A Meta-analysis of the Efficacy of Second-Generation Antipsychotics

John M. Davis, MD; Nancy Chen, MS; Ira D. Glick, MD

Background: Consensus panel recommendations regarding choice of an antipsychotic agent for schizophrenia differ markedly, but most consider second-generation antipsychotics (SGAs) as a homogeneous group. It has been suggested that SGAs seem falsely more efficacious than first-generation antipsychotics (FGAs) as a result of reduced efficacy due to use of a high-dose comparator, haloperidol. We performed (1) a meta-analysis of randomized efficacy trials comparing SGAs and FGAs, (2) comparisons between SGAs, (3) a dose-response analysis of FGAs and SGAs, and (4) an analysis of the effect of efficacy of an overly high dose of an FGA comparator.

Methods: Literature search of clinical trials between January 1953 and May 2002 of patients with schizophrenia from electronic databases, reference lists, posters, the Food and Drug Administration, and other unpublished data. We included 124 randomized controlled trials with efficacy data on 10 SGAs vs FGAs and 18 studies of comparisons between SGAs. Two of us independently extracted the sample sizes, means, and standard deviation of the efficacy data.

Results: Using the Hedges-Olkin algorithm, the effect sizes of clozapine, amisulpride, risperidone, and olanzapine were 0.49, 0.29, 0.25, and 0.21 greater than those of FGAs, with P values of $2 \times 10^{-7}$, $3 \times 10^{-7}$, $2 \times 10^{-12}$, and $3 \times 10^{-9}$, respectively. The remaining 6 SGAs were not significantly different from FGAs, although zotepine was marginally different. No efficacy difference was detected among amisulpride, risperidone, and olanzapine. We found no evidence that the haloperidol dose (or all FGA comparators converted to haloperidol-equivalent doses) affected these results when we examined its effect by drug or in a 2-way analysis of variance model in which SGA effectiveness is entered as a second factor.

Conclusion: Some SGAs are more efficacious than FGAs, and, therefore, SGAs are not a homogeneous group.

Arch Gen Psychiatry. 2003;60:553-564
Cochrane dataset of 33 trials (6464 patients) of SGA vs FGA efficacy.15-19

Additional methods, search strategies, tables, figures, discussions, citations, and sensitivity analyses can be accessed on our Web site (herein referred to as “Web") (http://www.psych.uic.edu/faculty/davis/meta_analysis) (University of Illinois at Chicago, Department of Psychiatry, 2003) or by requesting a copy from the authors. We plan to update our meta-analysis on the Web quarterly.

METHODS

SELECTION AND STUDY CHARACTERISTICS

We selected random-assignment, controlled clinical trials of patients with schizophrenia or schizoaffective disorder with no restriction on publication date, language, or sample size of 10 SGAs ( amisulpride, aripiprazole, clozapine, olanzapine, quetiapine fumarate, remoxipride hydrochloride, risperidone, serontide, ziprasidone hydrochloride, and zotepine) compared with either FGAs or another SGA, and a dose-response comparison of FGAs and SGAs. Previous research30 has established that all FGAs are equally efficacious. We analyzed the medical literature in its original language. We performed a sensitivity analysis and generated funnel plots to assess the possibility of publication bias.

SEARCH STRATEGY

Modeled after the search strategy of Cochrane reviews, we searched the following databases: MEDLINE (January 1, 1966, to May 31, 2002), International Pharmaceutical Abstracts (January 1, 1970, to March 31, 2002), CINAHL (January 1, 1982, to April 30, 2002), PsychINFO (January 1, 1987, to January 31, 2002). We also searched the Cochrane Database of Systematic Reviews (Issue 2, 2002) and reference lists in journal articles. The Quality of Reporting of Meta-Analyses statement21 and the PRISMA statement22 were used to ensure that the literature was systematically obtained.

We conducted extensive sensitivity analyses to determine whether results were altered by excluding certain studies or by meta-regression. We explored the effects of study design, report completeness (qualitatively), peer-reviewed publication vs non-peer-reviewed publication (including data from posters and the FDA Web site), quality of study, global rating vs PANSS/BPRS continuous scales, and exclusion of certain drugs (Web, “Sensitivity Analysis”). To evaluate data extraction, we compared our effect sizes with those calculated from the Cochrane reviews33-37 and Geddes et al1 effect sizes.

