tended to produce akathisia more frequently than the other drugs. The highest ranked drugs were aripiprazole and risperidone, and the lowest ranked drugs were amisulpride and quetiapine.

### Percentage of antiparkinsonism medication use

A striking contrast was observed among the percentage of antiparkinsonism medication use. Table 4 showed that the participants who took risperidone and ziprasidone were most frequently prescribed with antiparkinsonism medication. In contrast, quetiapine and clozapine had the lowest risk. The percentage of antiparkinsonism medication use of risperidone was significantly higher than clozapine (OR=2.77, 95% CI=1.54-5.16), olanzapine (OR=1.58, 95% CI=1.20-2.11) and quetiapine (OR=2.44, 95% CI=1.51-4.09). In addition, that of ziprasidone was significantly higher than clozapine (OR=2.63, 95% CI=1.33-5.40), olanzapine (OR=1.50, 95% CI=1.01-2.26) and quetiapine (OR=2.33, 95% CI=1.29-4.33). Overall, clozapine, quetiapine and aripiprazole were associated with the lowest risk of antiparkinsonism medication use.

### Percentage of clinically important weight gain

In most studies, the clinically significant weight gain had been defined as more than 7% increase in total body weight during the study period. Although many antipsychotics were known to carry the risk of significant weight gain, our results confirmed the general impression that clozapine and olanzapine have the highest risk. The percentage of significant weight gain in patients who had taken clozapine and olanzapine were significantly higher than any of the remaining drugs. The only insignificant difference was between clozapine and olanzap-

**Table 1.** The number of clinical trials included in MTC analyses

Measured variable	No. of trials
No clinically important response	40
Change in PANSS total score	25
Presence of akathisia	33
Use of antiparkinson medication	38
More than 7% increase in body weight	38
Drop out due to any reasons	56
Total number of analyzed trials	73

Since the original trials had used diverse range of measure variables, the different sets of trial had to be used in the analyses for each measure variable

**Table 2.** Change of PANSS total score during the trial period

	Ami	Ari	Clo	Ola	Que	Ris	Zip
Ami*		-1.93 (-7.06-3.28)	0.41 (-4.77-5.73)	1.21 (-3.49-5.84)	-2.01 (-7.34-3.4)	0.35 (-4.3-5.09)	-1.97 (-6.84-2.92)
Ari*	1.28 (0.86-1.94)*		2.33 (-1.53-6.21)	3.14 (0.03-6.00)	-0.08 (-4.06-4.01)	2.27 (-0.79-5.40)	-0.05 (-3.79-3.55)
Clo*	1.00 (0.67-1.49)*	0.79 (0.50-1.19)*		0.81 (-2.38-3.78)	-2.41 (-6.15-1.39)	-0.06 (-2.99-2.96)	-2.38 (-6.33-1.42)
Ola*	1.05 (0.80-1.4)*	0.83 (0.60-1.11)*	1.05 (0.77-1.46)*		-3.22 (-6.43-0.23)	-0.87 (-3.07-1.65)	-3.19 (-6.60-0.40)
Que*	1.42 (0.96-2.09)*	1.11 (0.72-1.67)*	1.42 (0.94-2.13)*	1.35 (0.98-1.84)*		2.35 (-0.57-5.24)	0.03 (-4.24-4.05)
Ris*	1.22 (0.91-1.62)*	0.95 (0.67-1.31)*	1.21 (0.88-1.66)*	1.15 (0.94-1.41)*	0.86 (0.64-1.13)*		-2.32 (-5.63-0.69)
Zip*	1.61 (1.05-2.43)*	1.26 (0.78-1.98)*	1.61 (1.01-2.52)*	1.53 (1.07-2.13)*	1.13 (0.72-1.77)*	1.33 (0.91-1.91)*	

All the numbers should be read as drugs listed in row compared with those listed in column. For example, in the lower left corner, "1.61 (1.05-2.43)" indicates OR when ziprasidone was compared with amisulpride. In this case, ziprasidone was far more likely to get "no clinically important response", therefore, OR is greater than one. In the upper right corner, "-1.97 (-6.84-2.92)" indicates WMD when amisulpride was compared with ziprasidone. The change of PANSS total score was less negative (=smaller reduction) in ziprasidone, thus WMD is less than one. \*odds ratios (ORs) for the risk of "no clinically important response" and weighted mean differences (WMDs) in change of PANSS total score during the trial period. Ami: amisulpride, Ari: aripiprazole, Clo: clozapine, Ola: olanzapine, Que: quetiapine, Ris: risperidone, Zip: ziprasidone, PANSS: Positive and Negative Syndrome Scale

ine, which suggested that the two drugs have equivalent propensity to induce weight gain (Table 5). The expected ranks revealed that ziprasidone, aripiprazole and amisulpride were relatively safe in terms of weight gain.

**Percentage of dropout due to any reason**

The percentage of dropout due to any reason could be regarded as a summary variable for overall effectiveness combined with tolerance. The most remarkable finding was that the risk of dropout for olanzapine was the lowest and that of ziprasidone was the highest. The remaining drugs revealed fairly similar risk of dropout. Olanzapine carried the significantly lower dropout risk more than any other drugs except for clozapine and amisulpride. Meanwhile, ziprasidone carried the

significantly higher dropout risk compared to any other drugs, with the exception of quetiapine (Table 6).

**Consistency checking of the obtained results**

In order to illustrate the possible inconsistency between direct and indirect comparisons, the consistency checking procedure was performed on the measured variable “no clinically important response”. Although there were no significant differences between the logit values obtained from direct comparison and those obtained from indirect comparison (all of the 95% CIs with difference in the logit values between the two types of comparisons contained zero), several comparisons changed their signs. For example, clozapine was found to be inferior (logit value=0.51) to quetiapine in direct comparison,

**Table 3.** Odds ratios (ORs) for the risk of inducing “akathisia”

	Ami	Ari	Clo	Ola	Que	Ris
Ami						
Ari	2.46 (0.86–6.88)					
Clo	1.51 (0.53–4.97)	0.61 (0.28–1.63)				
Ola	1.74 (0.74–4.41)	0.71 (0.42–1.32)	1.16 (0.57–2.1)			
Que	1.14 (0.41–3.07)	0.46 (0.23–0.91)	0.76 (0.30–1.56)	0.65 (0.36–1.05)		
Ris	2.27 (0.88–5.93)	0.92 (0.52–1.69)	1.50 (0.66–2.96)	1.30 (0.85–1.88)	1.99 (1.28–3.31)	
Zip	1.85 (0.68–5.15)	0.75 (0.38–1.55)	1.22 (0.47–2.73)	1.06 (0.57–1.87)	1.62 (0.86–3.2)	0.81 (0.46–1.42)

All the numbers should be read as drugs listed in row compared with those listed in column. Ami: amisulpride, Ari: aripiprazole, Clo: clozapine, Ola: olanzapine, Que: quetiapine, Ris: risperidone, Zip: ziprasidone

**Table 4.** Odds ratios (ORs) for the risk of using “antiparkinson medication”

	Ami	Ari	Clo	Ola	Que	Ris
Ami						
Ari	0.55 (0.15–2.06)					
Clo	0.43 (0.20–0.90)	0.79 (0.20–3.03)				
Ola	0.76 (0.46–1.26)	1.38 (0.39–4.84)	1.75 (0.98–3.22)			
Que	0.49 (0.25–0.95)	0.89 (0.24–3.34)	1.13 (0.54–2.41)	0.65 (0.38–1.07)		
Ris	1.19 (0.74–1.96)	2.18 (0.65–7.42)	2.77 (1.54–5.16)	1.58 (1.20–2.11)	2.44 (1.51–4.09)	
Zip	1.14 (0.65–2.04)	2.07 (0.57–7.63)	2.63 (1.33–5.40)	1.50 (1.01–2.26)	2.33 (1.29–4.33)	0.95 (0.62–1.46)

All the numbers should be read as drugs listed in row compared with those listed in column. Ami: amisulpride, Ari: aripiprazole, Clo: clozapine, Ola: olanzapine, Que: quetiapine, Ris: risperidone, Zip: ziprasidone

**Table 5.** Odds ratios (ORs) for the risk of “clinically important weight gain” defined as more than 7% increase in total body weight during the study period

	Ami	Ari	Clo	Ola	Que	Ris
Ami						
Ari	0.55 (0.06–4.60)					
Clo	21.03 (3.84–111.86)	38.53 (4.57–331.81)				
Ola	15.06 (4.16–53.81)	27.58 (4.28–184.06)	0.72 (0.22–2.39)			
Que	2.18 (0.44–10.38)	3.99 (0.52–31.21)	0.10 (0.02–0.47)	0.14 (0.05–0.42)		
Ris	1.49 (0.42–5.22)	2.73 (0.47–16.5)	0.07 (0.02–0.25)	0.10 (0.05–0.21)	0.68 (0.23–2.09)	
Zip	0.44 (0.08–2.57)	0.80 (0.09–7.59)	0.02 (0–0.12)	0.03 (0.01–0.10)	0.20 (0.04–0.99)	0.29 (0.07–1.21)

All the numbers should be read as drugs listed in row compared with those listed in column. Ami: amisulpride, Ari: aripiprazole, Clo: clozapine, Ola: olanzapine, Que: quetiapine, Ris: risperidone, Zip: ziprasidone

**Table 6.** Odds ratios (ORs) for the risk of “drop out due to any reasons”

	Ami	Ari	Clo	Ola	Que	Ris
Ami						
Ari	1.10 (0.78–1.55)					
Clo	0.96 (0.73–1.28)	0.87 (0.65–1.19)				
Ola	0.83 (0.66–1.05)	0.75 (0.59–0.97)	0.86 (0.72–1.03)			
Que	1.33 (1.00–1.75)	1.21 (0.91–1.64)	1.38 (1.09–1.77)	1.61 (1.34–1.93)		
Ris	1.15 (0.91–1.46)	1.04 (0.80–1.36)	1.19 (0.97–1.46)	1.38 (1.21–1.58)	0.86 (0.72–1.02)	
Zip	1.54 (1.17–2.03)	1.40 (1.04–1.92)	1.60 (1.26–2.04)	1.86 (1.56–2.22)	1.16 (0.91–1.44)	1.34 (1.10–1.63)

All the numbers should be read as drugs listed in row compared with those listed in column. Ami: amisulpride, Ari: aripiprazole, Clo: clozapine, Ola: olanzapine, Que: quetiapine, Ris: risperidone, Zip: ziprasidone

**Table 7.** The expected ranks of the second generation antipsychotics in terms of each measure variable

	No clinically important response	Change in PANSS total	Akathisia	Use of antiparkinson medication	≥ 7% body weight increase	Dropout d/t any reasons
Amisulpride	1	4	1	5	3	3
Aripiprazole	5	5	7	3	2	4
Clozapine	2	3	3	1	7	2
Olanzapine	3	1	4	4	6	1
Quetiapine	6	7	2	2	5	6
Risperidone	4	2	6	7	4	5
Ziprasidone	7	6	5	6	1	7

The expected ranks were calculated by the surface under the cumulative ranking (SUCRA) method. The higher ranking (ranks 1 or 2) indicates more effective and safest drugs. PANSS: Positive and Negative Syndrome Scale

but superior in indirect comparison (logit value=-0.43). Careful inspection of this inconsistency revealed that the distribution of the sampled log odds ratio for direct comparison was so broad that it lacked sufficient accuracy (Figure 1). Overall, no obvious inconsistency was found between direct and indirect comparison results.

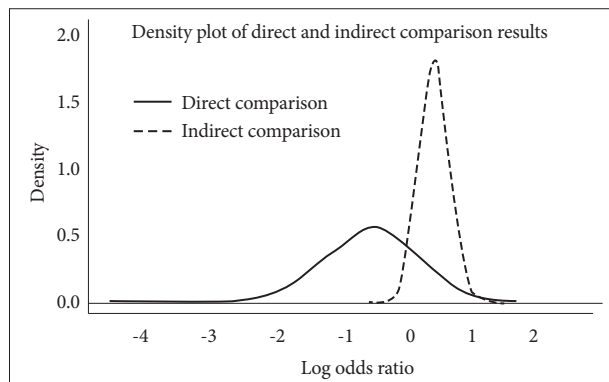
## DISCUSSION

In this study, we tried to integrate data from head-to-head clinical trials of second-generation antipsychotics as comprehensively as possible and to synthesize the evidence in a single integrated analysis framework called MTC. Unlike the traditional pairwise meta-analysis, the MTC framework enables simultaneous comparisons of all the included entities, and, therefore, presents a clearer and complete picture of the comparative hierarchies among second-generation antipsychotics.<sup>6</sup>

The output of the analysis was expressed as a series of pairwise ORs or WMDs. Some of these statistics could not be calculated from the pairwise meta-analysis because relevant clinical trials were rare or simply did not exist. For example, in actual clinical trials, amisulpride had never been compared with aripiprazole, clozapine and quetiapine. The MTC methodology can fill in these lacunae of information. MTC can also provide the most plausible hierarchies among the compared

drugs using the SUCRA method.<sup>36</sup> Using this method, we ranked seven second-generation antipsychotics in terms of their efficacy and risk of major side effects. The obtained hierarchies confirmed the general impression that all antipsychotics have fairly similar efficacy levels but differ widely in their side effects potentials.

In terms of efficacy, amisulpride, clozapine and olanzapine were ranked highest, and aripiprazole, quetiapine and ziprasidone were ranked lowest. Previous MTC results<sup>16,17</sup> also pointed out amisulpride, clozapine and olanzapine as the most effective drugs. However, the superior efficacy of amisulpride is quite unexpected. In a meta-analysis that compared the effect size of second-generation antipsychotics with first-generation antipsychotics, clozapine and amisulpride ranked the first and the second with the greatest effects, respectively, but with a fairly large margin.<sup>1</sup> The MTC results of the more recent study of Leucht et al. also confirmed the superiority of clozapine over amisulpride.<sup>17</sup> As mentioned above, our amisulpride results were based on our insufficient number of comparison trials, because we did not include single-arm placebo controlled studies. The lack of a sufficient number of comparisons precludes any valid conclusion regarding the actual superiority of amisulpride.<sup>25</sup> In this regard, the merits or disadvantages of using a placebo group as the common reference point must be mentioned. Since head-to-head trials are scarce, ex-



**Figure 1.** The inconsistency between the results obtained from direct comparison data and those from indirect comparison data. The lines depict the distribution of the obtained log odds ratios from the two types of data. This illustrative example compared the “percentage of no clinically important result” between quetiapine and clozapine. In direct comparison, quetiapine seemed to have superior efficacy, but in indirect comparison, clozapine was found to be superior. However, the narrowness of the graph indicates that the indirect data is more reliable.

cluding single-arm placebo studies generates the problem of an insufficient number of comparable trials. Also, some authors warn against excluding any kind of treatment modality in the MTC procedure.<sup>37</sup> However, placebo treatment is an artificial situation that is used only in clinical trial settings. It even differs from non-treatment.<sup>38</sup> The placebo response can be highly influenced by the expectations of the participants or by other circumstances irrelevant to the efficacy of the active drug.<sup>39</sup> These points suggest that the placebo response in one clinical trial cannot be simply equated with those in other trials.

Although the percentage of antiparkinson medication was not high with aripiprazole, it ranked as the worst drug in terms of the akathisia potential. In past literature, a relatively higher incidence of akathisia with aripiprazole was demonstrated.<sup>40</sup> However, whether this phenomenon should be interpreted as akathisia or agitation has yet to be settled.<sup>41</sup> Had the clinicians viewed the behavioral change of the patients as agitation but not as akathisia, proper management with antiparkinsonism medication would not have been implemented. Since most clinical trials used antiparkinson drug use as an indirect measure of extrapyramidal symptoms, the apparently low percentage of antiparkinson drug use with aripiprazole might have been due to the misinterpretation of akathisia as agitation or exacerbation of symptoms. Meanwhile, clozapine and quetiapine were preferred as the safest drugs in terms of extrapyramidal symptoms or akathisia; however, they were not desirable in weight gain potential. Ironically, drugs with relatively lower efficacy such as ziprasidone and aripiprazole were less prone to induce weight gain. The odds ratio of weight gain potential between the best drug (ziprasidone) and the worst drug (clozapine) was as high as 50 (Table 5). Despite the

high risk of weight gain, olanzapine and clozapine were ranked highest in terms of the low overall drop-out rate, but quetiapine and ziprasidone were ranked lowest. These results coincide with those of other MTC studies and with the CATIE study that demonstrated that the discontinuation of the treatment for any cause was significantly lower in olanzapine group.<sup>17,42</sup>

