Efficacy, Acceptability, and Tolerability of Antipsychotics in Treatment-Resistant Schizophrenia: A Network Meta-analysis

Myrto T. Samara, MD; Markus Dold, MD; Myrsini Gianatsi, MSc; Adriani Nikolakopoulou, MSc; Bartosz Helfer, MSc; Georgia Salanti, PhD; Stefan Leucht, MD

Importance
In treatment-resistant schizophrenia, clozapine is considered the standard treatment. However, clozapine use has restrictions owing to its many adverse effects. Moreover, an increasing number of randomized clinical trials (RCTs) of other antipsychotics have been published.

Objective
To integrate all the randomized evidence from the available antipsychotics used for treatment-resistant schizophrenia by performing a network meta-analysis.

Data Sources
MEDLINE, EMBASE, BIOSIS, PsycINFO, PubMed, Cochrane Central Register of Controlled Trials, World Health Organization International Trial Registry, and clinicaltrials.gov were searched up to June 30, 2014.

Study Selection
At least 2 independent reviewers selected published and unpublished single- and double-blind RCTs in treatment-resistant schizophrenia (any study-defined criterion) that compared any antipsychotic (at any dose and in any form of administration) with another antipsychotic or placebo.

Data Extraction and Synthesis
At least 2 independent reviewers extracted all data into standard forms and assessed the quality of all included trials with the Cochrane Collaboration’s risk-of-bias tool. Data were pooled using a random-effects model in a Bayesian setting.

Main Outcomes and Measures
The primary outcome was efficacy as measured by overall change in symptoms of schizophrenia. Secondary outcomes included change in positive and negative symptoms of schizophrenia, categorical response to treatment, dropouts for any reason and for inefficacy of treatment, and important adverse events.

Results
Forty blinded RCTs with 5172 unique participants (71.5% men; mean [SD] age, 38.8 [3.7] years) were included in the analysis. Few significant differences were found in all outcomes. In the primary outcome (reported as standardized mean difference; 95% credible interval), olanzapine was more effective than quetiapine (−0.29; −0.56 to −0.02), haloperidol (−0.29; −0.44 to −0.13), and sertindole (−0.46; −0.80 to −0.06); clozapine was more effective than haloperidol (−0.22; −0.38 to −0.07) and sertindole (−0.40; −0.74 to −0.04); and risperidone was more effective than sertindole (−0.32; −0.63 to −0.01). A pattern of superiority for olanzapine, clozapine, and risperidone was seen in other efficacy outcomes, but results were not consistent and effect sizes were usually small. In addition, relatively few RCTs were available for antipsychotics other than clozapine, haloperidol, olanzapine, and risperidone. The most surprising finding was that clozapine was not significantly better than most other drugs.

Conclusions and Relevance
Insufficient evidence exists on which antipsychotic is more efficacious for patients with treatment-resistant schizophrenia, and blinded RCTs—in contrast to unblinded, randomized effectiveness studies—provide little evidence of the superiority of clozapine compared with other second-generation antipsychotics. Future clozapine studies with high doses and patients with extremely treatment-refractory schizophrenia might be most promising to change the current evidence.

Published online February 3, 2016.
A

s many as one-third of patients with schizophrenia experience persistent psychotic symptoms despite adequate treatment with antipsychotics. At present, clozapine is considered the standard treatment for patients with refractory forms of schizophrenia, which has been codified in international guidelines. Unanimity concerning the superiority of clozapine is based mainly on the landmark study by Kane et al which resulted in an impressive effect size of clozapine compared with chlorpromazine hydrochloride and led to clozapine’s reintroduction in clinical practice. Although subsequent comparisons with first-generation antipsychotics (FGAs) showed smaller effect sizes than the dramatic superiority in the study by Kane et al, the superiority of clozapine compared with FGAs for treatment-resistant schizophrenia has been corroborated by an early meta-analysis of 12 trials and a Cochrane review. Moreover, researchers found that clozapine was the most efficacious antipsychotic in a network meta-analysis (NMA) of trials for treatment of nonrefractory schizophrenia, but the superiority was again driven by old comparisons with FGAs. Indeed, in a Cochrane review using a simple pairwise meta-analysis, clozapine was not more efficacious than the second-generation antipsychotics (SGAs) olanzapine, risperidone, or ziprasidone.

The reintroduction of clozapine led to the development of numerous other SGAs, some of which showed superior efficacy in meta-analyses not restricted to patients with treatment-resistant disease. Moreover, effectiveness studies suggested that some older agents, such as perphenazine or sulpiride, might be as efficacious as SGAs, but their effects in treatment-resistant schizophrenia are unknown. Given the restrictions of clozapine use owing to its many adverse effects, such as agranulocytosis, metabolic problems, or sedation, and in light of a continuously increasing number of randomized clinical trials (RCTs) of other SGAs, a new systematic review and meta-analysis on the effects of the various antipsychotics in treatment-resistant schizophrenia appeared to be necessary.