VALIDITY ASSESSMENT

We selected random-assignment, controlled clinical trials of patients with schizophrenia or schizoaffective disorder with no restriction on publication date, language, or sample size of 10 SGAs ( amisulpride, aripiprazole, clozapine, olanzapine, quetiapine fumarate, remoxipride hydrochloride, risperidone, serontide, ziprasidone hydrochloride, and zotepine) compared with either FGAs or another SGA, and a dose-response comparison of FGAs and SGAs. Previous research30 has established that all FGAs are equally efficacious. We analyzed the medical literature in its original language. We performed a sensitivity analysis and generated funnel plots to assess the possibility of publication bias.

SEARCH STRATEGY

Modeled after the search strategy of Cochrane reviews, we searched the following databases: MEDLINE (January 1, 1966, to May 31, 2002), International Pharmaceutical Abstracts (January 1, 1970, to March 31, 2002), CINAHL (January 1, 1982, to April 30, 2002), PsychINFO (January 1, 1987, to January 31, 2002). We also searched the Cochrane Database of Systematic Reviews (Issue 2, 2002) and reference lists in journal articles. The Quality of Reporting of Meta-Analyses statement21 and the PRISMA statement22 were used to ensure that the literature was systematically obtained.

We conducted extensive sensitivity analyses to determine whether results were altered by excluding certain studies or by meta-regression. We explored the effects of study design, report completeness (qualitatively), peer-reviewed publication vs non-peer-reviewed publication (including data from posters and the FDA Web site), quality of study, global rating vs PANSS/BPRS continuous scales, and exclusion of certain drugs (Web, “Sensitivity Analysis”). To evaluate data extraction, we compared our effect sizes with those calculated from the Cochrane reviews33-37 and Geddes et al1 effect sizes.

VALIDITY ASSESSMENT

We selected random-assignment, controlled clinical trials of patients with schizophrenia or schizoaffective disorder with no restriction on publication date, language, or sample size of 10 SGAs ( amisulpride, aripiprazole, clozapine, olanzapine, quetiapine fumarate, remoxipride hydrochloride, risperidone, serontide, ziprasidone hydrochloride, and zotepine) compared with either FGAs or another SGA, and a dose-response comparison of FGAs and SGAs. Previous research30 has established that all FGAs are equally efficacious. We analyzed the medical literature in its original language. We performed a sensitivity analysis and generated funnel plots to assess the possibility of publication bias.

SEARCH STRATEGY

Modeled after the search strategy of Cochrane reviews, we searched the following databases: MEDLINE (January 1, 1966, to May 31, 2002), International Pharmaceutical Abstracts (January 1, 1970, to March 31, 2002), CINAHL (January 1, 1982, to April 30, 2002), PsychINFO (January 1, 1987, to January 31, 2002). We also searched the Cochrane Database of Systematic Reviews (Issue 2, 2002) and reference lists in journal articles. The Quality of Reporting of Meta-Analyses statement21 and the PRISMA statement22 were used to ensure that the literature was systematically obtained.

We conducted extensive sensitivity analyses to determine whether results were altered by excluding certain studies or by meta-regression. We explored the effects of study design, report completeness (qualitatively), peer-reviewed publication vs non-peer-reviewed publication (including data from posters and the FDA Web site), quality of study, global rating vs PANSS/BPRS continuous scales, and exclusion of certain drugs (Web, “Sensitivity Analysis”). To evaluate data extraction, we compared our effect sizes with those calculated from the Cochrane reviews33-37 and Geddes et al1 effect sizes.

VALIDITY ASSESSMENT

We selected random-assignment, controlled clinical trials of patients with schizophrenia or schizoaffective disorder with no restriction on publication date, language, or sample size of 10 SGAs ( amisulpride, aripiprazole, clozapine, olanzapine, quetiapine fumarate, remoxipride hydrochloride, risperidone, serontide, ziprasidone hydrochloride, and zotepine) compared with either FGAs or another SGA, and a dose-response comparison of FGAs and SGAs. Previous research30 has established that all FGAs are equally efficacious. We analyzed the medical literature in its original language. We performed a sensitivity analysis and generated funnel plots to assess the possibility of publication bias.
Table 1. Two-Factor Analysis of Variance: Drug (3 Groups) × Haloperidol Dose (2 Groups) on Efficacy of SGAs vs FGAs*

