npg

Issues that May Determine the Outcome of Antipsychotic Trials: Industry Sponsorship and Extrapyramidal Side Effect

John M Davis*[1,2], Nancy Chen[1] and Ira D Glick[3]

[1]Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, USA; [2]Maryland Psychiatric Research Center, Baltimore, MD, USA; [3]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA

This study presents a meta-analysis of the influence of several potentially biasing factors (eg industry support, extrapyramidal side effects) on efficacy of studies comparing second-generation antipsychotic (SGA) with first-generation antipsychotic (FGA) medications. We used the dataset from our previously published meta-analysis of 124 randomized controlled trials (RCTs) comparing SGAs with FGAs, to evaluate whether certain possible biases could influence the actual outcome on the total score of the Positive and Negative Syndrome Scale (PANSS), Brief Psychiatric Rating Scale (BPRS), and Clinical Global Impressions (CGI) scores. Industry sponsorship was determined by contact with authors or publication statement. We calculated whether (1) industry sponsorship, (2) study quality, (3) extrapyramidal symptoms (EPS) properties, or (4) prophylactic antiparkinsonian medications influenced SGA vs FGA efficacy for each drug and averaged overall by two Hedges and Olkin-based meta-analyses. The analysis found that none of the factors was significantly associated with a particular outcome. While industry-sponsored articles may conclude their medication to be more favorable than that of a competitor in an RCT, we found that the observed efficacy was not influenced by sponsorship. Many attribute the finding that SGAs appears to be more efficacious than FGAs to be a result of EPS-decreasing efficacy (or its measurement). We were unable to confirm that the drug's EPS properties or antiparkinsonian management altered actual efficacy.
Neuropsychopharmacology (2008) 33, 971–975; doi:10.1038/sj.npp.1301493; published online 4 July 2007

Keywords: antipsychotic; meta-analysis; bias; pharmaceutical industry

INTRODUCTION

Currently, the pharmaceutical industry is being closely scrutinized and special attention is directed at potential biases in clinical trials that favor the trial sponsor (Bodenheimer, 2000). Indeed many studies in general medicine find that the sponsor's drug is superior to the comparator drug (Bodenheimer, 2000). Heres *et al* (2006) found that sponsors conclude their second-generation antipsychotics (SGAs) to be superior to their competitors' SGAs in 90% of the studies. In addition, Montgomery has reported similar results of SGAs *vs* first-generation antipsychotics (FGAs; Montgomery *et al*, 2004).

Although randomized controlled trials (RCTs) remain the gold standard in shaping clinical choices and FDA approval, many are concerned that the pharmaceutical industry may be subtly influencing the actual outcome of controlled random-assignment double-blind studies to favor their product.

In this paper, we explore several potential biasing factors and whether they affect the outcome of comparisons between SGAs and FGAs.

First, there are a number of studies that were used by the pharmaceutical industry for registration of a compound as well as studies that are phase IV postmarketing studies to expand knowledge on a compound after approval. We explored whether pharmaceutical industry-sponsored studies report more favorable 'actual' results (referring here to the 'numerical' empirical data rather than the conclusions) than non-industry-funded (ie 'academic') studies in trials of SGAs *vs* FGAs (we only explored who paid for the study itself, not personal support or affiliation of senior author). Second, since most investigators would hypothesize that the SGA is more efficacious than the FGA, this bias might be more apparent in the less well-controlled studies. We also tested whether poorer quality RCTs would be more likely to produce biased results than more carefully controlled studies. Third, many investigators (Geddes *et al*, 2000; Rosenheck, 2005) believe that the better outcome observed with some SGAs is not due to increased efficacy of these SGAs but rather that this is an artifactual result of the influence of extrapyramidal symptoms (EPS) on the ratings of the efficacy of FGAs. Underlying this hypothesis are observations that EPS may be confused with certain symptoms of schizophrenia (eg parkinsonian akinesia confused with schizophrenia apathy) (Davis and Chen, 2001, 2002; Marder *et al*, 1997). In addition, EPS do lead to early dropout and patients who drop out early will have had less time to improve on study drug. A variant of this hypothesis relevant to the same issue was tested by

Rosenheck in an intriguing way. He administered to every patient a clearly adequate dose of a prophylactic anti-parkinsonian drug, thus minimizing EPS (Rosenheck, 2005), and found no difference in efficacy between olanzapine and haloperidol.