In an attempt to integrate all the evidence from the available antipsychotics in treatment-resistant schizophrenia, we performed simple pairwise meta-analyses and an NMA (also called a multiple-treatments meta-analysis), a technique that allows the comparison of the relative effectiveness among all antipsychotics that have been examined in at least 1 trial. A particular advantage of this technique is that one can obtain an effect estimate for the relative effectiveness between 2 interventions, even if they have not been compared directly in any trial, as long as they are part of a connected network. This technique uses indirect comparison of the 2 treatments of interest via 1 or more intermediate comparators. Moreover, even when outcome data for a particular comparison exist, the use of all available information, direct and indirect, renders the effect estimate more precise than the pairwise meta-analysis counterpart. Such advantages have established NMA as a powerful evidence synthesis tool, and its use is crucial in health care areas where many drugs are available, such as in treatment-resistant schizophrenia.

Key Points

Question: What is the most effective and acceptable antipsychotic for treatment-resistant schizophrenia?

Findings: A certain pattern of superiority was found for olanzapine, clozapine, and risperidone in various efficacy outcomes, such as mean change in overall and positive symptoms, response rates, and dropouts owing to inefficacy, but treatment efficacy differences were not definitive and clozapine was no more efficacious than most other second-generation antipsychotics. Although substantial evidence was available for clozapine, haloperidol, olanzapine, and risperidone, data on the other included antipsychotics were limited.

Meaning: Insufficient evidence exists on which antipsychotic is more efficacious for patients with treatment-resistant schizophrenia.

Methods

Participants and Interventions
An a priori written study protocol can be found in eAppendix 1 in the Supplement. Our analysis included all single- and double-blind RCTs of adult patients with a treatment-resistant form of schizophrenia, schizophreniform disorder, or schizoaffective disorder. We planned a subgroup analysis a priori based on the degree of treatment resistance (discussed below in the Statistical Analysis subsection). All worldwide available antipsychotics, at any dose and in any form of administration that were compared with another antipsychotic or placebo, were included if the antipsychotics were used as monotherapy. Minimum duration of RCTs was set at 3 weeks. Institutional review board approval was not necessary for this review.

Search Strategy and Selection Criteria
We identified RCTs in patients with treatment-resistant schizophrenia through a comprehensive, systematic literature search in MEDLINE, EMBASE, Biosis, PsycINFO, PubMed, Cochrane Central Register of Controlled Trials, World Health Organization International Trial Registry, and clinicaltrials.gov up to June 30, 2014 (eAppendix 2 in the Supplement). Moreover, we inspected the reference lists of the included studies and previous reviews on antipsychotics in general schizophrenia. We included published and unpublished blinded RCTs because unblinded trials systematically favored SGAs in a previous analysis. In the case of crossover studies, we only used the first crossover phase to avoid the problem of carryover effects. We excluded cluster-randomized trials. Furthermore, studies that demonstrated a high risk for bias for sequence generation or allocation concealment were excluded. If a trial was described as double-blind but randomization was not explicitly mentioned, we assumed that study participants were randomized, and we excluded the trial in a sensitivity analysis. Study quality was independently assessed by 2 reviewers (M.D. and B.H.), who used the Cochrane Collaboration’s risk-of-bias tool.

We excluded studies from mainland China to avoid a systematic bias because many of these studies do not use appropriate randomization procedures and do not report their
methods.24,25 We sent emails to the first and corresponding authors of all included studies to ask for missing data.

Outcome Measures and Data Extraction
The primary outcome was the mean change from baseline to end point in overall symptoms of schizophrenia as measured by the Positive and Negative Syndrome Scale (PANSS),26 the Brief Psychiatric Rating Scale,27 or any other validated scale for the assessment of overall schizophrenia symptoms. If change data were not available, we used the mean score at study end point of these scales. Intention-to-treat data sets were used whenever available.

Secondary outcome was a clinically important response to treatment that was defined primarily as at least a 20% reduction of PANSS or Brief Psychiatric Rating Scale score or at least minimal improvement on the Clinical Global Impressions Scale.28,29 In patients with treatment-resistant schizophrenia, even a slight improvement might represent a clinically significant effect, justifying the use of the 20% cutoff, which roughly corresponds to at least minimal improvement on the Clinical Global Impressions Scale.30 For the typical patient with acute schizophrenia, the 50% cutoff, corresponding to much improvement on the Clinical Global Impressions Scale, might be preferable.31 Other secondary outcomes were the mean change in positive and negative symptoms of schizophrenia, dropouts owing to any reason (all-cause discontinuation), dropouts owing to inefficacy of treatment, the occurrence of important adverse effects (ie, weight gain, extrapyramidal symptoms, and sedation), quality of life, ability to work, and economic outcomes.

Study selection and data extraction were performed independently by at least 2 reviewers (M.T.S., M.D., B.H., and S.L.). Missing SDs were estimated from P values or substituted by the mean SD of the other included studies.

Statistical Analysis
Data were analyzed from October 13, 2014, to March 3, 2015. We performed pairwise meta-analyses and NMA in a Bayesian setting using WinBUGS.32-35 We used the random-effects model and assumed common heterogeneity across all comparisons. For continuous outcomes, the effect sizes were calculated as Hedges adjusted g standardized mean differences (SMDs).36 For binary outcomes, the effect sizes were calculated as odds ratios (ORs). Both types of effect sizes were presented along with their 95% credible intervals (CRIs). To enhance understandability, SMDs were also transformed in original units and ORs in numbers needed to benefit (NNTB) and numbers needed to harm (NNTH), which were estimated with the mean occurrence of an outcome as the baseline risk. The NMA combined direct and indirect evidence of all relative treatment effects, provided estimates with maximum power, and allowed the ranking of the various antipsychotics based on the surface under the cumulative ranking and the mean ranks.37

The assumption underlying the performance of NMA is transitivity, which suggests that all pairwise comparisons in the network do not differ with respect to the distribution of effect modifiers.19,38,39 Important intransitivity might be manifested in the data as inconsistency between direct and indirect evidence. Consistency was assessed by calculating the difference between direct and indirect treatment effects in all closed loops and assuming loop-specific heterogeneity.40-43 The magnitude of inconsistency factors and their respective P values were used to examine the presence of inconsistency in each loop; we judged loops to be inconsistent with a significant disagreement between direct and indirect evidence (P < .10). To assess the evidence of inconsistency in the entire network, we used the design-by-treatment model,44-46 which enabled us to examine the presence of loop and design inconsistency.