<table>
<thead>
<tr>
<th>Source</th>
<th>Heterogeneity Test</th>
<th>Q</th>
<th>df</th>
<th>P Value</th>
<th>Model</th>
<th>Q</th>
<th>df</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present study</td>
<td>Direct effect of drug group on efficacy</td>
<td>99.2</td>
<td>79</td>
<td>.06</td>
<td>F</td>
<td>75.94</td>
<td>2</td>
<td>$3 \times 10^{-12}$</td>
</tr>
<tr>
<td></td>
<td>Direct effect of haloperidol dose on efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.278</td>
<td>1</td>
<td>.60</td>
</tr>
<tr>
<td></td>
<td>Does haloperidol dose differentially affect efficacy in different drugs (interaction)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.170</td>
<td>2</td>
<td>.12</td>
</tr>
<tr>
<td>Cochrane reviews</td>
<td>Direct effect of drug group on efficacy</td>
<td>32.5</td>
<td>14</td>
<td>.003</td>
<td>R</td>
<td>2.184</td>
<td>2</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>Direct effect of haloperidol dose on efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.019</td>
<td>1</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>Does haloperidol dose differentially affect efficacy in different drugs (interaction)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.136</td>
<td>2</td>
<td>.34</td>
</tr>
<tr>
<td>Geddes et al</td>
<td>Direct effect of drug group on efficacy</td>
<td>21.9</td>
<td>17</td>
<td>.19</td>
<td>F</td>
<td>33.594</td>
<td>2</td>
<td>$5 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Direct effect of haloperidol dose on efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.057</td>
<td>1</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>Does haloperidol dose differentially affect efficacy in different drug (interaction)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.111</td>
<td>2</td>
<td>.13</td>
</tr>
<tr>
<td>Present study (includes nonhaloperidol FGAs)</td>
<td>Direct effect of drug group on efficacy</td>
<td>193.7</td>
<td>114</td>
<td>.000</td>
<td>R</td>
<td>58.104</td>
<td>2</td>
<td>$2 \times 10^{-13}$</td>
</tr>
<tr>
<td></td>
<td>Direct effect of haloperidol-equivalent dose on efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.397</td>
<td>1</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Does haloperidol-equivalent dose differentially affect efficacy in different drugs (interaction)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.943</td>
<td>2</td>
<td>.14</td>
</tr>
</tbody>
</table>

Abbreviations: F, fixed-effects model; FGA, first-generation antipsychotic; R, random-effects model; SGA, second-generation antipsychotic.
*Effect of 3 efficacy groups (1, clozapine; 2, amisulpride, olanzapine, and risperidone; and 3, aripiprazole, quetiapine, remoxipride, sertindole, ziprasidone, and zotepine) and haloperidol comparator dose (≤12 vs >12 mg) on differential efficacy. This analysis tests simultaneously drug group and dose of comparator in the same model. The third line of each triplet indicates whether the dose of haloperidol comparator affects the efficacy of the 3 drug groups. There is no effect of haloperidol comparator in all 3 datasets when drug group is included in the model. Analysis of the dose of all FGA comparators converted to haloperidol-equivalent doses also failed to show that high or low dose affected differential efficacy.

Efficacy, and consequently we tested this with a meta-analysis of a high dose vs medium dose of randomized, double-blind, fixed-dose FGA studies (Figure 1, points C and B, respectively). If true, the higher dose should be less efficacious. We also analyzed randomized, double-blind clozapine dose-response and plasma level studies.

The haloperidol dose, chlorpromazine hydrochloride dose, and other comparator doses (converted to haloperidol equivalents) were investigated as continuous and dichotomous variables (based on the haloperidol cutoff point of ≤12 vs >12 mg/d of Geddes et al) by using MetaWin and Comprehensive Meta-Analysis across all drugs and then for each drug individually (Web, Table 6 and Table 7). Second, meta-analysis based on a 2-factor analysis of variance was conducted to analyze the effect of the dichotomized haloperidol (or all drug) dose for 3 drug groups (Table 1) using the method of Wang and Bushman.

