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)

Peter B. Jones, MD, PhD; Thomas R. E. Barnes, MD, DSc; Linda Davies, MSc; Graham Dunn, PhD; Helen Lloyd, BA; Karen P. Hayhurst, MSc; Robin M. Murray, MD, DSc; Alison Markwick, BA; Shôn W. Lewis, MD

Context: Second-generation (atypical) antipsychotics (SGAs) are more expensive than first-generation (typical) antipsychotics (FGAs) but are perceived to be more effective, with fewer adverse effects, and preferable to patients. Most evidence comes from short-term efficacy trials of symptoms.

Objective: To test the hypothesis that in people with schizophrenia requiring a change in treatment, SGAs other than clozapine are associated with improved quality of life across 1 year compared with FGAs.

Design: A noncommercially funded, pragmatic, multisite, randomized controlled trial of antipsychotic drug classes, with blind assessments at 12, 26, and 56 weeks using intention-to-treat analysis.

Setting: Fourteen community psychiatric services in the English National Health Service.

Participants: Two hundred twenty-seven people aged 18 to 65 years with *DSM-IV* schizophrenia and related disorders assessed for medication review because of inadequate response or adverse effects.

Interventions: Randomized prescription of either FGAs or SGAs (other than clozapine), with the choice of individual drug made by the managing psychiatrist.

Main Outcome Measures: Quality of Life Scale scores, symptoms, adverse effects, participant satisfaction, and costs of care.

Results: The primary hypothesis of significant improvement in Quality of Life Scale scores during the year after commencement of SGAs vs FGAs was excluded. Participants in the FGA arm showed a trend toward greater improvements in Quality of Life Scale and symptom scores. Participants reported no clear preference for either drug group; costs were similar.

Conclusions: In people with schizophrenia whose medication is changed for clinical reasons, there is no disadvantage across 1 year in terms of quality of life, symptoms, or associated costs of care in using FGAs rather than nonclozapine SGAs. Neither inadequate power nor patterns of drug discontinuation accounted for the result.

Arch Gen Psychiatry. 2006;63:1079-1087

Author Affiliations: Department of Psychiatry, University of Cambridge and Cambridgeshire and Peterborough Mental Health NHS Trust, Cambridge (Dr Jones); Division of Neurosciences and Mental Health, Imperial College, London (Dr Barnes and Ms Lloyd); Department of Psychiatry, University of Manchester, Manchester (Mss Davies, Hayhurst, and Markwick and Drs Dunn and Lewis); and Institute of Psychiatry, London (Dr Murray), England.

ANTIPSYCHOTIC DRUGS HAVE been the mainstay of schizophrenia treatment for almost 50 years. However, many people with schizophrenia receiving typical or first-generation antipsychotics (FGAs) have had a suboptimal outcome, with symptomatic relapses and disabling adverse effects, particularly sedation and extrapyramidal symptoms (EPSs).[1]

Atypical or second-generation antipsychotics (SGAs) were hailed as a major advance, principally because of their lower liability for EPSs. The first atypical drug, clozapine, is the most efficacious of all antipsychotics but is restricted to treatment-resistant schizophrenia because of adverse effects. Therapeutic differences between the other SGAs and FGAs are less certain. Two systematic reviews[2,3] showed that the 2 groups of drugs are generally equivalent in terms of efficacy against positive symptoms, whereas another study[4] found evidence of superiority for SGAs. Claims of superiority for SGAs in terms of the treatment of negative symptoms, cognitive enhancement, fewer EPSs, and improved subjective experience and tolerability[5] have led to a general shift away from FGAs in the treatment of schizophrenia. Nevertheless, meta-analyses[6,7] have raised questions about the size and significance of these effects. Like FGAs, SGAs (apart from clozapine) are usually grouped as a class in clinical guidelines, despite pharmacologic heterogeneity.[8,9] The SGAs are much more expensive.