To assess the effect of potential moderator and quality variables on the primary outcome, we planned several meta-regression, subgroup, and sensitivity analyses in advance. In separate network metaregression models, the variables’ publication date, quality of randomization and allocation concealment, dose of antipsychotics in chlorpromazine equivalents,47 and severity of illness at baseline were included as covariates. Other analyses were added post hoc, including the dose of antipsychotics using 2 alternative methods to calculate chlorpromazine equivalents,48,49 dose of clozapine, trial duration, and study sponsorship.

The definition of treatment resistance varied across studies. In a preplanned subgroup analysis, the included studies were classified into the following 3 groups according to the applied criteria for defining treatment resistance: (1) no response to 1 previous antipsychotic, (2) no response to at least 2 retrospective periods of antipsychotic treatment, and (3) no response to a combination of retrospective and prospective criteria for treatment resistance.

In preplanned sensitivity analyses, we used the fixed-effect model, and we excluded trials with only completer data, trials that included patients who previously demonstrated treatment intolerance, trials that did not apply operationalized criteria to diagnose schizophrenia, and single-blind trials. In a post hoc sensitivity analysis, we performed pairwise meta-analyses and NMA for the primary outcome that included 5 studies published before 1990 that had to be excluded from the main analysis because they led to inconsistency (described in the Consistency of the Network subsection of the Results). The motivation underlying this sensitivity analysis was to show the full evidence of antipsychotic efficacy and to justify that the decision to exclude studies conducted before 1990 in the main analysis did not introduce bias. Other post hoc sensitivity analyses involved the exclusion of trials with a high risk for attrition bias and those with a high risk for reporting bias.

We investigated the presence of small-study effects for the primary outcome to assess whether important differences in treatment effect estimates were found between precise and imprecise studies. We produced 2 comparison-adjusted funnel plots; comparisons have been directed according to the age and effectiveness of the treatments, assuming that newer and more effective treatments, respectively, are favored in small trials.50,51 Potential asymmetry would indicate a form of small-study effects depending on the defined direction, whereas symmetry in the funnel plot would indicate a lack of evidence of small-study effects.
Results

Description of Included Studies
We identified 40 unique blinded RCTs \(^6-8,51-87\) with 5172 unique participants published from 1968 to March 2012 through the literature search. The PRISMA flowchart is shown in Figure 1 and with details of all included studies in eAppendix 3 in the Supplement. Of 4813 patients with sex indicated, 3442 were men (71.5%). The mean (SD) age of participants was 38.8 (3.7) years; the mean (SD) duration of illness, 16.2 (4.3) years; and the mean (SD) number of previous hospitalizations, 6.9 (3.1). The median trial duration was 11 weeks. The assessment of risk for bias is presented in eAppendix 4 in the Supplement. Overall, the trial reports often did not provide details about randomization procedures and allocation concealment; 4 studies \(^69,76,78,84\) were single-blind and the remaining were double-blind. The mean dropout rate was 30.2% for the studies included in our analysis, and we found some indication of selective reporting in 18 studies (45.0%). Figure 2 shows the network of eligible comparisons for the primary outcome; network plots for secondary outcomes are presented in eAppendix 5 in the Supplement. The mean dose of the 12 antipsychotics included in the analysis was 794 mg/d in chlorpromazine equivalents.\(^{47}\) The drug involved in most comparisons was clozapine (20 of 40 trials), followed by haloperidol (15 of 40 trials), olanzapine (14 of 40 trials), and risperidone (12 of 40 trials), whereas few trials were available for most other drugs. Three antipsychotics (aripiprazole, perphenazine, and thiothixene hydrochloride), although included in the systematic review, were not included in the meta-analysis because one of the 2 relevant studies was not connected to the network\(^{51}\) and the other\(^{52}\) did not provide usable data (further explanations can be found in eAppendix 3 in the Supplement).

Consistency of the Network
When we integrated the evidence from all trials in the NMA, we found a severe inconsistency between direct and indirect evidence (eAppendix 6 in the Supplement). Inspection of the data showed that the inconsistency was owing to 3 loops that included 2 older studies published before 1990.\(^5,6,53\) Because the methods of antipsychotic trials changed greatly\(^88-90\) after the study by Kane et al,\(^5,6\) cohort effects were the most likely explanation (explained in detail below in the Discussion section). Because consistency (transitivity) is a central assumption of NMA, we had to remove the 5 trials published\(^5,6,52-55\) before 1990 from the network of all outcomes post hoc, which resolved the inconsistency in all outcomes (\(P > .10\)). Assessment of inconsistency is presented in detail in eAppendix 7 in the Supplement. Moreover, we performed a sensitivity analysis of the primary outcome, including all trials (described in the Sensitivity Analyses of the Primary Outcome subsection below and eAppendix 6 in the Supplement).