### RESULTS

The effect sizes (95% confidence intervals [CIs]) from our meta-analysis of clozapine, amisulpride, risperidone, and olanzapine, respectively, and each was highly statistically significant—the best evidence of difference (P=10^{-7}–10^{-12}) (Figure 2, Table 2, and Web, Tables 1-3). Clozapine produced a better response than FGAs with effect size d=0.49, whereas amisulpride, risperidone, and olanzapine clustered around 0.25 effect size units (corresponding to 4-6 PANSS points or 3-4 BPRS points). For perspective, based on 7 studies performed contemporaneously with recently released SGAs, the mean haloperidol–placebo effect size was 0.60 (95% CI, 0.44-0.76) (corresponding to 12.6 PANSS points or 7.8 BPRS points; Web, Figure 4). Thus, the effects of amisulpride, risperidone, and olanzapine vs FGAs are somewhat less than half the effect size of FGAs over placebo or clozapine over FGAs. The large risperidone and olanzapine studies found consistent differences vs FGAs. The outliers were small exploratory studies. Examination of funnel plots for publication bias showed no gross asymmetry, except for those of clozapine and risperidone, which indicate that smaller studies reported better efficacy (greater effect sizes) for SGAs (Web, Figure 10). Sensitivity analyses omitting open-label randomized or non-peer-reviewed studies, including studies with low-dose con-

---

©2003 American Medical Association. All rights reserved.

Downloaded From: http://archpsyc.jamanetwork.com/pdfaccess.ashx?url=/data/journals/psych/5178/ on 06/18/2017
dations, and meta-regression with study quality, or PANSS/BPRS vs global rating, study duration, use of global vs continuous measures, etc, as moderator variables showed essentially identical results (Web, “Sensitivity Analysis”).

We rejected the assertion of Geddes et al1 that the SGAs were equivalently efficacious as a homogeneous group because the amount of variance attributable to the different SGAs was large (Q9=58.8; P=10-9 random-effects model). Aripiprazole,112-114 quetiapine,108,115-118 remoxipride,119-135 sertrindole,109-111,136-138 and ziprasidone107,139-142 show similar efficacy to FGAs in the sense that the improvement scores produced by these SGAs were not statistically significantly better than those of FGAs (Table 2). Failure to find a statistically significant difference does not prove that these drugs are equal to FGAs because there is a possibility that further studies could demonstrate this. We place substantial weight on the ziprasidone data from the FDA.107 Because data for 3 of 4 ziprasidone studies (1341 patients) and all 3 aripiprazole studies (560 patients) were poster data, although sufficient data exist to warrant inclusion, definitive judgment regarding differential efficacy must await publication of poster or FDA data. The 12 studies comparing zotepine with FGAs showed no clear evidence of superiority. There is some variability between studies (with most studies clustering at an efficacy similar to that of FGAs, with 2 outliers); thus, conclusions are limited.143-152 Zotepine’s effect size of 0.15 computed using MetaWin (95% CI, −0.01 to 0.30) and MetaView (95% CI, −0.02 to 0.45) just missed significance, whereas the effect size computed using Comprehensive Meta-Analysis was significant (95% CI, 0.01 to 0.28; P=0.03; Web, Tables 1-3). Most of the studies were short-term studies, but data were available on a few long-term studies (n=16), suggesting that long-term studies produce the same differential efficacy (Web, “Reviews of A amisulpride, Risperidone, and Zotepine”).

DOSE-RESPONSE STUDIES

We examined double-blind trials with patients randomly assigned to medium/high or very high doses of FGAs in a reanalysis of the data of Bollini et al153 and also of 24 trials (Web, “Dose-response Analyses and the Pooling of Doses in Fixed-Dose Studies”). Neither the analysis by Bollini et al153 nor our analysis of the average efficacy between the high/very high dose FGA and the medium/high doses was statistically significant. Indeed, the trend was opposite to that postulated by Geddes et al.1

One clozapine dose-response study155 found that 600 mg/d was somewhat superior to 300 mg/d, which in turn was superior to 100 mg/d in a small sample study, and some patients clinically needed 900 mg/d. Plasma level studies of 400 mg of clozapine (and one with a high clozapine dose) showed that patients with higher clozapine plasma levels had an excellent response, whereas those with lower clozapine plasma levels had a poor response, suggesting that many patients require doses greater than 400 mg. When the dose of the poor responders was increased, most patients’ responses increased. The doses of clozapine used in risperidone or olanzapine comparisons were generally 400 mg or less (sometimes much less).138-164

META-ANALYSES OF COMPARISONS OF SGAs

Meta-analyses of olanzapine vs clozapine66,138,159 and risperidone vs clozapine66,160-163,165,166 showed no significant differences (Table 3). Meta-regression of risperidone vs clozapine showed that clozapine dose was a statistically significant moderator variable (P=0.007). Clozapine tended to be more efficacious than risperidone in studies that used a higher dose of clozapine (Web, Figure 1). Our clozapine dose-response study and plasma level studies suggest that overly low clozapine doses were used in most comparisons of SGAs. Consequently, our meta-analysis does not exclude the possibility that adequate doses of clozapine could be superior to other SGAs.

Six olanzapine vs risperidone studies92,167-171 yielded a nonsignificant effect size (effect size d=0.10; 95% CI, −0.06 to 0.26) (Table 3). Two studies172,173 showed amisulpride to be similar to risperidone (effect size d=−0.10; 95% CI, −1.27 to 1.07), and single studies of clozapine vs zotepine,174 olanzapine vs amisulpride,175 olanzapine vs ziprasidone,176 remoxipride vs clozapine,123 and risperi-

### Table 2. Effect Sizes of 10 Second-Generation Antipsychotics Compared With First-Generation Antipsychotics

<table>
<thead>
<tr>
<th>Antipsychotic Agent</th>
<th>Present Study</th>
<th>Cochrane Reviews15–19</th>
<th>Geddes et al1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Studies, No. (n = 124)</td>
<td>Effect Size (95% CI)</td>
<td>Studies, No. (n = 33)</td>
</tr>
<tr>
<td>A amisulpride</td>
<td>12 F</td>
<td>0.286 (0.16 to 0.41)</td>
<td>0 NA</td>
</tr>
<tr>
<td>Aripiprazole</td>
<td>3 F</td>
<td>−0.03 (−0.39 to 0.38)</td>
<td>0 NA</td>
</tr>
<tr>
<td>Clozapine</td>
<td>31 R</td>
<td>0.494 (0.32 to 0.67)</td>
<td>14 0.38 (0.18 to 0.59)</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>14 F</td>
<td>0.211 (0.14 to 0.28)</td>
<td>8 0.27 (0.18 to 0.35)</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>5 F</td>
<td>−0.008 (−0.17 to 0.16)</td>
<td>3 −0.10 (−0.25 to 0.06)</td>
</tr>
<tr>
<td>Remoxipride</td>
<td>17 F</td>
<td>−0.089 (−0.20 to 0.02)</td>
<td>0 NA</td>
</tr>
<tr>
<td>Sertindole hydrochloride</td>
<td>22 F</td>
<td>0.252 (0.18 to 0.33)</td>
<td>4 0.09 (−0.04 to 0.22)</td>
</tr>
<tr>
<td>Sertindole hydrochloride</td>
<td>4 R</td>
<td>0.028 (−0.34 to 0.39)</td>
<td>0 NA</td>
</tr>
<tr>
<td>Ziprasidone</td>
<td>4 F</td>
<td>−0.038 (−0.15 to 0.06)</td>
<td>0 NA</td>
</tr>
<tr>
<td>Zolpine</td>
<td>12 F</td>
<td>0.146 (−0.01 to 0.30)</td>
<td>4 0.40 (0.14 to 0.67)</td>
</tr>
<tr>
<td>Haloperidol vs placebo</td>
<td>7 NA</td>
<td>0.60 (0.44 to 0.76)</td>
<td>0 NA</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; F, fixed-effects model; NA, not applicable; R, random-effects model.
done vs aripiprazole did not display significant differences (Web, Figure 3).

### COMPARISONS BETWEEN META-ANALYSES

#### Reliability of Data Extraction

The correlations between the effect sizes of the Cochrane reviews and that of Geddes et al (r = 0.95) and between each of these and our effect sizes (r = 0.92 and 0.93, respectively) show a high level of agreement in data extraction (Web, “Reliability of Data Extraction”).

#### Consistency of Data Synthesis

We performed meta-analyses using data from the Cochrane reviews and Geddes et al (our data synthesis of their effect size data). Table 2 gives the results of our meta-analysis of all 3 datasets. Our calculation of the overall effect sizes was similar to those of the Cochrane reviews and that of Geddes et al using 5 different software programs (Web, Tables 1-5). Differences in conclusions are not a result of different statistical methods of data synthesis per se, as our results were virtually identical. Geddes et al found that the same 4 SGAs (amisulpride, clozapine, olanzapine, and risperidone) were more efficacious than FGAs. Our P values and CIs are smaller owing to a much larger total sample size.