We report a pragmatic, open, multicenter, randomized controlled trial of FGAs vs SGAs for schizophrenia, with blind rating of outcomes across 1 year. The trial was funded by the Health Technology Assessment Program of the United Kingdom National Health Service and received no financial support from the pharmaceutical industry. The key question was whether the additional acquisition costs of SGAs over FGAs would be offset by improvements in health-related quality of life or savings in the use of other health and social care services in people with schizophrenia for whom a change in drug treatment was being considered for clinical reasons, most commonly suboptimal efficacy or adverse effects.

The trial concerned the relative clinical effectiveness of the 2 groups of drugs rather than the efficacy of individual drugs. The primary hypothesis was that the use of SGAs would be associated with a clinically significant improvement in quality of life across 1 year compared with the use of FGAs. Secondary questions concerned whether this improvement would be associated with fewer symptoms and adverse effects, improved patient satisfaction, and lower total health care costs.

METHODS

PROTOCOL AND RATIONALE OF TRIAL DESIGN

This pragmatic, multicenter, rater-blinded, randomized controlled trial was designed to test effectiveness in routine clinical practice: (1) trial entry was defined by the psychiatrist deciding to change drug management, (2) broad inclusion criteria reflected normal clinical practice, and (3) there was nonintensive follow-up with 1 primary outcome. The trial included an economic component and was called the Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1).

Participants were randomized to receive either an FGA or an SGA. The FGAs were chlorpromazine hydrochloride, flupenthixol, haloperidol, loxapine, methotrimeprazine, sulpiride, trifluoperazine hydrochloride, zuclopenthixol, and the depot preparations of fluphenazine decanoate, flupentixol decanoate, haloperidol decanoate, pipothiazine palmitate, and zuclopenthixol decanoate. Thioridazine hydrochloride and droperidol were also included initially but were withdrawn from licensed use during the trial. The SGAs were risperidone, olanzapine, amisulpride, zotepine, and quetiapine fumarate (ziprasidone has not been licensed in England). The responsible consultant psychiatrists (specialist physicians in secondary care) chose the individual drug in each class before randomization.

Five medical schools in England were recruited, covering 14 National Health Service Trusts in northwestern England, Nottingham, western London, southeastern London, and Cambridge. The North West Multi-Center Research Ethics Committee (Manchester) granted ethical approval.

The inclusion criteria were *DSM-IV*[10] schizophrenia, schizoaffective disorder, or delusional disorder; age 18 to 65 years; at least 1 month since the first onset of positive psychotic symptoms; and psychiatrist electing to change the current FGA or SGA treatment because of inadequate clinical response or intolerance. The exclusion criteria were substance misuse or a medical disorder considered clinically to be the major cause of positive psychotic symptoms and a history of neuroleptic malignant syndrome.

RANDOMIZATION AND ASSIGNMENT

Randomization to FGAs or SGAs was concealed via a remote telephone service, undertaken after baseline assessment. After stratifying by treatment center, the method of allocation was randomized, permuted blocks within strata. Participants were recruited over 30 months from July 12, 1999, to January 18, 2002.

The psychiatrists initiated the first dose of randomized treatment as soon as possible and were urged to keep patients in their randomized treatment arm for a minimum of 12 weeks, and preferably for 52 weeks. If a treatment change was required, the psychiatrist was instructed to initiate an alternative from the same class. Adjunctive medication was allowed, but antipsychotic polypharmacy was discouraged. Psychiatrists had access to a custom-made best-prescribing handbook.

OUTCOME MEASURES

The primary outcome was the total score on the Quality of Life Scale (QLS),[11] an instrument used widely in psychopharmacologic treatment trials for schizophrenia,[12,13] assessed blindly at baseline and 12, 26, and 52 weeks. Based on a semistructured interview, the QLS has 21 items rated on a 7-point scale from 0 to 6 with descriptive anchors; high scores reflect normal functioning. Probe questions explore items in 4 areas: interpersonal relations (household, friends, acquaintances, social activity, social network, social initiative, withdrawal, and sociosexual behavior), instrumental role (occupational role, work functioning, work level, and work satisfaction) intrapsychic foundations (sense of purpose, motivation, curiosity, anhedonia, aimless inactivity, empathy, and emotional interaction), and commonplace objects and activities. The sum of the mean scores from each area yields a total score. The QLS takes approximately 45 minutes to complete; interrater reliabilities are high, and confirmatory factor analysis has been conducted.[11]