Primary Outcome
The results of the pairwise meta-analysis and the NMA for the primary outcome, mean change in overall symptoms, are summarized in Figure 3. The surface under the cumulative ranking is given in eAppendix 8 in the Supplement. Drugs are reported in Figure 3 in order of their efficacy ranking. As a rule of thumb, Cohen\(^{91}\) recommended that an SMD of −0.20 is small, −0.50 is medium, and −0.80 is large. Few statistically significant differences were found. In the simple pairwise meta-analytic comparisons, only olanzapine was significantly better than haloperidol (SMD, −0.29, corresponding to −6.08 PANSS points; 95%
### Positive and Negative Symptoms

Results for the secondary outcome, reduction in positive symptoms, are shown in eAppendix 9 in the Supplement. In pairwise comparisons, only risperidone was statistically significantly better than fluphenazine hydrochloride and quetiapine (SMDs, −0.73 [95% CRI, −1.48 to −0.02] and −0.93 [95% CRI, −1.72 to −0.11], respectively; corresponding to −5.16 and −6.57 PANSS points, respectively) based on a single small trial with 38 participants. In the NMA, risperidone, clozapine, and olanzapine were significantly more efficacious than haloperidol (SMDs, −0.83 [95% CRI, −1.90 to −0.75], −0.67 [95% CRI, −1.60 to −0.04], and −0.67 [95% CRI, −1.60 to −0.04], respectively; corresponding to −2.42 and −2.33 PANSS points, respectively). In addition, risperidone and clozapine were significantly more efficacious than haloperidol (SMDs, −0.29 [95% CRI, −0.54 to −0.07] and −0.27 [95% CRI, −0.46 to −0.09], respectively; corresponding to −1.91 and −2.05 PANSS points, respectively).

Results for the secondary outcome, reduction in negative symptoms, are shown in eAppendix 9 in the Supplement. In pairwise comparisons, olanzapine was significantly more efficacious than haloperidol (SMD, −0.26, corresponding to −1.46 PANSS points; 95% CRI, −0.84 to −0.02) and ziprasidone (SMD, −0.04, corresponding to −0.26 PANSS points; 95% CRI, −0.33 to −0.001). In the NMA, olanzapine was better than clozapine.

---

**Table: Pairwise and Network Meta-Analysis Comparisons**

<table>
<thead>
<tr>
<th>Drug</th>
<th>-0.09 (0.28 to 0.12) (7 studies; 596 patients)</th>
<th>-0.21 (-0.73 to 0.29) (1 study; 841 patients)</th>
<th>-0.19 (-0.61 to 0.22) (2 studies; 112 patients)</th>
<th>-0.29 (-0.53 to -0.06) (4 studies; 693 patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olanzapine</td>
<td>-0.04 (-0.34 to 0.24)</td>
<td>-0.02 (-0.44 to 0.40) (1 study; 147 patients)</td>
<td>-0.07 (-0.42 to 0.28) (1 study; 306 patients)</td>
<td>-0.15 (-0.37 to 0.10) (4 studies; 646 patients)</td>
</tr>
<tr>
<td>Ziprasidone</td>
<td>-0.07 (-0.21 to 0.08)</td>
<td>-0.02 (-0.29 to 0.26)</td>
<td>-0.12 (-0.41 to 0.12)</td>
<td>-0.23 (-0.50 to 0.03) (2 studies; 496 patients)</td>
</tr>
<tr>
<td>Clozapine</td>
<td>-0.12 (-0.33 to 0.08)</td>
<td>-0.08 (-0.31 to 0.17)</td>
<td>-0.05 (-0.30 to 0.18)</td>
<td>-0.12 (-0.92 to 0.68) (1 study; 25 patients)</td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>-0.14 (-0.41 to 0.23)</td>
<td>-0.10 (-0.41 to 0.23)</td>
<td>-0.08 (-0.25 to 0.10)</td>
<td>-0.12 (-0.58 to 0.12) (3 studies; 165 patients)</td>
</tr>
<tr>
<td>Risperidone</td>
<td>-0.29 (-0.56 to -0.02)</td>
<td>-0.24 (-0.54 to 0.05)</td>
<td>-0.22 (-0.47 to 0.02)</td>
<td>-0.14 (-0.50 to 0.22) (1 study; 288 patients)</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>-0.29 (-0.44 to -0.13)</td>
<td>-0.25 (-0.53 to 0.05)</td>
<td>-0.22 (-0.38 to -0.07)</td>
<td>-0.14 (-0.50 to 0.22) (1 study; 288 patients)</td>
</tr>
<tr>
<td>Haloperidol</td>
<td>-0.38 (-0.85 to 0.03)</td>
<td>-0.35 (-0.88 to 0.17)</td>
<td>-0.33 (-0.79 to 0.12)</td>
<td>-0.25 (-0.71 to 0.21) (1 study; 47 patients)</td>
</tr>
<tr>
<td>Fluphenazine</td>
<td>-0.46 (-0.80 to -0.06)</td>
<td>-0.42 (-0.85 to 0.04)</td>
<td>-0.40 (-0.74 to -0.04)</td>
<td>-0.32 (-0.63 to -0.01) (1 study; 165 patients)</td>
</tr>
<tr>
<td>Sertindole</td>
<td>-0.40 (-0.75 to 0.10)</td>
<td>-0.34 (-0.75 to 0.10)</td>
<td>-0.32 (-0.63 to -0.01)</td>
<td>-0.18 (-0.59 to 0.25) (1 study; 288 patients)</td>
</tr>
</tbody>
</table>