**Interpretation by Geddes and Colleagues**

Geddes et al arrived at the opposite conclusion by meta-regression; they suggested that this efficacy difference was caused by an overly high dose of the comparator haloperidol, which reduced its efficacy. Using their data, we replicated the results of Geddes et al using our meta-analysis programs (Table 4). Our results show a small effect of comparator (P = .02), but the test for heterogeneity was highly significant (P < .001) (Table 5). Our meta-regression of the discontinuous haloperidol dose did not

**Table 3. Comparisons Between Second-Generation Antipsychotics Using MetaWin**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Studies, No.</th>
<th>Patients, No.</th>
<th>Effect Size</th>
<th>(95% CI)</th>
<th>Q</th>
<th>df</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olanzapine vs clozapine</td>
<td>3</td>
<td>397</td>
<td>0.089</td>
<td>(-0.34 to 0.52)</td>
<td>0.47</td>
<td>2</td>
<td>.79</td>
</tr>
<tr>
<td>Olanzapine vs risperidone</td>
<td>6</td>
<td>1043</td>
<td>0.097</td>
<td>(-0.06 to 0.26)</td>
<td>1.60</td>
<td>5</td>
<td>.90</td>
</tr>
<tr>
<td>Risperidone vs clozapine</td>
<td>7</td>
<td>836</td>
<td>-0.109</td>
<td>(-0.31 to 0.01)</td>
<td>11.31</td>
<td>6</td>
<td>.08</td>
</tr>
<tr>
<td>Risperidone vs amisulpride</td>
<td>2</td>
<td>472</td>
<td>-0.102</td>
<td>(-1.27 to 1.07)</td>
<td>0.56</td>
<td>1</td>
<td>.46</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

**Table 4. Summary Table**

<table>
<thead>
<tr>
<th>Data and Analysis</th>
<th>Our Statistical Method on</th>
<th>Cochrane Reviews</th>
<th>Data of Geddes et al</th>
<th>Present Data</th>
<th>Table or Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of SGAs reviewed</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of studies reviewed</td>
<td>29</td>
<td>30</td>
<td>124</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amisulpride, clozapine, olanzapine, and risperidone are more efficacious than FGAs</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other SGAs are equally as efficacious as FGAs</td>
<td>No†</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of continuous haloperidol dose</td>
<td>Overall</td>
<td>No significant effect</td>
<td>Higher dose, worse outcome</td>
<td>No significant effect</td>
<td></td>
</tr>
<tr>
<td>Effect of discontinuous haloperidol dose</td>
<td>Overall</td>
<td>No significant effect</td>
<td>No significant effect</td>
<td>No significant effect</td>
<td></td>
</tr>
<tr>
<td>Individual drugs</td>
<td>No significant effect</td>
<td>No significant effect</td>
<td>No significant effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does haloperidol dose affect efficacy differently in different drugs (2-factor ANOVA)</td>
<td>No significant effect</td>
<td>No significant effect</td>
<td>No significant interaction effect</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** ANOVA, analysis of variance; FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.


†Zotepine was significantly more efficacious than FGAs in the pooled Cochrane data (P = .003).

‡Although Geddes et al assert that higher doses of haloperidol produce less efficacy, their finding on the dichotomized haloperidol dose (<12 vs >12 mg) seems not to be statistically significant in their Figure 1 as there is considerable overlap between the 2 effect sizes. Our recalculation of the Geddes et al data agree closely with data presented in their Figure 1, but the high- and the low-dose haloperidol groups were not significantly different from each other.

§The differential effect of FGA dose (including nonhaloperidol FGAs using haloperidol-equivalent doses) and SGA group on efficacy was also examined, and no significant interaction effect was observed. There were few studies with chlorpromazine as a comparator; the effect of chlorpromazine dose could not be analyzed because, for many drugs, only low-dose chlorpromazine was used.
find a significant effect of haloperidol dose even with the data of Geddes et al \(^1\) (\(Q_1=2.34; P=.13\); Web, Table 6).

To explain the results of Geddes et al \(^1\), note that clozapine is used in treatment-resistant patients for whom a high dose of haloperidol comparator was often used. Seven of 9 studies of their 2 most effective SGAs (clozapine and amisulpride) used high haloperidol doses, whereas only 1 of 5 studies of quetiapine and sertindole (“similarly” effective SGAs) used high haloperidol doses (Figure 3B). We believe that the superiority of clozapine and some of the other SGAs are an important finding and that the effect of dose of comparator is an artifact because most studies with high comparator doses were clozapine or amisulpride studies. This is a “Which came first, the chicken or the egg?” problem. Geddes et al \(^1\) suggest that the effect of haloperidol dose explains the better effect of clozapine and some SGAs.