Secondary outcome measures were (1) Positive and Negative Syndrome Scale (PANSS) score,[14] (2) Calgary depression scale score,[15] (3) participant attitudes and adherence ratings using the Drug Attitudes Inventory[16] and a 7-point drug adherence scale,[17] (4) Global Assessment of Functioning scale score,[10] (5) scores on adverse effects scales (Simpson-Angus extrapyramidal adverse effects rating scale[18] to assess parkinsonism, Barnes Akathisia Rating Scale,[19] Abnormal Involuntary Movements Scale for tardive dyskinesia,[20] and Antipsychotic Non-Neurological Side-Effects Rating Scale[21] [a new scale developed to assess the adverse effects of antipsychotic drugs, including nonneurologic adverse effects found with SGAs rather than FGAs]), and (6) participant satisfaction rated at 12 and 52 weeks regarding the new antipsychotic medication, mental health, and adverse effects.

Interrater reliability was assessed using 10 videotaped QLS and PANSS interviews. An initial assessment of interrater reliability (for 9 trained raters) yielded an intraclass correlation coefficient of 0.91 for the QLS total score and 0.75 for the PANSS total score. Further training and assessment yielded interrater reliability of 0.99 for QLS total score and 0.84 for PANSS total score. For the QLS subscales, the intraclass correlation coefficients were 0.98 for interpersonal relations, 0.75 for instrumental role, and 0.99 for intrapsychic foundations. For the PANSS, the intraclass correlation coefficients were 0.94, 0.85, and 0.84 for the positive, negative, and general subscales, respectively. There were weekly discussions of ratings within medical centers, monthly intercenter video conferences, and face-to-face intercenter meetings every 3 months where fidelity was discussed.

MASKING TO ALLOCATION AND COSTS

The following measures were taken to maintain the blinding: isolation of the offices of the clinical assessors from other team members, use of passwords for electronic data, encryption of e-mails for randomization, restriction of discussions about patients within research teams, and the secure storage of all case report forms. Participants were reminded to avoid open discussion of treatment assignment. Follow-up assessments were performed blinded to randomized allocation at 12, 26, and 52 weeks. Telephone interviews were performed on a few occasions. Participants were deemed to be lost to follow-up only after a minimum of 4 failed visits.

We collected cost information about the use of all services, including hospital inpatient and outpatient services, primary and community care services, and prescribed medications. Direct costs were measured as resource use multiplied by unit cost.

STATISTICAL ANALYSIS

We estimated the intention-to-treat (ITT) effect in the primary analyses. Allowance was made for different patterns of loss to follow-up using multiple imputations, assuming the missing data to be ignorable or missing at random.[22] Routine data exploration was performed using SPSS for Windows 10 (SPSS Inc, Chicago, Ill). Further analysis was performed using Stata Version 7 (StataCorp, College Station, Tex).

Longitudinal analysis of covariance (ANCOVA) was used to estimate the differences between the treatment arms in QLS total scores at each of the 3 assessments (12, 26, and 52 weeks), using study center and baseline QLS score as covariates. Unstructured correlations between repeated measures were assumed. Treatment arm differences for nonlongitudinal, secondary, binary outcome measures were evaluated using Pearson χ^2. Treatment arm differences in ordinal outcomes (eg, patient satisfaction) were evaluated using the Mann-Whitney test.

For the primary analysis, we analyzed QLS scores using the longitudinal ANCOVA first in an analysis on available data, without attempting to impute missing information, and second after imputation of the missing data. Multiple imputations for this second model of QLS involved the generation of 5 full data sets using the propensity score method in Solas version 3.2, each of which was then analyzed as described previously herein (combining the results as suggested by Rubin and Schenker[23]). Separate multiple imputations were performed for each arm, and the complete data from the 2 arms of the trial were then combined to continue the analysis. Variables used to impute missing values included nonmissing QLS and PANSS total scores, study center, reason for referral to study (poor clinical response or intolerance and adverse effects), and whether first episode, current alcohol misuse, and current drug misuse.