To summarize, we re-evaluated the influence of industry sponsorship, study quality, influence of EPS, and anti-parkinsonian medication. Although these are different issues, they all bear on efficacy and are evaluated using the same methodology, even though they test distinctly different hypotheses. As such, the paper is aimed to be useful in interpreting influences on antipsychotic drug efficacy.

METHOD

Source of Data

We used a previously published dataset from our meta-analysis of SGAs vs FGAs of 124 RCTs and the methodology has been previously reported in detail (Davis et al, 2003). Our primary outcome was improvement as measured with the Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987). If this score was not available, we used the Brief Psychiatric Rating Scale (BPRS; Overall and Gorham, 1962), and if neither was available, we used the Clinical Global Impressions (CGI; Guy, 1976). Thus we used the term 'actual' result to denote the numerical difference on measured efficacy to distinguish it from the overall conclusion. We included not only published but also unpublished data obtained from the FDA by the Freedom of Information Act or from the FDA website (www.fda.gov), so as to minimize the frequently encountered 'file drawer problem' (ie the suppression of, or failure to publish, negative data) in our dataset. We did not update our database as most of the new studies focus on comparing SGA vs SGA. We plan a similar analysis of SGA vs SGA in the future.

Evaluation of Three Types of Moderator Variables as Potential Sources of Bias

Industry sponsorship. We dichotomously classified sponsorship as either industry supported or non-industry supported. We classified a study as industry supported if it was substantially industry funded. Government- and/or foundation-funded studies were also classified as non-industry supported and may include some minor personal industry support (ie speakers fees, consultantship, etc) or industry-supplied medication but were clearly not principally funded by industry. We used the author statement of support. However, since sponsorship information was frequently not stated in the published paper, we contacted the author(s) by mail or e-mail; if there was no response, we made repeated efforts (telephone call, contacting their department, or even personal contact at international meetings). Of note, most of the studies that failed to mention sponsorship were industry sponsored. In this manner, we obtained a 98% response regarding the sponsorship. Where there was 'mixed' funding such as speaker fees, it was relatively straightforward to determine where most of the funding (and ideas) originated. Using this method, there were very few situations in which the sponsorship was not clear. For example, Conley (Conley and Mahmoud, 2001) studied

olanzapine, which is manufactured by Eli Lilly and Company, under a NIMH grant. Although Lilly did provide the study medication, this was clearly an academic study with the core faculty of Spring Grove State Hospital provided by the state and the primary funding was public, even though some of the authors received speaker's fees, etc from industry. Sometimes the drug company would provide medication; they rarely provided partial support by funding one aspect. Our contrast is principal support vs minor support. We did not examine whether speaker's fees or consultantship influence results. We did not examine whether the senior author was a drug company employee who affects the results. Consequently our dichotomy provided two levels of influence, not the absence of influence.

Results from the CATIE trial (Lieberman et al, 2005), a large NIMH non-industry-supported trial, were published after we had completed our analysis. We were interested in whether the results of the CATIE trial would alter our findings comparing the influence of industry vs academic sponsorship. Therefore, we performed a second meta-analysis of industry vs non-industry sponsorship, which included the CATIE results. Our meta-analysis was based on improvement in the PANSS/BPRS/CGI. Since some CATIE patients without tardive dyskinesia (TD) were used in the perphenazine arm, and adequate information was available to measure effect size in the patient without TD receiving SGA, we used the CATIE improvement scores 'duration of successful treatment' using the Cox model in Table 2 (p 1217; Lieberman et al, 2005).