In deciding between 2 alternatives, we first tested the effect of haloperidol dose on efficacy for each SGA considered separately. Dose of haloperidol comparator, as a continuous (Table 5) or dichotomous (Web, Tables 6 and 7) variable, did not reliably affect differential efficacy of any SGA using data from Geddes et al \(^1\), Cochrane, or us. We also examined all FGA comparators converted to haloperidol-equivalent doses (Web, Table 9). All \(P\) values were nonsignificant \((P>.05)\). If the identity of the drug is held constant, effect of comparator dose disappears, a finding consistent with our interpretation. The overall effect of continuous dose of haloperidol comparator was not significant using Cochrane data (coefficient for dose effect=0.005; \(P=.63\)) or our data (coefficient for dose effect=0.003; \(P=.59\)). The effect of the dichotomous haloperidol dose also did not significantly affect efficacy (Figure 3 and Figure 4): present study—\(Q_1=2.5; P=.11\), Cochrane—\(Q_1=0.63; P=.43\), Geddes et al \(^1\)—\(Q_1=2.3; P=.13\) (Table 1). When drug group and comparator group are both included in the model, drug is significant and dose of comparator is not, even in the data from Geddes et al \(^1\) Figure 3 depicts the effect sizes for the 3 groups of drugs by high and low haloperidol dose for data from Geddes et al \(^1\) and our data. Figure 4 shows the same for FGAs converted to haloperidol-equivalent doses. The effect sizes are not very different for trials using 12 mg or less of haloperidol vs those using greater than 12 mg of haloperidol or haloperidol equivalents of all SGAs. We replicated our finding in sensitivity analyses when we used a different meta-regression model or when we omitted studies, that is, 3 single-blend studies, various drugs, non-peer-reviewed studies, etc. Sensitivity analyses using meta-regression with outcome variable (Clinical Global Impressions or PANSS/BPRS) and study quality as the moderator variable also showed no difference (Web, “2-Way Meta-Regression” and “Sensitivity Analysis”).

### RELIABILITY OF META-ANALYSIS

We found a robust correlation (approximately 0.93) among the effect sizes found by Cochrane, Geddes et al \(^1\), and us. The agreement on data extraction and the statistical methods (for each drug separately) supports the validity of meta-analysis and is itself an important finding. It is easier to “spin” a narrative review, which can quote select articles to support a position. The Cochrane reviews are particularly thorough, with many methodological safeguards, including evaluation of indirect measures of efficacy, such as dropouts due to failure to respond. Since our present meta-analysis focuses on overall differential efficacy, it supplements but does not sub-

---

**Table 5. Effect of Comparator Dose on Efficacy of SGAs vs FGAs**

<table>
<thead>
<tr>
<th>Source and Drug</th>
<th>Heterogeneity Test</th>
<th>Effect of Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Q_{all} )</td>
<td>(df )</td>
</tr>
<tr>
<td>Present study: all grouped</td>
<td>179.60</td>
<td>84</td>
</tr>
<tr>
<td>Cochrane reviews: all grouped</td>
<td>52.85</td>
<td>19</td>
</tr>
<tr>
<td>Geddes et al: all grouped</td>
<td>59.66</td>
<td>22</td>
</tr>
</tbody>
</table>

**Abbreviations:** F, fixed-effects model; FGA, first-generation antipsychotic; R, random-effects model; SGA, second-generation antipsychotics; tot, total. \(b_1\) is the regression coefficient.
stitute for the rigor of Cochrane reviews, such as, the classic clozapine meta-analysis. One qualification is that almost all studies have been sponsored by the pharmaceutical industry. It is possible that bias from this source (or others) could be present despite randomized double-blind methods (Web, “Potential Sources of Bias in Meta-analysis”). Consequently, trials independent of the pharmaceutical industry are needed (ie, the National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness [CATIE] project).