A secondary, exploratory analysis of 12-week QLS scores was undertaken to investigate the effect of switching between arms during that initial phase. First, a conventional ITT analysis was performed using ANCOVA as previously described but restricting the outcome to 12-week scores. Then, in a per-protocol analysis, participants who switched from their allocated arm before the 12-week follow-up were dropped and the ANCOVA was repeated.

SAMPLE SIZE AND POWER

The principal outcome, QLS total score, was used to determine sample size. Two assumptions were made a priori: first, that there would be a correlation of 0.5 between baseline and 52-week QLS scores and, second, that a clinically meaningful difference in QLS scores between the 2 arms would be 5 points from baseline to 12 months (difference in 12-month means of 40 vs 45). This was predicated on a common standard deviation of 18 for baseline and 52-week QLS scores and underpinned the primary hypothesis of an advantage for SGAs. A posteriori, the correlation between baseline and 12-month total scores was found to be higher than assumed (0.75 rather than 0.5). The within-group standard deviations were as expected. The higher baseline to 12-month correlation implied that the within-group standard deviation for the change score was approximately 13. Thus, using 80% power, 95% confidence, and 2-tailed assumptions, the target sample size for detecting a difference of 5 points was 110 patients in each of the 2 arms, requiring a total of 254 participants to account for the projected follow-up rate of 75%.

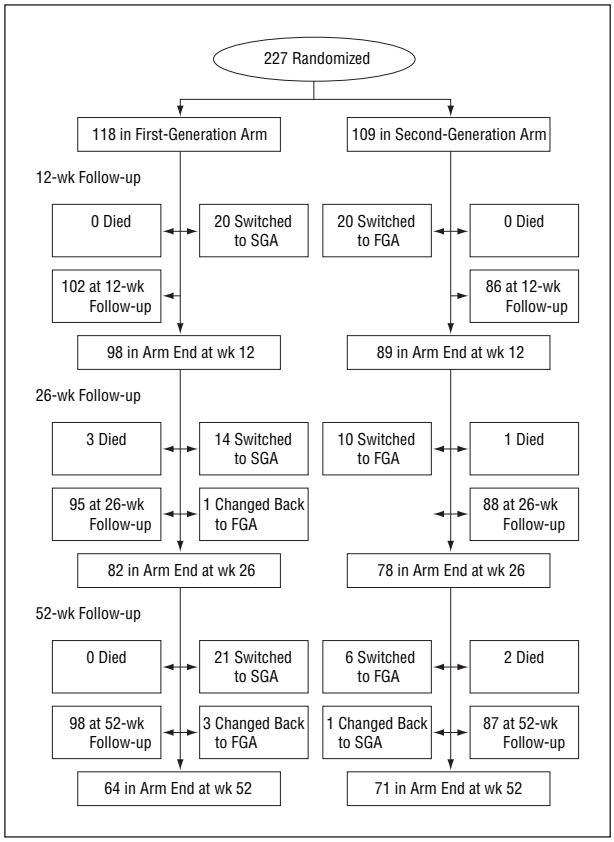

Figure 1. Progress of the randomized patients through the study. FGA indicates first-generation antipsychotic; SGA, second-generation antipsychotic.

RESULTS

PARTICIPANT FLOW

Two hundred seventy-five patients were referred. Of these, 9 (3%) were ineligible, 1 (0.4%) was unable to give consent, and 36 (13%) refused to give consent; 2 psychiatrists each withdrew a referral (1%). Thus, 227 patients, referred by 73 psychiatrists, were randomized. **Figure 1** shows the patients' subsequent progress through the trial. One protocol violation, a patient randomized before the referring psychiatrist reformulated the diagnosis, was included in the final analysis in his or her randomized treatment arm.