CRI, −0.53 to −0.06 based on 4 studies

In the NMA, olanzapine was significantly more effective than quetiapine fumarate (SMD, −0.29, corresponding to −6.08 PANSS points; CRI, −0.56 to −0.02), haloperidol (SMD, −0.29, corresponding to −6.08 PANSS points; CRI, −0.44 to −0.13), and sertindole (SMD, −0.46, corresponding to −9.64 PANSS points; 95% CRI, −0.80 to −0.06); clozapine was significantly more effective than haloperidol (SMD, −0.22, corresponding to −4.61 PANSS points; 95% CRI, −0.38 to −0.07) and sertindole (SMD, −0.40, corresponding to −8.39 PANSS points; 95% CRI, −0.74 to −0.04); and risperidone was significantly more effective than sertindole (SMD, −0.32, corresponding to −6.71 PANSS points; 95% CRI, −0.63 to −0.01). In terms of rankings, olanzapine was ranked first, followed by ziprasidone and clozapine, but results for ziprasidone were based on only 2 studies (n = 433) and did not show a significant difference compared with any other antipsychotic (Figure 3).
### Figure 4. Secondary Outcome of Response Rates

<table>
<thead>
<tr>
<th>Treatment</th>
<th>OR (95% CRI)</th>
<th>Study Count</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ziprasidone</td>
<td>1.11 (0.49 to 2.57)</td>
<td>1 study; 106 patients</td>
<td>95% CRI, 0.11-4.73</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>7.17 (0.34 to 143.12)</td>
<td>1 study; 84 patients</td>
<td>95% CRI, 0.14-183.46</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>1.77 (0.74 to 4.19)</td>
<td>1 study; 288 patients</td>
<td>95% CRI, 0.23-14.60</td>
</tr>
<tr>
<td>Sertindole</td>
<td>1.61 (0.13 to 50.00)</td>
<td>1 study; 47 patients</td>
<td>95% CRI, 0.18-21.50</td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>1.77 (0.74 to 4.19)</td>
<td>1 study; 288 patients</td>
<td>95% CRI, 0.23-14.60</td>
</tr>
<tr>
<td>Haloperidol</td>
<td>1.86 (1.01 to 4.00)</td>
<td>2 studies; 494 patients</td>
<td>95% CRI, 0.61-3.03</td>
</tr>
</tbody>
</table>

Pairwise (upper right portion) and network (lower left portion) meta-analytic results are shown for the secondary outcome. Drugs are reported in order of efficacy ranking, and outcomes are expressed as odds ratios (ORs) (95% credible intervals). For the pairwise meta-analyses, ORs of greater than 1 indicate that the treatment specified on the row has more responders. For the network meta-analysis (NMA), ORs of greater than 1 indicate that the treatment specified in the column has more responders than the treatment on the row. Bold results indicate statistical significance. The overall heterogeneity (τ) is 0.27 for pairwise results and 0.30 for NMA results. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken.

(SMD, −0.14, corresponding to −0.79 PANSS points; 95% CRI, -0.30 to -0.01), risperidone (SMD, −0.24, corresponding to −1.35 PANSS points; 95% CRI, −0.44 to −0.02), haloperidol (SMD, −0.24, corresponding to −1.35 PANSS points; 95% CRI, −0.40 to −0.04), chlorpromazine (SMD, −0.26, corresponding to −1.46 PANSS points; 95% CRI, −0.51 to −0.02), and sertindole (SMD, −0.44, corresponding to −2.48 PANSS points; 95% CRI, −0.81 to −0.08). Ziprasidone was better than chlorpromazine (SMD, −0.26, corresponding to −1.46 PANSS points; 95% CRI, −0.53 to −0.04) and sertindole (SMD, −0.44, corresponding to −2.48 PANSS points; 95% CRI, −0.88 to −0.01).

### Discontinuation
Treatment discontinuation owing to all causes was used as a measure of acceptability. Results are shown in Appendix 9 in the Supplement. In pairwise comparisons, only olanzapine was better than haloperidol (OR, 0.52; 95% CRI, 0.24–0.97). In the NMA, no difference among antipsychotics was found apart from olanzapine being better than haloperidol (OR, 0.56; 95% CRI, 0.33–0.87; NNTH, 9) and fluphenazine (OR, 0.24; 95% CRI, 0.03–0.87; NNTH, 5).

In pairwise comparisons of discontinuation owing to inefficacy, clozapine was better than risperidone (OR, 0.32; 95% CRI, 0.14–0.81) and haloperidol (OR, 0.18; 95% CRI, 0.08–0.46), and olanzapine was better than haloperidol (OR, 0.32; 95% CRI, 0.10–0.99). In the NMA, clozapine was better than risperidone, quetiapine, haloperidol, and fluphenazine (OR range, 0.44 [95% CRI, 0.19–0.91] to 0.08 [95% CRI, 0.01–0.35]; NNTH range, 6–10); chlorpromazine and olanzapine were better than haloperidol and fluphenazine (OR range, 0.04 [95% CRI, 0.01–0.76] to 0.27 [95% CRI, 0.11–0.60]; NNTH range, 5–7); and risperidone was better than fluphenazine (OR, 0.19; 95% CRI, 0.02–0.81; NNTH, 7) (eAppendix 9 in the Supplement).