**Efficacy Differences**

Some SGAs (clozapine, amisulpride, risperidone, and olanzapine) are significantly more efficacious than FGAs, whereas others are not proven to be so. Some SGAs produce a better functional recovery than FGAs and are cost-effective because reduction of other costs (hospitalization, etc) offsets these much greater medication costs. If efficacy differences are a “myth,” it is a myth that reduces costs. Because there are qualitative and quantitative adverse effect and efficacy differences among SGAs, we believe that most guidelines that group SGAs as a homogeneous class are imprecise. Some researchers suggest that the property of blocking serotonin receptors, characteristic of most SGAs, accounts for the improved efficacy. However, many SGAs (ziprasidone, quetiapine, sertindole, etc) seem to have about the same efficacy as FGAs despite being potent serotonin receptor blockers, and amisulpride, although not a serotonin receptor blocker, is more efficacious than FGAs. This questions serotonin receptor blockade as the primary cause of efficacy differences.

Our meta-analyses on the raw data of the registration studies of olanzapine and risperidone revealed that both SGAs were slightly superior to FGAs on positive symptoms but moderately superior on negative symptoms, cognitive symptoms (thought disorder), mood, and impulse control/excitement, improving many symptoms that were untouched by FGAs. So that the disagreement is not merely semantic, those who argue that SGAs are as efficacious as FGAs on positive symptoms while recognizing that SGAs may be more efficacious on negative symptoms, cognition, or mood hold a somewhat similar position as ours.

There is good evidence that negative studies are more likely to go unpublished. One variant of failure to pub-
lish is incomplete publication where only favorable results such as a good effect on negative symptoms are published, but the unfavorable results on total score are omitted. We have made considerable efforts in obtaining complete data (from the Freedom of Information Act, FDA Web site, posters, etc).

TOLERABILITY DIFFERENCES

Geddes et al(1) argue, “when we controlled for the higher than recommended dose of conventional antipsychotics . . . the differences in efficacy and overall tolerability disappear.” We disagree because their tolerability is based on the number of total dropouts. Because the less effective drug has substantially more dropouts due to lack of efficacy, this is a different phenomenon from dropouts due to adverse effects (total number of dropouts confounds 2 issues: adverse effects and efficacy). Furthermore, dropouts from a double-blind study often reflect concern about “unknown” toxicity in experimental drugs (Web, “Significance of Dropout Rate”). There is no one-to-one correspondence between meta-analyses and treatment recommendations. One limitation of meta-analysis is that it cannot balance qualitative differences (apples and oranges) such as between adverse effects. Clinicians need to weigh the medical seriousness and reversibility of rare but serious adverse effects (e.g., agranulocytosis with clozapine and cardiac conduction disturbance changes with sertindole) vs the frequency and seriousness of more common adverse effects (e.g., weight gain and diabetes mellitus found with olanzapine and clozapine, prolactin elevation with risperidone, etc) in the context of long-term use. Rare adverse effects cannot be accurately estimated from trials with small sample sizes. A few fixed-dose studies show that some SGAs (i.e., risperidone and amisulpride) cause dose-related extra-pyramidal symptoms (EPS). Other SGAs cause so few EPS that their incidence fades into that of placebo.

Substantially fewer EPS results in better acceptance and long-term risk-benefit ratios and is clinically more important than the efficacy differences. We do not believe that it is valid to infer efficacy differences between 2 or more SGAs from effect size comparisons between SGAs and FGAs. Head-to-head comparisons are necessary for proof. Nevertheless, if some SGAs were empirically more efficacious than others with equally few EPS, we believe that they should be recommended above other SGAs with just a low EPS advantage. Some SGAs are more efficacious than FGAs because they alleviate a greater variety of symptoms, resulting in more complete rehabilitation. Consequently, at this time efficacy and EPS advantages necessitate the consideration of olanzapine, risperidone, and amisulpride as first-line drugs. For further discussion on this article, please see the Web projects from Eli Lilly & Co, Indianapolis, Ind; Janssen Pharmaceutical Products LP, Titusville, NJ; AstraZeneca Pharmaceuticals LP, Wilmington, Del; Otsuka America Pharma Inc, Rockville, Md; and Pfizer Inc, New York, NY.

We thank Michael E. Bennett, BS, for assistance with references and manuscript preparation.

Corresponding author and reprints: John M. Davis, MD, the Psychiatric Institute (MC 912), University of Illinois at Chicago, 1601 W Taylor, Chicago, IL 60612 (e-mail: Jdavis@psych.uic.edu).

REFERENCES

20. Lan...


per presented at: American Psychiatric Association 2000 Annual Meeting; May 15, 2000; Chicago, Ill.


(Reprinted) Arch Gen Psychiatry/Vol 60, June 2003 www.archgenpsychiatry.com

©2003 American Medical Association. All rights reserved.