Quality of study. We subjectively rated the quality of a study using a global rating on an 8-point ordinal (Likert) scale based on factors such as independent blinding and/or randomization conducted at a different location than site, use of quantitative rating (BPRS/PANSS), reliability training, large sample size, quality of statistical analysis, and systematic presentation of results. Since we only included RCTs, our scale evaluated very good quality studies vs good quality studies.

Extrapyramidal side effects (EPS). We evaluated this potential bias five different ways. First, we constructed a continuous 5-point semiquantitative scale to measure the degree to which FGAs produce EPS, scored from high to low: ie $5 =$ high dose of high potency FGAs (high dose is defined as above 12 mg haloperidol, the cutoff point proposed by Geddes et al (2000)); $4 =$ low dose of high potency FGAs ($\leqslant 12.5$ mg haloperidol); $3 =$ high dose (defined as above 600 mg chlorpromazine equivalents) of low potency FGAs; $2 =$ low dose of low potency FGAs; or $1 =$ receiving thioridazine (which has a particularly low incidence of EPS). Second, we also classified the SGAs by their propensity to cause EPS on a four-point semiquantitative scale: ie $4 =$ amisulpride, risperidone, or remoxipride (which do cause some EPS, although less than FGAs; Marder et al, 1997; Davis et al, 2003; Davis and Chen, 2004); $3 =$ olanzapine, sertindole, ziprasidone, or aripiprazole; $2 =$ quetiapine; or $1 =$ clozapine. The propensity scale was based primarily on the double-blind, randomized registrational studies of FGA or SGA vs placebo. This included those registrational studies that randomly assigned drug to different doses. There are data for almost all the SGAs and the more common FGAs from these large double-bind

studies. There were a few uncommon FGAs used and we used data from the reported studies of those drugs as summarized (on p 96) by Klein and Davis 1969. We used previously described methods for drug and dose equivalence (Davis and Chen, 2004). We used Geddes *et al*'s (2000) division of high- *vs* low-dose haloperidol (12 mg or lower *vs* above 12 mg) adjusting other FGAs by its equivalence. The Cochrane group and ourselves have reviewed this literature. Third, to measure the difference in EPS properties between SGA and FGA, we subtracted the SGA score from the FGA score, which yielded an 8-point scale (-3 to $+4$) in which the highest score would thus be the pairing of a high dose of a high potency FGA *vs* clozapine ($5-1=4$). The lowest score is the pairing of thioridazine *vs* amisulpride, risperidone, or remoxipride ($1-4=-3$). Fourth, since a high dose of FGA (or of haloperidol in those studies using haloperidol as comparator) would surely produce more EPS than a low dose, we evaluated this as well. Fifth, for the antiparkinsonian strategies, we dichotomized this category into 'continuous prophylactic antiparkinsonian drug' *vs* 'PRN administration.' As mentioned, we recognize the issue that some of the factors may be related, rather than independent. For example, it is likely that EPS propensity and use of antiparkinsonian drugs are correlated.

Statistical Methods

Since some SGAs (eg clozapine) may be more efficacious than others, we held constant the SGA and evaluated the effects of the potential source of bias (eg industry *vs* non-industry sponsorship) for each SGA separately, and then averaged these differences. This was performed by using the SGA–FGA effect size difference found in each study adjusted for the overall mean difference of that SGA by subtracting this mean from each study score of that drug. For example, Buchanan *et al* (1998) found that clozapine produces an efficacy of 0.40. This was 0.05 effect size units smaller than the average effect size of clozapine and therefore it was entered as -0.05 in the non-industry-sponsored studies. Kane *et al* (1988) found clozapine to be 0.88 effect size units higher than the average clozapine study, so 0.43 was entered: EPS of typical drug rated 3 in EPS of FGA rating and 1 in EPS of SGA rating and was in the highest quality rating category. The adjusted SGA–FGA effect size difference (the deviance of each study from each drug's mean) was then used in the meta-analysis with the given continuous or dichotomous measures of bias as a moderator variable using the Hedges and Olkin (1985) based models. We used MetaWin for sensitivity analysis (Rosenberg *et al*, 2000) and Comprehensive Meta-Analysis II as the primary method (Borenstein *et al*, 2005).

Sensitivity Analysis

We did extensive sensitivity analysis, analysis with different meta-analytic methods or using dichotomous variables (see web supplement) (Wang and Bushman, 1999).

Interpretation of Meta-Analysis Size Effects

Since it has been suggested that the observed differences in efficacy between drugs might be completely explained by

Table 1 How Potential Biases Influence Efficacy of SGA *vs* FGA

Potential biases	Effect size	χ^2	*p*
Industry *vs* nonindustry sponsorship	-0.04	0.3	NS
Quality of study	0.01	0.7	NS
EPS of FGAs	0.01	0.3	NS
EPS of SGAs	0.00	0.0	NS
EPS difference between SGAs and FGAs	0.02	0.6	NS
FGA dose	0.01	3.1	NS
Haloperidol dose	0.00	0.0	NS
Antiparkinsonian agent: PRN *vs* continous administration	-0.13	2.6	NS

We evaluated whether potential biases could alter the degree to which SGAs were superior to FGAs in efficacy. A positive effect size indicates that the bias distorted effect size such that the sponsor drug is more efficacious. For example, industry sponsorship actually produces a slightly negative effect, in that, on average, industry studies found SGA less efficacious than academic studies by 0.04 effect sizes, or about 0.5 PANSS total points. Similarly, the larger the effect size, the greater the degree of high EPS propensity influences the differential efficacy of SGA>FGA efficacy. The Q statistic, distributed as χ^2 with one degree of freedom, evaluated the influence of the bias. Effect size changed to +0.04 when the CATIE trial was included.

possible biases, it is important to keep in mind how large such a bias would have to be to account for noticeable differences between certain SGAs and their FGA comparator. Roughly speaking, clozapine has been found to be almost 0.5 effect size units better than FGAs; and olanzapine, risperidone, and amisulpride about 0.25 effect size units better than FGAs (Davis *et al*, 2003). For these empirical results to be completely explained by such an artifact, the bias would need to produce approximately a 0.5 effect size unit (which corresponds essentially to a 12-point PANSS total score) difference of SGA *vs* FGA to explain clozapine's greater efficacy (or one-half that to explain the difference of olanzapine or risperidone *vs* an FGA (Davis and Chen, 2002)). For each potential bias (Table 1), we present the results. A positive number indicates bias in the hypothesized direction. Numbers close to zero indicate no bias and a negative number suggests a bias in the opposite direction. If the results fail to show a given bias, the effect size measure would be close to zero. If the bias was near $+0.25$ effect size units (about 6 PANSS points), it would be a reasonable alternate explanation of the observed difference of some SGAs *vs* FGAs.

RESULTS

Industry Sponsorship

About one-fifth of the studies were academically sponsored and almost all of the rest industry sponsored. We found that SGAs were about as efficacious in industry-sponsored studies as found in the non-industry-sponsored studies (Table 1). The difference was (1) close to zero, ie -0.04 or about a 1-point PANSS total score, (2) in the opposite direction of our hypothesis, and (3) was not statistically significant. Since the difference is in the opposite direction, there is no indication of a sponsorship effect. The significance of the difference is evaluated by the meta-analytic Q statistic (distributed as χ^2), which is almost zero

(Q = 0.3, d.f. = 1, *p* = 0.57). Our meta-analysis (Davis *et al*, 2003) found similar results to the CATIE trial for three SGAs, ie we agree that quetiapine and ziprasidone had similar efficacy to FGAs. We also concur with the CATIE (phase I) finding that olanzapine produced a better improvement than FGA on efficacy, with virtually the exact same effect size. In contrast, the CATIE (phase I) found risperidone to be about equally efficacious to FGAs. Overall the results including those of the CATIE study did not alter the comparison of industry *vs* non-industry sponsorship to a statistically significant degree. The overall effect size of sponsorship with the CATIE study included change from −0.04 to +0.04, but we found the sponsorship effect remains close to zero and was not significant (Q = 0.7, d.f. = 1, *p* = 0.42).

Study Quality

The rating of study quality was not significantly correlated with relative SGA efficacy.

EPS Properties of SGA and/or FGA

We tested five variations with the hypothesis that EPS influence efficacy rating: (1) the EPS propensity of the FGA used; (2) the EPS propensity of the SGA used (some do produce some EPS); (3) the difference in the EPS propensity between each SGA *vs* FGA; (4) the dose of the FGA (higher doses of FGAs would be expected to produce more EPS), and the haloperidol dose for those studies using haloperidol as the comparator; and (5) use of PRN *vs* continuous prophylactic antiparkinsonian agents. The results are presented in Table 1. In meta-regression, the slope of continuous variables is a measure of how much one unit of each bias affects the differential outcome of SGA *vs* FGA effect size. For example, a change in the slope of +0.01 units indicates that this bias produces a false increase in efficacy of SGA over FGA of about 0.1 PANSS total score points. We observed no significant effects from any of the EPS-related variables. Indeed the effect sizes were close to zero. We found no significant effect on efficacy of the mode (either prophylactic or as needed (PRN)) of antiparkinsonian agent administration.

DISCUSSION

We find no evidence that industry sponsorship, study quality, or EPS properties actually influenced the findings of the relative efficacy of SGAs *vs* FGAs. This is an important finding because many believe RCT results are tainted due to industry bias. Since RCTs influence both FDA and clinical decisions, had any of the biases altered actual results, this would have been an important finding.

This said, we recognize the limitations of our study. The evidence presented here is restricted to only the biases examined, by our method, in this dataset (SGA *vs* FGA), and to the outcome examined (actual observed numerical change on PANSS/BPRS/CGI improvement score, not the study author's conclusion). The fact that only 20% of trials were not industry sponsored and few used prophylactic antiparkinsonian medication limits statistical power. The evaluation of study quality was not done on a validated scale

and represents no more than our subjective ratings. Since only good quality studies were included, we evaluated only the upper end of the quality dimensions. We can make no inference about the effect of lower quality studies. Nor can we comment on the efficacy results in other Axis I (eg bipolar) disorders or on recent studies of SGAs *vs* FGAs published since our literature search (Davis *et al*, 2003). Most studies of SGAs *vs* FGAs were conducted for registrational purposes; consequently the pressure on industry is to use an adequate dose of the first-generation comparator that will satisfy the regulatory agencies. In contrast, the competitive pressure on the company in phase IV trials is to put a positive face on their drug by finding it superior to another SGA in efficacy, or through counter-marketing, by asserting that the competitor's drug is more dangerous. The sponsor has more to lose with SGA *vs* SGA comparisons. Most likely SGAs cause less EPS than FGA, so there is at least one outcome where an SGA can be superior. Another limitation is that even though great efforts were undertaken to locate studies, we cannot be sure that other unpublished studies do not exist (ie the file drawer problem). Meta-analytic ANOVA or regression techniques have the limitations of any observational studies on attribution of causation. They do not have the protection of randomization. We think the conclusion made in some of the studies did reflect industry bias, but we find the numerical results on efficacy do not reflect such bias.

We found in examination of individual patient data from trials of several SGAs that there is a low but consistent correlation between the ratings of EPS and the ratings of symptoms. We find that this relationship goes both ways, ie negative symptoms can influence the rating of EPS and vice versa (Davis and Chen, 2001, 2002, 2004; Marder *et al*, 1997). Early dropouts for whatever reason can influence improvement because patients who drop out have less time to manifest improvement. We agree with those who raise the question of the potential for this type of bias. Some authors (Geddes *et al*, 2000; Rosenheck, 2005) have suggested that the empirical findings that certain SGAs are more efficacious than FGAs are an artifact of the FGA-associated EPS altering the effective rating. Rosenheck failed to find olanzapine superior to haloperidol using prophy-lactic antiparkinsonian drugs to minimize this artifact. Our evidence for the lack of influence of prophylactic anti-parkinsonian drug is weak since few studies used this mode of administration. The fact that we could find no evidence that these potential biases do alter actual results does not definitively disprove this hypothesis. It only provides evidence against the possibility that they do. Although EPS may be confused with negative symptoms, we find no evidence that the net effect is large enough to completely explain the difference between some SGAs *vs* FGAs.

Since CATIE was a very large NIMH-sponsored study (Lieberman *et al*, 2005), it is obviously pertinent to consider whether the results of the CATIE study would alter our findings. The CATIE study found virtually identical results for all drugs to our meta-analysis (Davis *et al*, 2003), except for risperidone in phase I, in which it was used in an arguably suboptimal dose (Davis and Chen, 2004). Our second meta-analysis with the CATIE data showed that the overall effect of CATIE did not significantly alter the industry *vs* non-industry comparison, which remains close to zero.

Empirically, both Montgomery *et al* (2004) and ourselves find that biases do creep into the conclusions of RCTs of SGA (Heres *et al*, 2006). It is likely that a much greater degree of bias is present in the information the practitioners receive directly from industry (such as pharmaceutical representatives, industry-sponsored papers or symposia, Internet events, or speaker bureaus). Indeed, the widespread publications of massive numbers of industry-sponsored events, usually focusing on the advantages of each sponsor's drug, have created in clinicians a skepticism about the accuracy of industry-sponsored study findings. Our finding is based on a much larger sample of studies than that of Geddes *et al* (2000). We failed to replicate their analysis (Davis *et al*, 2003) using their methods. Here we used a different method, holding the efficacy of each drug constant. Two studies with very similar results can be made to look different by selection of findings favoring the sponsor's drug and by 'spinning' the discussion and conclusions on these findings. Can it be said that even in pure science, authors never focus on what they feel is their salient findings? In all of science, the critical investigators closely scrutinize the data as well as weigh the paper's discussion and conclusions. One paradox of evidence-based medicine is that if a clinician knows what the data are, the clinician is empowered to use clinical intuition. Focus on the data will help separate empirical results from the marketer's spin. It follows that preserving the integrity of the data and protecting the data from bias is important.

We found no evidence of bias in the actual results of RCTs. We thus conclude that we have evidence that RCT data yield accurate results and that therefore the RCTs are a valid guide to treatment decisions. We found no evidence of possible effects of industry sponsorship, study quality, EPS influence, method by which antiparkinsonian drugs are administered, on the measured PANSS/BPRS/CGI results of the relative efficacy of SGAs *vs* FGAs.

ACKNOWLEDGEMENTS

DISCLOSURE/CONFLICT OF INTEREST

John Davis and Nancy Chen have no financial conflict of interest and have not received or anticipate receiving any monetary support from the pharmaceutical industry including compensation for professional services, research support, honorarium, speaker's fees, royalties, stock or stock options, consultantship and reimbursement for travel expenses, over the past 5 years. Ira Glick, over the past 5 years, either has served on advisory boards, received research support, or been on speakers' bureaus for Solvay, Shire, Glaxo, Janssen, Lilly, AstraZeneca, Pfizer, BMS-Otsuka, and the NIMH. Dr Glick also has stock in Johnson & Johnson and Forrest Laboratories.

REFERENCES

Bodenheimer T (2000). Uneasy alliance—clinical investigators and the pharmaceutical industry (see comment). *N Engl J Med* 342: 1539–1544.

Borenstein M, Hedges L, Higgins J, Rothstein H (2005). *Comprehensive Meta-Analysis*, version 2. Biostat: Englewood, NJ.

Buchanan RW, Breier A, Kirkpatrick B, Ball P, Carpenter WT (1998). Positive and negative symptom response to clozapine in schizophrenic patients with and without the deficit syndrome. *Am J Psychiatry* 155: 751–760.

Conley RR, Mahmoud R (2001). A randomized double-blind study of risperidone and olanzapine in the treatment of schizophrenia or schizoaffective disorder. *Am J Psychiatry* 158: 765–774.

Davis JM, Chen N (2001). The effects of olanzapine on the 5 dimensions of schizophrenia derived by factor analysis: combined results of the North American and international trials. *J Clin Psychiatry* 62: 757–771.

Davis JM, Chen N (2002). Clinical profile of an atypical antipsychotic: risperidone. *Schizophr Bull* 28: 43–61.

Davis JM, Chen N (2004). Dose-response and dose equivalence of antipsychotics. *J Clin Psychopharmacol* 24: 192–208.

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

Geddes J, Freemantle N, Harrison P, Bebbington P (2000). The National Schizophrenia Guideline Development Group. Atypical antipsychotics in the treatment of schizophrenia systematic overview and meta-regression analysis. *Br Med J* 321: 1371–1376.

Guy W (1976). Clinical Global Impressions (CGI). *ECDEU Assessment Manual for Psychopharmacology*. In: US Department of Health, Education, and Welfare publication (ADM) 76-338. National Institute of Mental Health: Rockville MD, pp 218–222.

Hedges LV, Olkin I (1985). *Statistical Methods for Meta-Analysis*. Academic Press: Orlando, FL.

Heres S, Davis JM, Maino K, Jetzinger E, Kissling W, Leucht S (2006). Why olanzapine beats risperidone, risperidone beats quetiapine and quetiapine beats olanzapine again: an analysis of head-to-head studies on second-generation antipsychotics. *Am J Psychiatry* 163: 185–194.

Kane J, Honigfeld G, Singer J, Meltzer H (1988). Clozapine for the treatment-resistant schizophrenic: a double-blind comparison with chlorpromazine. The Clozaril Collaborative Study Group. *Arch Gen Psychiatry* 45: 789–796.

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

Klein DF, Davis JM (1969). *Diagnosis and Drug Treatment of Psychiatric Disorders*. Williams & Wilkins: Baltimore, MD.

Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO et al (2005). Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 353: 1209–1223.

Marder SR, Davis JM, Chouinard G (1997). The effects of risperidone on the five dimensions of schizophrenia derived by factor analysis: combined results of the North American trials (erratum appears in *J Clin Psychiatry* 1998;59(4):200). *J Clin Psychiatry* 58: 538–546.

Montgomery JH, Byerly M, Carmody T, Li B, Miller DR, Varghese F et al (2004). An analysis of the effect of funding source in randomized clinical trials of second generation antipsychotics for the treatment of schizophrenia. *Control Clin Trials* 25: 598–612.

Overall J, Gorham D (1962). The Brief Psychiatric Rating Scale. *Psychol Rep* 10: 799–812.

Rosenberg MS, Adams DC, Gurevitch J (2000). *MetaWin: Statistical Software for Meta-Analysis*, Version 2.0 edition. Sinauer Associates: Sunderland, MA.

Rosenheck RA (2005). Open forum: effectiveness versus efficacy of second-generation antipsychotics: haloperidol without anticholinergics as a comparator. *Psychiatr Serv* 56: 85–92.

Wang MC, Bushman BJ (1999). *Integrating Results Through Meta-Analytic Review Using SAS Software*. SAS Institute Inc: Cary, NC.