Table 1. Baseline Characteristics of the 227 Study Participants by Randomized Treatment Arm

	FGA Arm (n = 118)	SGA Arm (n = 109)
Age, y		
Mean (SD)	40.5 (11.3)	40.9 (11.1)
Median (range)	40.5 (18-63)	41.6 (19-62)
Length of illness, y		
Mean (SD)	13.3 (10.8)	14.4 (11.2)
Median (range)	11.9 (0-42)	11.6 (0-39)
No. of previous hospital admissions		
Mean (SD)	3.4 (4.7)	3.9 (3.9)
Median (range)	2 (0-40)	3 (0-20)
Sex, No. (%)		
M	81 (69)	73 (67)
F	37 (31)	36 (33)
Ethnicity, No. (%)		
White	87 (74)	83 (76)
Black and other minority groups	25 (21)	25 (23)
Other	6 (5)	1 (1)
Diagnosis, No. (%)		
Schizophrenia	85 (72)	85 (78)
Schizophreniform	5 (4)	3 (3)
Schizoaffective disorder	22 (19)	17 (16)
Delusional disorder	6 (5)	4 (4)
Patient status at baseline, No. (%)		
Inpatient	48 (41)	43 (39)
Day patient	1 (1)	5 (5)
Outpatient	69 (58)	61 (56)
First episode, No. (%)	15 (13)	11 (10)
No current drug misuse, No. (%)	86 (73)	84 (77)
No current alcohol misuse, No. (%)	74 (63)	65 (60)
Reason for referral, No. (%)		
Inadequate response	52 (44)	59 (54)
Adverse effects	35 (30)	13 (12)
Both	31 (26)	37 (34)
Antipsychotic drug treatments before randomization		
No. of FGAs	108	99
No. of depots	47 (44)†	37 (37)*
No. of SGAs	25	19
None	2 (<1)	2 (<1)
Antipsychotic polypharmacy, No. (%)	13 (11)	15 (14)

Abbreviations: FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.
*Percentage of total FGAs prescribed that were depot preparations.

Table 2. Drugs Prescribed by Treatment Arm: Numbers at Randomization and at 52 Weeks and End-Study Doses

Drug	Patients, No.		End-Study Dose, Mean (Range), mg
	At Randomization	At 52 wk*	
FGA arm (n = 118)			
Chlorpromazine	8	4	250 (200-300)
Droperidol	1	0	0
Flupentixol	1	3	4 (2-6)
Flupentixol decanoate	2	8	142 fortnightly (40 mo to 250 wk)
Fluphenazine decanoate	3	2	50 fortnightly
Haloperidol	8	2	22.5 (20-25)
Haloperidol decanoate	2	0	0
Loxapine	3	0	0
Methotrimeprazine	0	1	250
Pipotiazine palmitate	2	1	50 fortnightly
Sulpiride	58	31	813 (200-2400)
Thioridazine	1	0	0
Trifluoperazine hydrochloride	21	12	15 (6-30)
Zuclopenthixol	5	2	37 (20-50)
Zuclopenthixol decanoate	3	8	358 fortnightly (150-750 fortnightly)
SGA arm (n = 109)†			
Amisulpride	13	10	610 (200-1200)
Olanzapine	50	37	15 (5-30)‡
Quetiapine	23	11	450 (200-750)
Risperidone	22	13	5 (2-10)

Abbreviations: FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.
*Some patients were taking more than 1 antipsychotic drug concurrently.
†One missing data point at randomization.
‡The United Kingdom–licensed maximum olanzapine dose is 20 mg, although 30 mg is not uncommon in practice.

Of the 227 patients, 118 (52%) were randomized to receive an FGA and 109 (48%) to receive an SGA. These 2 groups were similar at baseline in terms of demographic and clinical characteristics (**Table 1**). Before randomization, FGAs were being prescribed to 207 patients and SGAs to 44. Eighty-four patients (37%) were taking depot FGAs before randomization; 47 (56%) were subsequently randomized to receive FGAs and 37 (44%) to receive SGAs. One patient was prescribed clozapine immediately before randomization. