

Second-generation antipsychotics for schizophrenia: can we resolve the conflict?

S. Leucht^{1*}, W. Kissling¹ and J. M. Davis²

¹ Department of Psychiatry and Psychotherapy, Technische Universität München, Germany

² Department of Psychiatry, University of Illinois at Chicago, USA

The initial enthusiasm about the second-generation or atypical antipsychotic drugs soon changed into criticism and debate, culminating in the controversial CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness), CUtLASS (Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study) and EUFEST (European First-Episode Schizophrenia Trial) effectiveness trials. This review summarizes the results of three recent meta-analyses that compared second-generation antipsychotics (SGAs) with placebo, with conventional antipsychotics, and with SGAs head-to-head. We compare the meta-analyses with previous reviews and put them in the perspective of CATIE, CUtLASS and EUFEST. The data show that the SGAs are not a homogeneous group and that this confusing classification should be abandoned. We find that, overall, the data are consistent but experts interpret the same results differently. The debate seems to be driven more by values than by data; some place an emphasis on cost, others focus on extrapyramidal side-effects (EPS), weight gain, or efficacy. In our opinion, the SGAs are not the breakthrough that industry would like to maintain. They have different properties, so a clinician may individualize a treatment plan to a given patient's problems, a decision that should be shared with the patient. However, these drugs are important contributions to treatment, and most psychiatrists, let alone patients, would probably not want to do without them.

Received 24 November 2008; Revised 4 February 2009; Accepted 4 February 2009; First published online 1 April 2009

Key words: Antipsychotic drugs, schizophrenia, effectiveness studies.

Introduction

In 1988 the proof that clozapine was more efficacious than first-generation antipsychotic drugs (FGAs) for the treatment of schizophrenia led to enormous enthusiasm about the development of other clozapine-like atypical antipsychotics (Kane *et al.* 1988), each new drug being vigorously marketed as an atypical or second-generation antipsychotic (SGA). Since the introduction of risperidone (Mattes, 1997), there has been a heated debate among clinicians and researchers on which drug is best, culminating in the, again vigorously debated, CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness), CUtLASS (Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study) and EUFEST (European First-Episode Schizophrenia Trial) effectiveness studies (Lieberman *et al.* 2005; Jones *et al.* 2006; Lewis *et al.* 2006; Kahn *et al.* 2008). There is indeed extreme controversy, with some experts lauding the new drugs and others saying that, overall, there is no difference. In this context, we have recently published updated meta-analyses on the

following topics: SGAs *versus* placebo, SGAs *versus* typical antipsychotics, and head-to-head comparisons of SGAs (Leucht *et al.* 2008a, 2009a, b), where we focused on adherence to the evidence without opinion. Here we explore reasons for the controversy and integrate the results. Our aim is to compare our numerical results with those of other meta-analyses (Geddes *et al.* 2000; Adams *et al.* 2008), to put them in perspective with the CATIE, CUtLASS and EUFEST studies, and to resolve or at least understand the controversy.

Method

We present a narrative summary of our previous meta-analyses comparing SGAs with placebo (a meta-analysis based on 38 studies with 7323 participants; Leucht *et al.* 2008a), FGAs (a meta-analysis based on 215 studies, 150 of those double-blind with 21 533 participants; Leucht *et al.* 2009a), and SGAs head-to-head (a meta-analysis of 78 studies with 13 558 participants; Leucht *et al.* 2009b).

Results

Tables 1–3 summarize the main findings of our reviews.

* Address for correspondence: S. Leucht, M.D., Department of Psychiatry and Psychotherapy, Technische Universität München, Klinikum rechts der Isar, Ismaningerstr. 22, 81675 München, Germany.

(Email: Stefan.Leucht@Lrz.tu-muenchen.de)

SGAs versus placebo (Table 1)*Efficacy*

All SGAs were more efficacious than placebo based on various outcomes. Nevertheless, two points were striking: first, the effect sizes for overall efficacy were medium sized. Hedges' *g* for mean Positive and Negative Syndrome Scale (PANSS)/Brief Psychiatric Rating Scale (BPRS) reduction ranged from 0.41 (aripiprazole) to 0.59 (risperidone; a single old clozapine study yielded an effect size of 1.64), and was 0.51 across drugs (Cohen's classification of effect sizes is 0.20 for small, 0.50 for medium and 0.80 for large). The number needed to treat (NNT) ranged from 4 to 7, and was 6 overall. It was surprising that haloperidol was efficacious for both negative symptoms and depression.

Side-effects

Overall, there were a few occasions where SGAs were more sedating than placebo, but there were no differences in drop-outs due to adverse events (except sertindole using risk difference). Most SGA ratings of extrapyramidal side-effects (EPS) were similar to placebo, but amisulpride was only studied in low doses (maximum 300 mg/day). Risperidone had a non-significant trend ($p=0.07$) to produce more EPS, and 32% of risperidone patients received anti-parkinsonian medication (compared to 26% of placebo patients), although clearly less so than haloperidol (48% compared to 20% in its placebo groups; Leucht *et al.* 2008a). A meta-analysis on SGAs for bipolar mania also suggested that some SGAs (aripiprazole, risperidone and ziprasidone; $p=0.06$) do induce some EPS (see fig. 3 in Scherk *et al.* 2007), perhaps because manic patients are more sensitive towards EPS. A more plausible explanation is that manic patients are less likely to be on antipsychotic medication at the beginning of the trial than patients with schizophrenia, which can lead to carry-over effects. Similarly, in an aripiprazole study on adolescents with schizophrenia, EPS clearly occurred (see Table 3 in Findling *et al.* 2008).

SGAs versus FGAs (Table 2)*Efficacy*

Four (amisulpride, clozapine, olanzapine, risperidone) out of nine SGAs were more efficacious than FGAs with small to medium effect sizes in terms of the primary outcome overall symptoms (ranging from an effect size of 0.13 for risperidone to 0.52 for clozapine, NNT 6–15). Of note, the four SGAs that were more efficacious overall were also more efficacious for the

specific positive symptoms and negative symptoms whereas the others were not (quetiapine was less effective for positive symptoms). We therefore concluded that negative symptoms cannot be a core component of 'atypicality'. The relatively few data available for depression were slightly different; aripiprazole and quetiapine were more efficacious whereas risperidone was not. Few data on quality of life and relapse were available (Leucht *et al.* 2009a).

Side-effects

All SGAs induced fewer EPS than haloperidol, most even when the latter was used in low doses below 7.5 mg/day. In the relatively few studies that compared SGAs with low-potency antipsychotics, the difference in EPS was less clear. When compared to haloperidol, clozapine, olanzapine, sertindole and zotepine induced the most weight gain whereas quetiapine and risperidone caused intermediate weight gain and amisulpride stimulated yet a smaller amount of weight gain. However, aripiprazole and ziprasidone induced no significant weight gain. Importantly, the low-potency FGAs also induce weight gain. The SGAs also differed in their sedative properties; some were more sedating than haloperidol but some were less sedating than low-potency FGAs.

SGAs versus SGAs head-to-head (Table 3)

Head-to-head comparisons of SGAs are needed to compare the various SGAs because indirect comparisons relative to a common standard can be influenced by many confounders. Overall, a similar efficacy pattern was observed as we found in the meta-analysis with FGA as comparator. Olanzapine was more efficacious than aripiprazole, quetiapine, risperidone and ziprasidone; risperidone was more efficacious than quetiapine and ziprasidone. Amisulpride was not statistically different from olanzapine or risperidone, but was more efficacious than ziprasidone in drop-out due to inefficacy. Thus, the only notable difference was clozapine, which proved to be superior only to zotepine, and to risperidone in drop-out due to inefficacy. Possibly too low a dose of clozapine may explain this surprising finding (see discussion; Leucht *et al.* 2009b). (Our results on comparative side-effects of the various SGAs have not yet been published.)

Other reviews comparing SGAs with FGAs

The overall efficacy results of the reviews by Geddes *et al.* (2000), Davis *et al.* (2003) and the Cochrane Schizophrenia Group (Adams *et al.* 2008) are presented in Fig. 1. Overall, the effect sizes are comparable to those of our updated reviews.

Table 1. Second-generation antipsychotic drugs (SGAs) and haloperidol versus placebo (Leucht et al. 2008a)

Outcome	Amisulpride	Aripiprazole	Clozapine	Haloperidol	Olanzapine	Quetiapine	Risperidone	Sertindole	Ziprasidone	Zotepine
Overall symptoms	N=1, n=241 SMD -0.54 (-0.81 to -0.27)**	N=7, n=1556 SMD -0.41 (-0.51 to -0.31)**	N=1, n=22 SMD -1.64 (-2.61 to -0.68)**	N=11, n=1540 SMD -0.53 (-0.64 to -0.43)**	N=6, n=992 SMD -0.59 (-0.83 to -0.01)**	N=5, n=735 SMD -0.42 (-0.72 to -0.13)**	N=7, n=977 SMD -0.59 (-0.78 to -0.39)**	N=3, n=629 SMD -0.42 (-0.58 to -0.25)**	N=4, n=584 SMD -0.48 (-0.65 to -0.32)**	N=3, n=304 SMD -0.55 (-0.89 to -0.21)**
Negative symptoms	N=5, n=603 SMD -0.65 (-0.73 to -0.39)**	N=7, n=1556 SMD -0.33 (-0.43 to -0.22)**	N=1, n=22 SMD -0.42 (-1.27 to -0.42)	N=11, n=1527 SMD -0.30 (-0.41 to -0.20)**	N=5, n=733 SMD -0.44 (-0.60 to -0.29)**	N=5, n=652 SMD -0.35 (-0.73 to 0.02)	N=6, n=935 SMD -0.42 (-0.59 to -0.26)**	N=3, n=614 SMD -0.28 (-0.44 to -0.11)**	N=4, n=584 SMD -0.42 (-0.58 to -0.25)**	N=3, n=304 SMD -0.26 (-0.50 to -0.02)*
Positive symptoms	N=3, n=402 SMD -0.16 (-0.58 to 0.26)	N=6, n=1489 SMD -0.36 (-0.46 to -0.25)**	N.A.	N=9, n=1367 SMD -0.54 (-0.65 to -0.43)**	N=5, n=668 SMD -0.56 (-0.72 to -0.41)**	N=5, n=735 SMD -0.41 (-0.67 to -0.16)**	N=5, n=828 SMD -0.83 (-1.13 to -0.53)	N=3, n=629 SMD -0.38 (-0.54 to -0.21)**	N=4, n=584 SMD -0.52 (-0.78 to -0.26)**	N=1, n=79 SMD -0.27 (-0.71 to -0.17)
Depression	N=2, n=261 SMD -0.50 (-0.75 to -0.24)**	N.A.	N=1, n=22 SMD -0.39 (-1.24 to 0.45)	N=2, n=299 SMD -0.33 (-0.56 to -0.11)**	N=3, n=479 SMD -0.28 (-0.47 to -0.10)**	N=2, n=333 SMD -0.08 (-0.31 to 0.14)	N=2, n=459 SMD -0.06 (-0.44 to 0.32)	N=1, n=98 SMD -0.30 (-0.69 to 0.10)	N=3, n=404 SMD -0.33 (-0.52 to -0.13)**	N=1, n=79 SMD -0.48 (-0.92 to -0.03)*
Use of anti-parkinson medication	N=3, n=487 RR 0.87 (0.24-3.20)	N=6, n=1310 RR 1.07 (0.81-1.41)	N.A.	N=11, n=1608 RR 2.34 (1.90-2.88)**	N=3, n=481 RR 1.23 (0.52-2.93)	N=3, n=509 RR 0.79 (0.46-1.35)	N=4, n=323 RR 1.24 (0.89-1.71)	N=3, n=661 RR 0.79 (0.51-1.23)	N=4, n=598 RR 1.33 (0.70-2.51)	N=2, n=227 RR 1.49 (0.60-3.72)
Sedation	N.A.	N=4, n=1107 RR 1.38 (0.82-2.34)	N.A.	N=6, n=970 RR 2.28 (1.11-4.67)*	N=3, n=408 RR 1.93 (0.76-4.90)	N=5, n=750 RR 2.02 (1.18-3.47)**	N=4, n=665 RR 1.29 (0.73-2.29)	N=2, n=315 RR 1.23 (0.53-2.87)	N=2, n=291 RR 2.08 (0.62-6.95)	N=3, n=312 RR 4.60 (1.21-17.50)*
Drop-out due to adverse events	N=4, n=514 RR 0.59 (0.20-1.78)	N=7, n=1615 RR 0.73 (0.49-1.08)	N=1, n=24 RR 0.50 (0.04-7.00)	N=10, n=1476 RR 1.31 (0.80-2.12)	N=5, n=914 RR 0.83 (0.18-3.84)	N=4, n=521 RR 1.10 (0.27-4.38)	N=4, n=599 RR 0.86 (0.35-2.13)	N=3, n=661 RR 1.86 (0.96-3.60)	N=4, n=598 RR 2.49 (0.33-18.89)	N=3, n=312 RR 1.88 (0.61-5.85)

N, Number of studies; n, number of participants; SMD, standardized mean difference (effect size) calculated as Hedges' g; RR, relative risk; EPS, extrapyramidal side-effects measured by use of antiparkinson medication at least once; N.A., not available.

Numbers in parentheses are 95% confidence intervals.

Negative SMDs reflect a superiority of SGAs.

* $p < 0.05$, ** $p < 0.01$.

Table 2. Second-generation antipsychotic drugs (SGAs) versus first-generation antipsychotic drugs (FGAs) (Leucht et al. 2009a)

Outcome	Amisulpride	Aripiprazole	Clozapine	Olanzapine	Quetiapine	Risperidone	Sertindole	Ziprasidone	Zotepine
Overall symptoms	<i>N</i> = 13, <i>n</i> = 1017 SMD -0.31 (-0.44 to 0.19)**	<i>N</i> = 5, <i>n</i> = 2049 SMD -0.05 (-0.14 to 0.05)	<i>N</i> = 23, <i>n</i> = 1997 SMD -0.52 (-0.75 to 0.29)**	<i>N</i> = 28, <i>n</i> = 4966 SMD -0.28 (-0.38 to 0.18)**	<i>N</i> = 11, <i>n</i> = 2412 SMD 0.04 (-0.04 to 0.12)	<i>N</i> = 34, <i>n</i> = 4173 SMD -0.13 (-0.22 to 0.05)**	<i>N</i> = 4, <i>n</i> = 1344 SMD 0.02 (-0.13 to 0.16)	<i>N</i> = 5, <i>n</i> = 980 SMD 0.04 (-0.08 to 0.17)	<i>N</i> = 15, <i>n</i> = 1125 SMD -0.10 (-0.27 to 0.06)
Positive symptoms	<i>N</i> = 4, <i>n</i> = 703 SMD -0.22 (-0.37 to 0.06)**	<i>N</i> = 4, <i>n</i> = 1983 SMD 0.03 (-0.06 to 0.12)	<i>N</i> = 10, <i>n</i> = 1080 SMD -0.36 (-0.56 to 0.16)**	<i>N</i> = 24, <i>n</i> = 4189 SMD -0.15 (-0.21 to 0.09)**	<i>N</i> = 9, <i>n</i> = 1742 SMD 0.14 (0.03-0.26)*	<i>N</i> = 28, <i>n</i> = 3286 SMD -0.13 (-0.20 to 0.05)**	<i>N</i> = 3, <i>n</i> = 1145 SMD 0.17 (-0.03 to 0.36)	<i>N</i> = 4, <i>n</i> = 728 SMD 0.03 (-0.20 to 0.26)	<i>N</i> = 2, <i>n</i> = 192 SMD 0.12 (-0.16 to 0.40)
Negative symptoms	<i>N</i> = 10, <i>n</i> = 929 SMD -0.27 (-0.40 to 0.14)**	<i>N</i> = 5, <i>n</i> = 2049 SMD -0.09 (-0.19 to 0.01)	<i>N</i> = 17, <i>n</i> = 1603 SMD -0.27 (-0.42 to 0.13)**	<i>N</i> = 24, <i>n</i> = 4187 SMD -0.32 (-0.47 to 0.16)**	<i>N</i> = 10, <i>n</i> = 1926 SMD -0.00 (-0.09 to 0.09)	<i>N</i> = 30, <i>n</i> = 3455 SMD -0.13 (-0.21 to 0.06)**	<i>N</i> = 4, <i>n</i> = 1198 SMD -0.11 (0.22-0.01)	<i>N</i> = 3, <i>n</i> = 691 SMD -0.09 (0.29-0.11)	<i>N</i> = 5, <i>n</i> = 450 SMD -0.23 (-0.46 to 0.0)*
Depression	<i>N</i> = 9, <i>n</i> = 900 SMD -0.37 (-0.51 to 0.24)**	<i>N</i> = 1, <i>n</i> = 1278 SMD -0.12 (-0.24 to 0.01)*	<i>N</i> = 6, <i>n</i> = 462 SMD -0.51 (-0.87 to 0.14)**	<i>N</i> = 12, <i>n</i> = 2893 SMD -0.27 (-0.35 to 0.19)**	<i>N</i> = 4, <i>n</i> = 442 SMD -0.23 (-0.41 to 0.04)*	<i>N</i> = 11, <i>n</i> = 1611 SMD -0.10 (-0.23 to 0.03)	<i>N</i> = 2, <i>n</i> = 574 SMD -0.04 (-0.22 to 0.14)	<i>N</i> = 3, <i>n</i> = 691 SMD 0.01 (-0.14 to 0.16)	<i>N</i> = 2, <i>n</i> = 134 SMD -0.14 (-0.48 to 0.20)
Quality of Life	<i>N</i> = 1, <i>n</i> = 194 SMD -0.31 (-0.60 to 0.03)*	<i>N</i> = 1, <i>n</i> = 206 SMD 0.06 (-0.22 to 0.33)	<i>N</i> = 1, <i>n</i> = 311 SMD -0.24 (-0.46 to 0.01)*	<i>N</i> = 5, <i>n</i> = 1450 SMD -0.07 (-0.23 to 0.09)	<i>N</i> = 2, <i>n</i> = 166 SMD -0.012 (-0.18 to 0.43)	<i>N</i> = 4, <i>n</i> = 330 SMD -0.16 (-0.23 to 0.20)	<i>N</i> = 1, <i>n</i> = 105 SMD -0.44 (-0.83 to 0.05)*	<i>N</i> = 1, <i>n</i> = 72 SMD 0.03 (-0.43 to 0.49)	<i>N</i> = 1, <i>n</i> = 122 SMD -0.27 (0.63-0.09)
EPS ^a	<i>N</i> = 11, <i>n</i> = 985 RR 0.61 (0.48-0.77)**	<i>N</i> = 5, <i>n</i> = 2094 RR 0.49 (0.36-0.66)*	<i>N</i> = 6, <i>n</i> = 322 RR 0.16 (0.05-0.49)**	<i>N</i> = 14, <i>n</i> = 3758 RR 0.38 (0.30-0.49)**	<i>N</i> = 10, <i>n</i> = 2031 RR 0.56 (0.40-0.79)**	<i>N</i> = 28, <i>n</i> = 3430 RR 0.62 (0.55-0.71)**	<i>N</i> = 4, <i>n</i> = 1472 RR 0.36 (0.29-0.54)**	<i>N</i> = 4, <i>n</i> = 807 RR 0.65 (0.39-1.07)	<i>N</i> = 9, <i>n</i> = 720 RR 0.76 (0.54-1.07)
Sedation ^a	<i>N</i> = 5, <i>n</i> = 622 RR 0.86 (0.41-1.79)	<i>N</i> = 3, <i>n</i> = 1902 RR -0.63 (0.44-0.89)**	<i>N</i> = 15, <i>n</i> = 1583 RR 1.31 (1.10-1.57)**	<i>N</i> = 8, <i>n</i> = 2879 RR 0.93 (0.81-1.06)	<i>N</i> = 9, <i>n</i> = 1834 RR 0.93 (0.49-1.76)	<i>N</i> = 20, <i>n</i> = 2533 RR 0.92 (0.77-1.09)	<i>N</i> = 3, <i>n</i> = 1127 RR 0.77 (0.44-1.34)	<i>N</i> = 2, <i>n</i> = 607 RR 0.99 (0.42-2.32)	<i>N</i> = 7, <i>n</i> = 554 RR 1.39 (1.01-1.93)*
Weight gain (kg) ^a	<i>N</i> = 4, <i>n</i> = 535 0.28 (-0.94 to 1.50)	<i>N</i> = 3, <i>n</i> = 1892 0.19 (-0.06 to 0.97)	<i>N</i> = 6, <i>n</i> = 402 2.08 (0.36-3.81)**	<i>N</i> = 11, <i>n</i> = 3040 3.46 (2.46-4.46)**	<i>N</i> = 5, <i>n</i> = 1171 1.26 (0.65-1.87)**	<i>N</i> = 11, <i>n</i> = 1545 1.72 (1.08-2.36)**	<i>N</i> = 2, <i>n</i> = 779 2.38 (-0.19 to 4.95)	<i>N</i> = 2, <i>n</i> = 608 -0.49 (-1.63 to 0.64)	<i>N</i> = 5, <i>n</i> = 522 2.41 (1.58-3.24)**

N, Number of studies; *n*, number of participants; SMD, standardized mean difference (effect size) calculated as Hedges' *g*; RR, relative risk; EPS, extrapyramidal side-effects measured by use of antiparkinson medication at least once.

Numbers in parentheses are 95% confidence intervals.

Negative SMDs reflect a superiority of SGAs.

^a SGAs compared to any FGA (haloperidol or low-potency), the text sometimes mentions the results compared to haloperidol or low-potency FGAs.

* $p < 0.05$, ** $p < 0.001$.

Table 3. Second-generation antipsychotics (SGAs) head-to-head, primary outcome PANSS/BPRS total score (modified from Leucht et al. 2009b)

	Amisulpride	Aripiprazole	Clozapine	Olanzapine	Quetiapine	Risperidone	Sertindole	Ziprasidone	Zotepine
Aripiprazole									
Clozapine									
Olanzapine	↔ N=4, n=701 SMD -0.07 (-0.28 to 0.14)	OLA ↑ N=2, n=794 SMD -0.22 (-0.36 to 0.08)**	↔ N=10, n=731 SMD -0.11 (-0.25 to 0.04)						
Quetiapine			↔ N=5, n=299 SMD -0.07 (-0.30 to 0.15)	OLA ↑ N=10, n=1449 SMD 0.23 (0.12-34)**					
Risperidone	↔ N=3, n=519 SMD 0.07 (-0.10 to 0.24)	↔ N=2, n=372 SMD -0.07 (-0.28 to 0.14)	↔ N=8, n=609 SMD 0.13 (-0.06 to 0.32)	OLA ↑ N=16, n=2438 SMD 0.11 (0.03-0.19)**	RIS ↑ N=10, n=1978 SMD -0.19 (-0.10 to 0.28)**				
Sertindole						↔ N=2, n=493 SMD 0.09 (-0.36 to 0.54)			
Ziprasidone	↔ N=1, n=122 SMD 0.15 (-0.20 to 0.51)		↔ N=1, n=146 SMD -0.02 (-0.35 to 0.30)	OLA ↑ N=5, n=1542 SMD 0.29 (0.16-0.41)**	↔ N=2, n=710 SMD 0.00 (-0.28 to 0.27)	RIS ↑ N=3, n=1016 SMD 0.21 (0.07-0.34)*			
Zotepine			CLO ↑ N=1, n=59 SMD 0.76 (0.21-0.1.30)**						

PANSS, Positive and Negative Syndrome Scale; BPRS, Brief Psychiatric Rating Scale; N, number of studies; n, number of participants; SMD, standardized mean difference (effect size) calculated as Hedges' g; CLO, clozapine; OLA, olanzapine; RIS, risperidone.

Numbers in parentheses are 95% confidence intervals.
 Negative SMDs reflect a superiority of the SGA listed in the first column.
 Blank fields indicate that no study is available.
 † Statistically significantly superior, ↔ no significant difference between groups.
 * p < 0.05, ** p < 0.001.

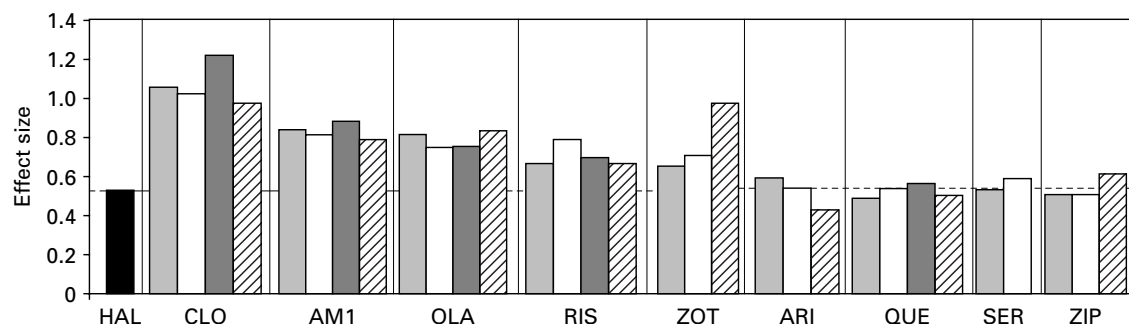


Fig. 1. Comparison of meta-analyses. The mean effect sizes (Hedges' *g*) are presented for the outcome overall efficacy [Positive and Negative Syndrome Scale (PANSS)/Brief Psychiatric Rating Scale (BPRS) total score] of each second-generation antipsychotic drug (SGA) versus conventional antipsychotics in the meta-analyses by Leucht et al. (2009a; ■), Davis et al. (2003; □), Geddes et al. (2000; ■) and the current versions of the Cochrane Reviews (Adams et al. 2008; ▨). For illustration, the effect sizes are added to the effect size found for haloperidol (HAL) versus placebo in Leucht et al. (2008a) (Hedges' *g* = 0.53), which is presented as a benchmark (see dotted line). AMI, amisulpride; ARI, aripiprazole; CLO, clozapine; OLA, olanzapine; QUE, quetiapine; RIS, risperidone; SER, sertindole; ZIP, ziprasidone; ZOT, zotepine.

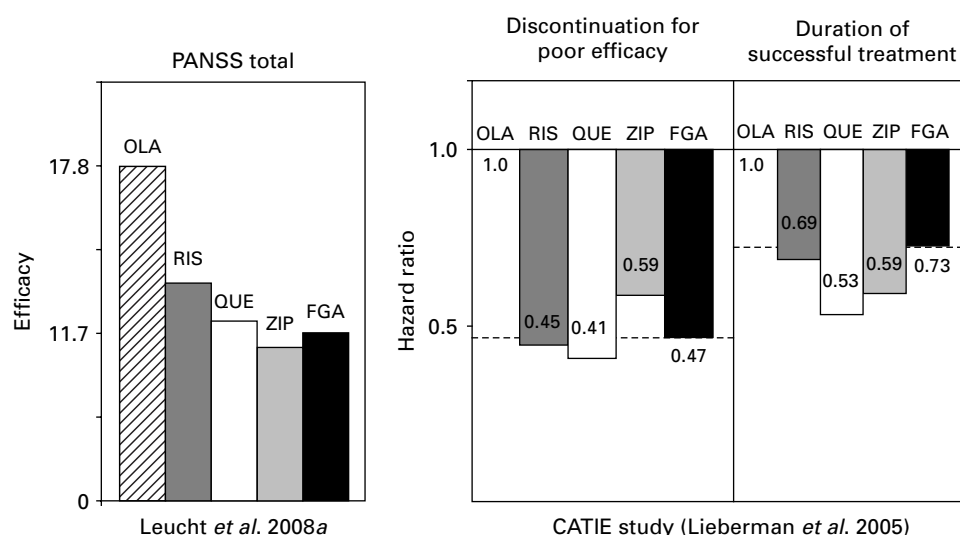


Fig. 2. Comparison with results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). The figure on the left presents the mean reduction on the Positive and Negative Syndrome Scale (PANSS) achieved by olanzapine (OLA), risperidone (RIS), quetiapine (QUE), ziprasidone (ZIP) and first-generation antipsychotic drugs (FGAs) (the antipsychotics also assessed in CATIE) in the meta-analysis by Leucht et al. (2008a). The figures on the right present the hazard ratios for the primary efficacy outcomes in CATIE (Lieberman et al. 2005; discontinuation due to poor efficacy, duration of successful treatment).

Comparison with CATIE

Fig. 2 shows the primary efficacy outcomes reported in the primary publication of CATIE (Lieberman et al. 2005).

Discussion

Are the data consistent?

Fig. 1 shows that the results of our recent meta-analysis are consistent with prior meta-analyses and Cochrane Reviews (Geddes et al. 2000; Davis et al.

2003; Adams et al. 2008). All reviews found that amisulpride, clozapine, olanzapine and risperidone were significantly more efficacious than FGAs whereas for the other SGAs there were no significant differences. The data are consistent with CATIE-I (Lieberman et al. 2005) in that olanzapine was superior in several efficacy outcomes (drop-out due to inefficacy, time on effective treatment, see Fig. 2), and (open-label) clozapine was better than other SGAs in CATIE-II (McEvoy et al. 2006) and in CUTLASS (Lewis et al. 2006). Our meta-analyses found a somewhat better result for risperidone, possibly because in CATIE the

mean modal dose was 3.9 mg/day and only 40% received the 6 mg/day dose, meaning that as many patients as 30–40% received 3 mg or 1.5 mg/day, less efficacious doses according to randomized dose studies (Davis & Chen, 2004) (guidelines suggest 2–8 mg/day; Lehman *et al.* 2004). Clozapine did not turn out to be superior to other SGAs as had been expected, but doses were usually well below 400 mg/day (five were <210 mg/day) and thus clearly lower than in the pivotal studies showing superiority to FGAs (Kane *et al.* 1988; Rosenheck *et al.* 1997; 600 and 523 mg/day respectively) and lower than the optimum doses (Simpson *et al.* 1999; Davis & Chen, 2004). A sufficiently dosed double-blind clozapine *versus* other SGAs trial is still needed.

There is minimal controversy about side-effects and the meta-analytic results are fairly consistent with CATIE.

Are the data flawed?

Several possible methodological flaws have been put forward. Here we discuss industry bias, EPS artifact, statistical methods, blinding and general problems of current clinical trials.

Industry bias

In an analysis of 33 industry-sponsored head-to-head comparisons of SGAs, our blind ratings of abstracts found that 90% favored the sponsor's drug, which provides an answer to our title 'why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine' (Heres *et al.* 2006). However, Davis *et al.* (2008) examined the efficacy effect sizes and found no difference between industry and non-industry sponsored trials, a finding consistent with our two later meta-analyses (SGA *versus* FGA and SGA *versus* SGA; Leucht *et al.* 2009a,b). We conclude that most of the 'industry sponsorship' effect (Heres *et al.* 2006) is due to the spin that the authors put on the data, which contributes to enormous confusion and creates misinformation. That being said, there are of course some studies with obviously flawed designs such as using too low clozapine doses or omitting a result. Failure to report a specific finding can blur interpretation; for example, positive symptoms have never been published in the single comparison of ziprasidone with amisulpride (Olie *et al.* 2006).

High potency FGA comparator artifact in the evaluation of efficacy and side-effects

Could the EPS from high-dose haloperidol or other high-potency FGAs mimic (negative) symptoms and artificially inflate any SGA superiority? Geddes *et al.*

(2000) observed the same effect sizes as we did but interpreted them differently, attributing the observed superiority to occur as an artifact of high haloperidol doses. Other larger meta-analyses could not replicate the dose effect with further studies (Davis *et al.* 2003; Leucht *et al.* 2009a; Table 1 in Davis *et al.* 2008). A Veterans Affairs (VA) collaborative study provided an intriguing investigation of the EPS artifact hypothesis by comparing prophylactic antiparkinsonian medication combined with haloperidol *versus* olanzapine and found little difference (Rosenheck *et al.* 2003). In our dataset there were 11 studies involving only three SGAs on the use of prophylactic antiparkinsonian medication, and these failed to clearly demonstrate this effect for clozapine and olanzapine (Leucht *et al.* 2009a).

In addition, a dose-response analysis by Davis & Chen (2004) found that higher doses of haloperidol were not associated with less efficacy than lower doses. Marder *et al.* 1997; see Table 5) and Davis & Chen (2001, p. 769) examined the original patient data and found no significant correlation between observed EPS and efficacy. We hasten to add that EPS can, and probably does, influence ratings of negative symptoms (or *vice versa*), but firm proof is not available and such an effect would have to be large enough to fully explain the difference.

Nevertheless, haloperidol was used in 95 out of 150 studies; partial justification being that it was the standard antipsychotic at the time of the SGAs' introduction. We find that high-potency FGAs clearly cause more EPS. Low-potency FGAs did not induce more EPS than some, but not all, SGAs, but they lead to weight gain and are sedating (Leucht *et al.* 2009a). It is a major limitation that only a few studies used mid-potency FGA comparators. We recommend that each new drug is compared with a low-potency, a mid-potency, and a high-potency FGA. However, in the absence of such data, 'no evidence of effect does not mean evidence of no effect' (Tarnow-Mordi & Healy, 1999).

Statistical methods

Many have suggested that the efficacy superiority of some SGAs may have been an artifact of the use of last observation carried forward (LOCF) analyses. The argument made concerns the following: when a patient terminates a study prematurely, in LOCF their last observation is used ('carried forward') as their end-point evaluation. If this happens more frequently with haloperidol, SGAs have more time to act on symptoms in an LOCF analysis. We had the opportunity to reanalyze original patient data from pivotal studies comparing amisulpride and olanzapine with

conventional antipsychotics, but did not find a clear LOCF bias (Leucht et al. 2007). Nevertheless, it would be of interest to do such reanalyses on head-to-head comparisons of SGAs. Participants on drugs with side-effects that occur early on, such as EPS (e.g. risperidone) or sedation (e.g. quetiapine), may drop out earlier than those on drugs with side-effects that appear later, such as weight gain (e.g. olanzapine).

Blinding

We found that open studies do have a bias in favor of the SGAs but our meta-analysis was based on just double-blind (Leucht et al. 2009a).

General methodological problems of current antipsychotic drug trials

The populations of patients who enter studies change. For example, risperidone and olanzapine studies usually occurred in subjects who had never received an SGA. Patients who entered later trials may have already been on one or more SGAs. This might make it more difficult to show a difference between SGAs introduced later (e.g. aripiprazole) and FGAs. Nevertheless, we found the same efficacy pattern in head-to-head comparisons of SGAs, where more participants would have already received risperidone/olanzapine.

There are other methodological problems with current antipsychotic drug trials, including: high drop-out rates, too short wash-out phases, chronicity of the participants, selected patient populations, and lack of standardized response criteria (for a review, see Leucht et al. 2008b). These limitations could either inflate or reduce differences, and we could find no evidence that the observed difference among some drugs was a methodological artifact.

Why is there a conflict?

If we accept that, overall, the data are fairly consistent and not entirely biased, we need to discuss the factors producing the controversy.

Promoting all new drugs as one group

The new drugs were all vigorously promoted as being better drugs, similar to clozapine but with pharmacological reasons why they should be better. However, many have never been shown to be more efficacious from the first registrational studies onwards. Such promotion keeps the drugs constantly in the physicians' minds as a class that is essentially different from anything available before. Companies consistently emphasize the disadvantages of their competitors' drugs (sometimes referred to as 'counter-marketing').

However, the most insidious effect of promotion is what is not said. All this promotion contributes to enormous confusion.

Cost

The SGAs cost US\$7.5 billion in the USA in 2003, as much as the cost of all US psychiatrists (Rosenheck, 2005). Without the high cost of the new agents as a key driver, the debate might not even exist. To be cost-effective, higher acquisition costs must be counter-balanced by better efficacy for a long enough time, leading to fewer hospitalizations. As CATIE and CUtLASS did not find important efficacy differences, it was not surprising that their cost-effectiveness analyses found the cheaper drugs to be superior. To find cost-effectiveness, the patients need to be on the initial drug long enough so that fewer hospitalizations can produce the cost offset. Some may suggest that the money spent on the SGAs should rather be spent on psychosocial therapies. Unfortunately, however, the money might be allocated to other areas of medicine such as cardiology or cancer.

How large are the efficacy differences?

Cohen classified effect sizes of 0.20 as small, 0.50 medium, and 0.80 large, but cautioned against over interpretation. The effect size of clozapine versus FGAs was 0.52 (NNT=6). The other significant effect sizes ranged between 0.11 and 0.31, with NNTs between 6 and 15. For perspective, the effect size of haloperidol versus placebo was 0.53 and the NNT equaled 9 (Leucht et al. 2008a). There are many differences between placebo-controlled and active comparator-controlled trials, such as higher drop-out rates in the former (Kemmler et al. 2005), that make it impossible to say that clozapine 'doubles' the efficacy compared with placebo (haloperidol versus placebo 0.53 + clozapine versus FGA 0.52). For another perspective, the NNT to reduce mortality by statins is 88, for aspirin it is 1000, and the effect size of antihypertensive drugs on blood pressure is 0.50. In our opinion, highly disturbed patients are rarely included in trials, and those who are allowed have undergone some partial stabilization. We speculate that this would decrease the effect sizes. Schizophrenia onset is in adolescence/early adulthood and afflicts patients for life, and even a small benefit for a long period of time could be important.

Is the debate more driven by values than by data?

Philosophers of science classify knowledge into three categories: true by definition (logic), empirically true (science), and other (values, ethics, religion). This

distinction is useful here because we feel that many statements on treatment of choice fall in the last category and are value driven; that is, which outcome is more important cannot be proven right or wrong by empirical evidence. For example, those interested in efficacy might value efficacy in particular. Those interested in, for example, weight gain or type 2 diabetes, might emphasize avoiding drugs with these side-effects; those concerned about cost may emphasize the importance of using low-cost drugs; and those who believe that we need better psychosocial treatments will want funds for psychosocial treatments, not expensive medication. The problem is that it is impossible to conclude which value is the most important, as they are qualitatively different. What an expert's value is should not be confused with what the data are.

CATIE, CUtLASS and EUFEST stimulated the conflict

CATIE. We described earlier how many *CATIE* results are compatible with the meta-analyses. *CATIE* is frequently misrepresented as having different results and the assumed reason is 'industry independent *versus* industry sponsored', but we did not find a clear sponsor effect. The misperception was that *all* SGAs are more efficacious, but even the registrational studies failed to find quetiapine and ziprasidone to be more efficacious.

In *CATIE*, there were more discontinuations due to EPS and a greater use of antiparkinsonian agents in the perphenazine group. Nevertheless, the EPS differences were smaller than in our meta-analysis, which mainly used haloperidol, and in *CATIE* there were no differences in rated EPS. It could be that EPS, which are complex phenomena, are not adequately measured in large multi-site trials. For example, the measurement of EPS is altered by variable use of antiparkinsonian medication or different skills of raters. A considerable number of *CATIE* patients were on prophylactic antiparkinsonian medication that was not withdrawn before the study. Perphenazine induces few dystonic reactions and was a wise choice, but we need more studies comparing it with the SGAs to define the exact EPS differences. We speculate that the EPS risk of perphenazine is close to that of risperidone (Hoyberg *et al.* 1993), but higher than that of quetiapine or olanzapine. Finally, there is no perfect study, and *CATIE*'s limitations have been discussed extensively elsewhere (e.g. Kasper & Winkler, 2006), reminding us of an old joke in the theatre community: How many actors does it take to change a light bulb? Twenty-six: one to change it and the other 25 to say they could have done it better. We highlight that the primary analysis based on survival analysis is valid,

but the 74% overall discontinuation rate creates problems to continuous variables. With so many non-random drop-outs, the groups become progressively no longer randomized. Seventy-four per cent of the data must be guessed by statistical modeling. As there were differences between some drugs in drop-outs due to poor efficacy, analysis of observed data at time points late in the study will differ because of a disproportional retention confounding efficacy evaluation. To illustrate this, consider a cancer trial where the drug reduces death over placebo; at the end-point there will still be patients on placebo who are healthy, but the few healthy controls will be as healthy as the many drug-treated patients.

CUtLASS. *CUtLASS* found no difference between SGAs and FGAs apart from the higher SGA cost. This finding was not entirely unexpected because, in *CUtLASS*, clinicians could choose among SGAs and FGAs (the only depots available at that time were FGAs), making conclusions about individual drugs impossible, and we found that not all SGAs are more efficacious. In addition, 60% of the participants of the FGA group were started on sulpiride, a drug chemically similar to amisulpride and therefore probably an old 'atypical'. It seems that you can do reasonably well when you choose carefully among old drugs. We would hasten to add that sulpiride was not well investigated, indeed even the optimum dose range is not known. Neither sulpiride nor perphenazine was the mainstay of treatment (and not even available in some countries). We fear that an interpretation of *CUtLASS* that all drugs are equal would make psychiatrists return to old bad habits, such as high-dose haloperidol.

The design used in many effectiveness studies (*CUtLASS*), many US VA collaborative studies and some *CATIE* analyses keeps patients in the drug group initially randomized to evaluate the patients after switching to other treatments. Although useful in certain situations, this design may blur drug efficacy and side-effects because many of the patients are on a different drug for part of the study. To illustrate this, consider an automobile race from Alaska to the tip of Argentina, where drivers were randomized to a very expensive BMW or the most inexpensive Ford (disclosure: BMW is a major funder of the Technical University of Munich). If the car broke down during the race, the driver could choose either a BMW or a Ford as a replacement car and could even keep the car at the end of the race. Many Fords quickly broke down and the drivers invariably replaced them with BMWs. The drivers assigned randomly to Fords did almost as well as the BMWs, possibly because most were driving BMWs for most of the race.

EUFEST. The *EUFEST* study found a similar rank order in that olanzapine and amisulpride turned out best in discontinuation due to inefficacy, but not identical because haloperidol did worse than ziprasidone. Side-effects were again even more consistent with previous evidence. Our finding that unblinded trials favored the SGAs may in part explain why those SGA equal to haloperidol in blinded studies were intermediate but statistically superior to haloperidol (Leucht *et al.* 2009a).

What are the implications of the data?

SGAs are not a homogeneous category and not a class

As SGAs differ in many properties, including efficacy, side-effects (even in the occurrence of EPS), cost (some are becoming generic), and pharmacology (amisulpride is not a serotonin blocker), they do not form a homogeneous class. Humans like to believe in a consistent body of evidence, avoiding cognitive dissonance. They focus on evidence consistent with their beliefs, and ignore or minimize evidence inconsistent with their beliefs. They therefore distort or simplify information to make it consistent with their dominant cognitions and beliefs. Therefore, because many may have initially thought that clozapine was a more efficacious drug, they concluded that all so-called SGAs were more efficacious. As recent evidence has shown that some new drugs are *not* more efficacious, others may try to make this dissonance consistent by saying that all antipsychotics are the same. As clozapine seems to be effective for negative symptoms, they attribute this to all SGAs, even though most were not shown to be more effective for negative symptoms from the registrational studies onwards. Some confuse empirical data with their conclusions, making a psychoanalytical interpretation as to what the data really show. Most importantly, forcing improper generalizations that all SGAs are the same creates confusion and the classification might have to be abandoned (Leucht *et al.* 2009a).

Implications for choice of drug

Our meta-analyses do not suggest that haloperidol should be used first line, and so does the Cochrane review on haloperidol *versus* placebo (Joy *et al.* 2007). Even in very low doses (2–4 mg) it led to more EPS than several FGAs (Zimbroff *et al.* 1997; Schooler *et al.* 2005; Kahn *et al.* 2008). Low-potency FGAs have a lower EPS risk but they induce weight gain, sedation and other side-effects [hypotension, QT-prolongation, sudden death (thioridazine)]. Some mid-potency antipsychotics may avoid problems of both low- and high-potency antipsychotics, but these drugs have not

been sufficiently studied (dose range, etc.). To avoid the side-effects of the more efficacious drugs clozapine, olanzapine (weight gain, etc.), amisulpride and risperidone (some EPS and more prolactin increase than haloperidol), physicians could start patients on aripiprazole or ziprasidone, which are just as efficacious as haloperidol, yet more tolerable. Quetiapine has very low EPS but intermediate weight gain. Nevertheless, it is useful to have many options available and we feel it is not justified to conclude that FGAs should never be used. We still use them in many circumstances, although usually second line.

There is substantial variability between individual patients in how they respond to antipsychotic drugs. Not all patients will develop weight gain on olanzapine or EPS on haloperidol. Therefore, the drug of choice must be tailored to the individual patient. Furthermore, we strongly recommend shared decision making with the patient (see review by Hamann *et al.* 2003). After all, it is the patients who take the medication.

Conclusion

Marketing by pharmaceutical companies has often promoted SGAs by smoke and mirrors. Many hopes in the SGAs, such as dramatically better efficacy, compliance, quality of life and no side-effects, have not been fulfilled. Overall, the data are consistent, but value judgments and spin have led to different interpretations. Nevertheless, we have no doubt that the introduction of these compounds has contributed to the treatment of schizophrenia and, as an article in 2003 in *The New York Times* said, 'few psychiatrists – and perhaps even fewer patients – would want to lose any of the newer generation of antipsychotics now on the market' (Goode, 2003).

Acknowledgements

Financial support was provided by a grant from the German Federal Ministry of Education and Research (no. FKZ: 01KG 0606, GZ: GFKG01100506) to S.L.; and a grant from the National Institute of Mental Health, Advanced Center for Intervention and Services Research Center (MH-68580), Grant No. 1 P01MH68580-01 CFDA #93.242, the Maryland Psychiatric Research Center to J.D.

Declaration of Interest

S.L. has received speaker and/or consultancy honoraria from SanofiAventis, BMS, EliLilly, Janssen/Johnson and Johnson, Lundbeck and Pfizer; and he has received funding for research projects from

EliLilly and SanofiAventis. W.K. has received speaker and/or advisory board/consultancy honoraria from Janssen, Sanofi-AVENTIS, Johnson and Johnson, Pfizer, Bayer, BMS, Astra Zeneca, Lundbeck, Novartis and EliLilly.

References

- Adams CE, Coutinho E, Davis JM, Duggan L, Li C, Leucht S, Tharyan P, Cochrane Schizophrenia Group** (2008). In *The Cochrane Library*. John Wiley & Sons Ltd: Chichester, UK.
- 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. *Journal of Clinical Psychiatry* **62**, 757–771.
- Davis JM, Chen N** (2004). Dose–response and dose equivalence of antipsychotics. *Journal of Clinical Psychopharmacology* **24**, 192–208.
- Davis JM, Chen N, Glick ID** (2003). A meta-analysis of the efficacy of second-generation antipsychotics. *Archives of General Psychiatry* **60**, 553–564.
- Davis JM, Chen N, Glick ID** (2008). Issues that may determine the outcome of antipsychotic trials: industry sponsorship and extrapyramidal side effect. *Neuropsychopharmacology* **33**, 971–975.
- Findling RA, Robb A, Nyilas M, Forbes RA, Jin N, Ivanova S, Marcus R, McQuade RD, Iwamoto T, Carson WH** (2008). A multiple-center, randomized, double-blind, placebo-controlled study of oral aripiprazole for treatment of adolescents with schizophrenia. *American Journal of Psychiatry* **165**, 1432–1441.
- Geddes J, Freemantle N, Harrison P, Bebbington P** (2000). Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *British Medical Journal* **321**, 1371–1376.
- Goode E** (2003). Leading drugs for psychosis come under new scrutiny. *The New York Times*, 20 May 2003.
- Hamann J, Leucht S, Kissling W** (2003). Shared decision making in psychiatry. *Acta Psychiatrica Scandinavica* **107**, 403–409.
- Heres S, Davis J, Maino K, Jetzinger E, Kissling W, Leucht S** (2006). Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: an exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *American Journal of Psychiatry* **163**, 185–194.
- Hoyberg OJ, Fensbo C, Remvig J, Lingjaerde O, Sloth-Nielsen M, Salvesen I** (1993). Risperidone versus perphenazine in the treatment of chronic schizophrenic patients with acute exacerbations. *Acta Psychiatrica Scandinavica* **88**, 395–402.
- Jones PB, Barnes TRE, Davies L, Dunn G, Lloyd H, Hayhurst KP, Murray RM, Markwick A, Lewis SW** (2006). Randomized controlled trial of the effect on Quality of Life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Archives of General Psychiatry* **63**, 1079–1086.
- Joy CB, Adams CE, Laurie S** (2007). Haloperidol versus placebo for schizophrenia. *Cochrane Database of Systematic Reviews*. Issue No. 4. Article No. CD003082.
- Kahn RS, Fleischhacker WW, Boter H, Davidson M, Vergouwe Y, Keet IPM, Gheorghe MD, Rybakowski JK, Galderisi S, Libiger J, Hummer M, Dollfus S, Lopez-Ibor JJ, Hranov LG, Gaebel W, Peuskens J, Lindfors N, Riecher-Rossler A, Grobbee DE** (2008). Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet* **371**, 1085–1097.
- Kane JM, Honigfeld G, Singer J, Meltzer H** (1988). Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry* **45**, 789–796.
- Kasper S, Winkler D** (2006). Addressing the limitations of the CATIE study. *World Journal of Biological Psychiatry* **7**, 126–127.
- Kemmler G, Hummer M, Widschwendter C, Fleischhacker W** (2005). Dropout rates in placebo-controlled and active-control clinical trials of antipsychotic drugs: a meta-analysis. *Archives of General Psychiatry* **62**, 1305–1312.
- Lehman AF, Lieberman JA, Dixon LB, McGlashan TH, Miller AL, Perkins DO, Kreyenbuhl J, McIntyre JS, Charles SC, Altshuler K, Cook I, Cross CD, Mellman L, Moench LA, Norquist G, Twemlow SW, Woods S, Yager J, Gray SH, Askland K, Pandya R, Prasad K, Johnston R, Ninger J, Peele R, Anzia DJ, Benson RS, Lurie L, Walker RD, Kunkle R, Simpson A, Fochtman LJ, Hart C, Regier D** (2004). Practice guideline for the treatment of patients with schizophrenia, second edition. *American Journal of Psychiatry* **161**, 1–56.
- Leucht S, Arbter D, Engel RR, Kissling W, Davis JM** (2008a). How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Molecular Psychiatry*. Published online: 8 January 2008. doi:10.1038/sj.mp.4002136.
- Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM** (2009a). Second-generation versus first-generation antipsychotics for schizophrenia: a meta-analysis. *Lancet* **373**, 31–41.
- Leucht S, Engel RR, Bauml J, Davis JM** (2007). Is the superior efficacy of new generation antipsychotics an artifact of LOCF? *Schizophrenia Bulletin* **33**, 183–191.
- Leucht S, Heres S, Hamann J, Kane JM** (2008b). Methodological issues in current antipsychotic drug trials. *Schizophrenia Bulletin* **34**, 275–285.
- Leucht S, Komossa K, Rummel-Kluge C, Corves C, Hunger H, Schmid F, Asenjo Lobos C, Schwarz S, Davis JM** (2009b). A meta-analysis of head-to-head comparisons of second-generation antipsychotics in the treatment of schizophrenia. *American Journal of Psychiatry* **166**, 152–163.
- Lewis SW, Barnes TR, Davies L, Murray RM, Dunn G, Hayhurst KP, Markwick A, Lloyd H, Jones PB** (2006). Randomized controlled trial of effect of prescription of clozapine versus other second-generation antipsychotic drugs in resistant schizophrenia. *Schizophrenia Bulletin* **32**, 715–723.

- Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RSE, Davis SM, Davis CE, Lebowitz BD, Severe J, Hsiao JK (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine* **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. *Journal of Clinical Psychiatry* **58**, 538–546.
- Mattes JA (1997). Risperidone: how good is the evidence for efficacy? *Schizophrenia Bulletin* **23**, 155–161.
- McEvoy JP, Lieberman JA, Stroup TS, Davis SM, Meltzer HY, Rosenheck RA, Swartz MS, Perkins DO, Keefe RSE, Davis CE, Severe J, Hsiao JK (2006). Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *American Journal of Psychiatry* **163**, 600–610.
- Olie JP, Spina E, Murray S, Yang R (2006). Ziprasidone and amisulpride effectively treat negative symptoms of schizophrenia: results of a 12-week, double-blind study. *International Clinical Psychopharmacology* **21**, 143–151.
- Rosenheck R, Cramer J, Xu WC, Thomas J, Henderson W, Frisman L, Fye C, Charney D (1997). A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. *New England Journal of Medicine* **337**, 809–815.
- Rosenheck R, Perlick D, Bingham S, Liu-Mares W, Collins J, Warren S, Leslie D, Allan E, Campbell EC, Caroff S, Corwin J, Davis L, Douyon R, Dunn L, Evans D, Frecska E, Grabowski J, Graeber D, Herz L, Kwon K, Lawson W, Mena F, Sheikh J, Smelson D, Smith-Gamble V (2003). Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: a randomized controlled trial. *Journal of the American Medical Association* **290**, 2693–2702.
- Rosenheck RA (2005). Effectiveness versus efficacy of second-generation antipsychotics: haloperidol without anticholinergics as a comparator. *Psychiatric Services* **56**, 85–92.
- Scherk H, Pajonk FG, Leucht S (2007). Second-generation antipsychotics in the treatment of acute mania: a systematic review and meta-analysis of randomized, controlled trials. *Archives of General Psychiatry* **64**, 442–455.
- Schooler N, Rabinowitz J, Davidson M, Emsley R, Harvey PD, Kopala L, McGorry PD, Van Hove I, Eerdeken M, Swyzen W, De Smedt G (2005). Risperidone and haloperidol in first-episode psychosis: a long-term randomized trial. *American Journal of Psychiatry* **162**, 947–953.
- Simpson GM, Josiassen RC, Stanilla JK, De Leon J, Nair C, Abraham G, Odom WA, Turner RM (1999). Double-blind study of clozapine dose response in chronic schizophrenia. *American Journal of Psychiatry* **156**, 1744–1750.
- Tarnow-Mordi WO, Healy MJ (1999). Distinguishing between ‘no evidence of effect’ and ‘evidence of no effect’ in randomised controlled trials and other comparisons. *Archives of Disease in Childhood* **80**, 210–211.
- Zimbroff DL, Kane JM, Tamminga CA, Daniel DG, Mack RJ, Wozniak PJ, Sebree TB, Wallin BA, Kashkin KB, Adan F, Ainslie G, Allan E, Atri P, Baker R, Beitman B, Brown G, Canive J, Carman J, Dott S, Edwards J, Fenton W, Freidli J, Funderburg L, Ereshefsky L, Gladson M, Hamilton J, Haque S, Hartford J, Horne R, Houck C, Jampala C, Labelle A, Larson G, Liesem M, Liskow B, Makela E, Moore N, Morphy M, Posever T, Risch S, Rotrosen J, Sheehan D, Silverstone P, Swann A, Tapp A, Thomas M, Volavka J, Vora S (1997). Controlled, dose response study of sertindole and haloperidol in the treatment of schizophrenia. *American Journal of Psychiatry* **154**, 782–791.