### Categorical Response to Treatment
The ORs for the secondary outcome, categorical response to treatment, are presented in Figure 4. In the pairwise comparisons, findings were significantly better for risperidone (OR, 9.68; 95% CRI, 1.11–183.46) and clozapine (OR 1.86; 95% CRI, 1.01–4.00) compared with haloperidol. In the network comparisons, significantly better results were found for risperidone (OR, 2.27; 95% CRI, 1.11–4.73), clozapine (OR, 2.09; 95% CRI, 1.26–3.82), and olanzapine (OR, 2.00; 95% CRI, 1.16–3.76) compared with haloperidol (NNTBs, 7, 8, and 8, respectively).
Adverse Events
Few significant differences were found in terms of adverse events (eAppendix 9 in the Supplement). As expected, the patterns were generally consistent with the previous NMA in patients with non-treatment-resistant schizophrenia, but fewer findings were statistically significant, which may be attributed to the lower number of included participants.

Other Outcomes
Only 5 studies provided data on quality of life. The pairwise meta-analysis did not indicate any significant difference among antipsychotics (eAppendix 9 in the Supplement), whereas conducting an NMA was not feasible. No data were available for ability to work and economic outcomes.

Sensitivity Analyses of the Primary Outcome
The most important sensitivity analysis was the inclusion of the 5 old studies that were excluded from the primary analysis owing to a violation of the consistency assumption. Including these trials did not change the results much (eAppendix 6 in the Supplement). The only change in pairwise comparisons was that clozapine was now significantly better than chlorpromazine (sensitivity analysis SMD, −0.74 [95% CRI, −0.99 to −0.49] vs main analysis SMD, −0.44 [95% CRI, −1.10 to 0.22]). In the NMA, the main change was a few differences in terms of statistical significance. Olanzapine was ranked first as in the main analysis, whereas clozapine was ranked second compared with third in the main analysis, but clozapine was still not significantly superior to all other SGAs. In the remaining sensitivity analyses, SMDs and rankings did not change considerably, or so few studies remained that hardly any result was significant (eAppendix 10 in the Supplement).

Subgroup and Metaregression Analyses for the Primary Outcome
For the degree of treatment resistance, the 3 subgroups of studies based on the criteria for defining treatment resistance (described in the Statistical Analysis subsection of the Methods section) were formed. No significant efficacy difference was shown among antipsychotics in any of these subgroups, probably because the available data per group were less than that of the main analysis (eAppendix 10 in the Supplement). Similar results were obtained when all trials, irrespective of their publication date, were included in this subgroup analysis (eAppendix 6 in the Supplement).

Moreover, we found no significant effects in metaregressions examining the impact of the possible effect modifiers (eAppendix 10 in the Supplement), but these analyses also had limited statistical power, and any interpretation should be made with great caution.

Small-Study Effects, Original Units, NNTBs, and NNTHs
We found no evidence of small-study effects for the primary outcome (eAppendix 11 in the Supplement). The SMDs back transformed to original units, and NNTBs, and NNTHs can be found in eAppendices 12 and 13 in the Supplement.

Discussion
We compared the effects of all antipsychotics in patients with treatment-resistant schizophrenia using pairwise meta-analyses and NMA. Olanzapine, clozapine, and risperidone were found to be significantly better than some other antipsychotics in various efficacy outcomes, but the results were not consistent and the effect sizes were usually small. The most surprising finding was that clozapine was not significantly more efficacious than most other drugs.

Clozapine’s superiority was originally demonstrated in a pivotal study in which it was clearly superior to chlorpromazine in treatment-resistant schizophrenia. Although some subsequent comparisons with FGAs were also statistically significant and although the superiority to FGAs has been confirmed by meta-analyses, the effect size of the original study by Kane et al. (−0.88) has never been replicated. Figure 5 presents the results of all single-comparison clozapine trials and illustrates that, of 21 comparisons, only 2 old studies compared with chlorpromazine led to a high inconsistency in the NMA, violating a key assumption of the method. We speculate that the inconsistency is in part owing to cohort effects in terms of the periods when these studies were performed because some evidence suggests that the clinical trial quality of psychopharmacologic studies changed significantly after 1990. This finding is also evident, among others, by an increasing placebo response and smaller drug-placebo differences. Nevertheless, when all trials, irrespective of their publication date, were included, results of the NMA did not substantially differ (eAppendix 6 in the Supplement).

Owing to several limitations, our NMA is not definitive. Clozapine was superior in 3 large effectiveness studies that could not be included. Essock et al. published a large cost-effectiveness study in which clozapine was superior to the group of FGAs in many measures of effectiveness, but the study was unblinded. Similarly, in phase 2 of the Clinical Antipsychotic Trials in Intervention Effectiveness (CATIE), clozapine was shown to be more effective than risperidone and quetiapine, but clozapine was the only open-label treatment arm. In the Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study Band (CUTLASS2), clozapine was found to be more effective than other SGAs as a group, but again the trial was unblinded and comparisons with individual SGAs were not presented. One could claim that, in open-label studies, expectation bias could be a reason for clozapine’s superiority, but results from real-world observational studies and population-based register studies confirm this finding, even regarding hard end points such as hospitalization and number of inpatient days. In these studies, better monitoring of patients receiving clozapine (eg, regular blood tests) is a possible bias. Therefore, another possible explanation for not finding clozapine superior to other SGAs is that patients who could benefit from clozapine the most (patients with the most treatment-
resistant disease) are rarely included in a blinded RCT, which causes sampling bias.