The overall validity of the MTC procedure is still being hotly debated. First, for an MTC result to be valid, several assumptions must be met. The most important is the absence of unacceptable heterogeneity.<sup>6</sup> The MTC procedure is more sensitive to homogeneity violation than the traditional meta-analysis, because, in MTC, all the included trials in the treatment network must be comparable—i.e., clinically and methodologically similar.<sup>12,43</sup> However, these requirements are unrealistic, because the inclusion or exclusion criteria, study duration, dosage schedule, used scales and assessment method differ across studies. Strictly imposing a standard, fixed protocol to all future clinical trials for comparability may not be possible or desirable. We have no choice but to interpret the obtained results with caution, bearing in mind the heterogeneity of each study.

Violations of the homogeneity assumption may lead to the second assumption of consistency. Inconsistency points to the contradiction between the direct and the indirect evidence. For example, if A is superior to B and B is superior to C, then A must be superior to C (indirect evidence). If the direct comparison suggests that C is superior, it is inconsistent. No standard analysis method to investigate inherent inconsistency has been agreed upon yet, but Dias et al.<sup>44</sup> proposed the “node-splitting” method to check the consistency of the MTC network. Using this method, we performed an inconsistency check for the “rate of no clinically important response” variable (Figure 1). The visual inspection clarified the relative contributions of the direct and indirect evidence to the final estimates and explained the found inconsistency between the two evidence.

Another assumption is the adequate number of included trials. The number of simultaneously compared trials in this study was at most 56 (for drop out due to any reasons). This number is far from ideal even in terms of traditional meta-analysis. Furthermore, amisulpride, aripiprazole, quetiapine and ziprasidone had rarely been included in head-to-head trials, so that collecting enough number of studies was impossible. With these drugs, the result of single unusual study can perturb the whole result. Ranking probabilities are especially known to be sensitive to small errors.<sup>37</sup> Therefore, not only the point estimate of the relative effects or the most probable rank, but also the margin of error reflected in the confidence interval, had to be carefully inspected.<sup>37</sup> The odds ratios from the comparisons of risperidone, olanzapine and clozapine had nar-

rower confidence intervals, but the other comparisons had broader confidence intervals.

In this study, we used a newer method of meta-analysis called MTC to synthesize the available evidence of the relative merits or shortcomings of second-degree antipsychotics. We confirmed the general impression that the efficacy levels of these drugs are fairly comparable, but their side-effect potentials widely varied. We also demonstrated the usefulness of the MTC framework in summarizing and integrating the available data. The MTC framework is a fast-developing methodology. Its future advancement and refinement may help solidify the evidence-based medicine groundwork.

## REFERENCES

- Davis JM, Chen N, Glick ID. A meta-analysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry* 2003;60:553-564.
- Leucht S, Komossa K, Rummel-Kluge C, Corves C, Hunger H, Schmid F, et al. A meta-analysis of head-to-head comparisons of second-generation antipsychotics in the treatment of schizophrenia. *Am J Psychiatry* 2009;166:152-163.
- Haddad PM, Sharma SG. Adverse effects of atypical antipsychotics: differential risk and clinical implications. *CNS Drugs* 2007;21:911-936.
- Scarff JR, Casey DA. Newer oral atypical antipsychotic agents: a review. *P T* 2011;36:832-838.
- Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *BMJ* 1995;310:1122-1126.
- Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 2009;181:488-493.
- Caldwell DM, Welton NJ, Ades AE. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *J Clin Epidemiol* 2010;63:875-882.
- Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50:683-691.
- Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313-2324.
- Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. 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* 2011;14:429-437.
- Cipriani A, Barbui C, Rizzo C, Salanti G. What is a multiple treatments meta-analysis? *Epidemiol Psychiatr Sci* 2012;21:151-153.
- Tu YK, Faggion CM Jr. A primer on network meta-analysis for dental research. *ISRN Dent* 2012;2012:276520.
- Sobieraj DM, Cappelleri JC, Baker WL, Phung OJ, White CM, Coleman CI. Methods used to conduct and report Bayesian mixed treatment comparisons published in the medical literature: a systematic review. *BMJ Open* 2013;3.
- Assiri A, Al-Majzoub O, Kanaan AO, Donovan JL, Silva M. Mixed treatment comparison meta-analysis of aspirin, warfarin, and new anticoagulants for stroke prevention in patients with nonvalvular atrial fibrillation. *Clin Ther* 2013;35:967-984.
- Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746-758.
- Klemp M, Tvete IF, Skomedal T, Gaasemyr J, Natvig B, Aursnes I. A review and Bayesian meta-analysis of clinical efficacy and adverse effects of 4 atypical neuroleptic drugs compared with haloperidol and placebo. *J Clin Psychopharmacol* 2011;31:698-704.
- Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013;382:951-962.
- Fava M, Evins AE, Dorer DJ, Schoenfeld DA. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom* 2003;72:115-127.
- Kemp AS, Schooler NR, Kalali AH, Alphas L, Anand R, Awad G, et al. What is causing the reduced drug-placebo difference in recent schizophrenia clinical trials and what can be done about it? *Schizophr Bull* 2010;36:504-509.
- Alphas L, Benedetti F, Fleischhacker WW, Kane JM. Placebo-related effects in clinical trials in schizophrenia: what is driving this phenomenon and what can be done to minimize it? *Int J Neuropsychopharmacol* 2012;15:1003-1014.
- Komossa K, Rummel-Kluge C, Hunger H, Schwarz S, Bhoopathi PS, Kissling W, et al. Ziprasidone versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2009:CD006627.
- Komossa K, Rummel-Kluge C, Schwarz S, Schmid F, Hunger H, Kissling W, et al. Risperidone versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2011:CD006626.
- Asenjo Lobos C, Komossa K, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, et al. Clozapine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006633.
- Khanna P, Komossa K, Rummel-Kluge C, Hunger H, Schwarz S, El-Sayeh HG, et al. Aripiprazole versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2013;2:CD006569.
- Komossa K, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, Silveira da Mota Neto JI, et al. Amisulpride versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006624.
- Komossa K, Rummel-Kluge C, Schmid F, Hunger H, Schwarz S, Sri-surapanont M, et al. Quetiapine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006625.
- Komossa K, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, Duggan L, et al. Olanzapine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006654.
- Komossa K, Rummel-Kluge C, Schmid F, Hunger H, Schwarz S, El-Sayeh HG, et al. Aripiprazole versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2009:CD006569.
- Komossa K, Rummel-Kluge C, Hunger H, Schwarz S, Schmidt F, Lewis R, et al. Sertindole versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2009:CD006752.
- Komossa K, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, Kissling W, et al. Zotepine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006628.
- Subramanian S, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, Kissling W, et al. Zotepine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev* 2010:CD006628.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;10:325-337.
- Plummer M. JAGS Version 1.0.3 manual. URL: [http://www-ice.iarc.fr/~martyn/software/jags/jags\\_user\\_manual.pdf](http://www-ice.iarc.fr/~martyn/software/jags/jags_user_manual.pdf) 2009.
- Valkenhoef G, Lu G, Brock B, Hillege H, Ades A, Welton NJ. Automating network meta-analysis. *Res Synth Methods* 2012;3:285-299.
- Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7:434-455.
- Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163-171.
- Mills EJ, Kanfers S, Thorlund K, Chaimani A, Veroniki AA, Ioannidis JP. The effects of excluding treatments from network meta-analyses: survey. *BMJ* 2013;347:f5195.



38. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. *J Intern Med* 2004;256:91-100.
39. Enck P, Benedetti F, Schedlowski M. New insights into the placebo and nocebo responses. *Neuron* 2008;59:195-206.
40. Kerwin R, Millet B, Herman E, Banki CM, Lublin H, Pans M, et al. A multicentre, randomized, naturalistic, open-label study between aripiprazole and standard of care in the management of community-treated schizophrenic patients Schizophrenia Trial of Aripiprazole: (STAR) study. *Eur Psychiatry* 2007;22:433-443.
41. Takeuchi H, Remington G. A systematic review of reported cases involving psychotic symptoms worsened by aripiprazole in schizophrenia or schizoaffective disorder. *Psychopharmacology (Berl)* 2013;228:175-185.
42. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 2005;353:1209-1223.
43. Jansen JP, Cope S. Meta-regression models to address heterogeneity and inconsistency in network meta-analysis of survival outcomes. *BMC Med Res Methodol* 2012;12:152.
44. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-944.