The same explanation could apply to clozapine’s dramatic superiority found in the study by Kane et al.5,6 At the time of their study, clozapine was off the market owing to agranulocytosis, so only patients considered “beyond the reach of conventional therapy”6 were recruited. In our analysis, most included studies were conducted at a time when less severely ill patients were generally enrolled in blinded RCTs, with more onerous consenting procedures and commercialization of the research conduct. Therefore, the criteria of treatment resistance varied from partial nonresponse to high levels of treatment resistance, and synthesizing this evidence could be problematic. For example, most trials examining olanzapine applied a more liberal definition of treatment resistance than clozapine trials, highlighting the shortcomings when comparing both compounds in a meta-analysis. The subgroup analysis based on the degree of treatment resistance failed to detect any significant efficacy difference among the various antipsychotics, but the data per group were very limited, which increased the possibility of failing to detect a difference, although it might be present (type II error).

Furthermore, some evidence shows clozapine having a dose-related efficacy at as much as 600 mg/d in patients with treatment-resistant schizophrenia.102-104 Also, recent guidelines1,4 underline the importance of adequate plasma concentrations of clozapine (>350 ng/mL; to convert to micromoles per liter, multiply by 0.003) and treatment duration (≥8 weeks) in patients failing to respond to clozapine. Clozapine plasma concentrations were not examined in most clozapine studies included in our analysis, but clozapine titration speed was high and similar to that of each comparator used. The mean clozapine dose in studies comparing clozapine with an SGA was 392 mg/d, which was significantly lower than the mean clozapine dose (511 mg/d) in studies using an FHA as a comparator (P = .03). The metaregression analyses using the antipsyc-

<table>
<thead>
<tr>
<th>Clozapine Comparator</th>
<th>No. of Patients</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorpromazine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong et al,37 1997</td>
<td>40</td>
<td>-0.44 (-1.07 to 0.19)</td>
</tr>
<tr>
<td>Honigfeld et al,35 1984</td>
<td>125</td>
<td>-0.64 (-1.00 to -0.28)</td>
</tr>
<tr>
<td>Kane et al,36 1988</td>
<td>263</td>
<td>-0.88 (-1.13 to -0.63)</td>
</tr>
<tr>
<td>Total</td>
<td>430</td>
<td>-0.75 (-0.97 to -0.53)</td>
</tr>
<tr>
<td>τ² = 0.01; χ² = 2.27; P = .32; I² = 12%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Haloperidol          |                |             |
| Buchanan et al,7 1998 | 75            | -0.14 (-0.59 to 0.31) |
| Kane et al,18 2001  | 34             | -0.26 (-0.98 to 0.46) |
| Rosenheck et al,8 1997 | 423         | -0.23 (-0.42 to -0.04) |
| Volavka et al,19 2002 | 77           | 0.13 (-0.31 to 0.57)   |
| Total                | 609            | -0.17 (-0.33 to -0.01) |
| τ² = 0.00; χ² = 2.19; P = .53; I² = 0% |

| Olanzapine           |                |             |
| Bitter et al,60 2004 | 140            | -0.01 (-0.34 to 0.32) |
| Conley et al,61 2003 | 13             | -0.55 (-1.69 to 0.59) |
| Meltzer et al,62 2008 | 40           | 0.03 (-0.59 to 0.65)   |
| Moresco et al,63 2004 | 15            | -0.43 (-1.48 to 0.62)  |
| Naber et al,64 2005  | 108            | 0.08 (-0.30 to 0.46)   |
| Tollefson et al,65 2001 | 176         | 0.14 (-0.16 to 0.44)   |
| Volavka et al,66 2002 | 79            | 0.47 (0.02 to 0.92)    |
| Total                | 571            | 0.10 (-0.07 to 0.27)   |
| τ² = 0.00; χ² = 5.36; P = .50; I² = 0% |

| Risperidone          |                |             |
| Azorin et al,56 2001 | 256            | -0.33 (-0.58 to -0.08) |
| Bondolfi et al,67 1998 | 86         | 0.18 (-0.25 to 0.61)   |
| Breier et al,67 1999  | 29             | -0.44 (-1.18 to 0.30)  |
| McGurk et al,68 2005  | 52             | 0.03 (-0.52 to 0.58)   |
| Volavka et al,69 2000 | 81            | 0.25 (-0.18 to 0.68)   |
| Wahlbeck et al,70 2000 | 19          | 0.63 (-0.29 to 1.55)   |
| Total                | 523            | 0.00 (-0.29 to 0.29)   |
| τ² = 0.07; χ² = 10.92; P = .05; I² = 54% |

| Ziprasidone          |                |             |
| Sacchetti et al,71 2009 | 144        | 0.02 (-0.31 to 0.35)   |
| Total                | 144            | 0.02 (-0.31 to 0.35)   |
| τ² = 0.09; χ² = 67.75; P < .001; I² = 70% |

| Combined             | 2277           | -0.11 (-0.28 to 0.06) |

The size of squares reflects the weight attributed to each study for every separate pairwise meta-analysis (per drug) and not vs all drugs combined. The diamonds illustrate the summary effect sizes (per separate drug and vs all drugs combined). The middle of each diamond sits on the value for the summary effect size, and the width of the diamond depicts the width of the overall CI. The last comparison (combined) presents the summary effect size for the pairwise meta-analysis of clozapine vs all other antipsychotics. Error bars indicate 95% CI; SMD, standardized mean difference.
Antipsychotics in Treatment-Resistant Schizophrenia

Conclusions

At present, insufficient blinded evidence exists on which antipsychotic is more efficacious for patients with treatment-resistant schizophrenia. Clozapine’s superiority over the FGAs has been demonstrated repeatedly, which establishes clozapine as the standard treatment in this specific population, but evidence from blinded RCTs for the comparison of clozapine with other SGAs is lacking. Our analysis suggests that more trials comparing clozapine with other SGAs in patients with more severe illness and using high clozapine doses are warranted. Moreover, the evidence on antipsychotics other than clozapine, haloperidol, olanzapine, and risperidone is scarce, and their results can change if further studies become published.

ARTICLE INFORMATION

Submitted for Publication: July 8, 2015; final revision received October 19, 2015; accepted November 15, 2015.


Author Contributions: Drs Samara and Leucht had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Dold, Leucht.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Samara, Leucht.

Critical revision of the manuscript for important intellectual content: Samara, Dold, Gianatsi, Nikolakopoulou, Helfer, Salanti.

Statistical analysis: Samara, Dold, Gianatsi, Nikolakopoulou, Salanti, Leucht.

Obtained funding: Leucht.

Administrative, technical, or material support: Samara, Dold, Helfer, Leucht.

Study supervision: Samara, Salanti, Leucht.

Conflict of Interest Disclosures: In the past 3 years, Dr Leucht reports receiving honoraria for lectures from Eli Lilly, Lundbeck (Institute), Pfizer, Janssen, BMS, Johnson and Johnson, Otsuka, Roche, Sanofi, ICON, AbbVie, AOP Orphan, and Servier; for consulting/advisory boards from Roche, Janssen, Lundbeck, Eli Lilly, Otsuka, and TEVA; and for the preparation of educational material and publications from Lundbeck, Institute and Roche. Eli Lilly has provided medication for a clinical trial led by Dr Leucht as the principal investigator. No other disclosures were reported.

Funding/Support: This study was supported by grant FKZ 01KG1214 from the German Federal Ministry of Education and Research.

Role of the Funder/Sponsor: The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: Samantha Roberts, Nottinghamshire Healthcare NHS Foundation Trust, conducted the literature search and received compensation for this role. Carlo Altamura, MD, Department of Mental Health, Psychiatry Clinic, University of Milan, Istvan Bitter, PhD, Department of Psychiatry and Psychotherapy, Semmelweis University, Robert W. Buchanan, MD, Maryland Psychiatric Research Center, University of Maryland School of Medicine, Hung-Yu Chan, PhD, Department of General Psychiatry, Taoyuan Hospital, Taoyuan, Department of Psychology, Chung Yuan Christian University, Chung-Li, and Department of General Psychiatry, National Taiwan University Hospital and University of Medicine, Robin Emsley, MD, University of Stellenbosch, Tygerberg, Deanna L. Kelly, PharmD, Maryland Psychiatric Research Center, University of Maryland, Bruce Kinon, MD, US Therapeutic Area Head, Psychiatry, Lundbeck, LLC, Rosa Maria Moresco, MSC, Institute of Molecular Bioimaging and Physiology, National Research Council, Tecnomed Foundation and Department of Health Sciences, University of Milan-Bicocca, and Experimental Imaging Center, lnstitute for Research, Hospitalization and Health Care San Raffaele Scientific Institute, Susan McGurk, PhD, Center of Psychiatric Rehabilitation and Department of Occupational Therapy, Boston University, Dieter Naber, MD, PhD, Department of Psychiatry and Psychotherapy, University of Hamburg, Pierre Baumann, PhD, Département de Psychiatrie, Département de Psychiatrie, Centre Hospitalier Universitaire Vaudois, Robert Rosenheck, MD, Department of Psychiatry, Yale University School of Medicine, Emilio Sacchetti, MD, University of Brescia, School of Medicine, and Department of Mental Health, Spedali Civili Hospital, Ronald E. See, PhD, Department of Neurosciences, Medical University of South Carolina, Robert C. Smith, MD, PhD, Nathan Kline Institute for Psychiatric Research, and Department of Psychiatry, New York University Medical School, Jan Volavka, MD, PhD, Department of Psychiatry, New York University School of Medicine, and William C. Wirshing, MD, UCLA Medical Center, assisted by replying to our requests. None of these contributors received compensation for this role.

REFERENCES


15. Davis JM, Chen N, Glick ID. A meta-analysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry*. 2003;60(6):553-564.


Antipsychotics in Treatment-Resistant Schizophrenia


65. Brunoni AR, Tadini L, Fregni F. Changes in clinical trials methodology over time: a systematic review. JAMA Psychiatry. March 2016 Volume 73, Number 3 209


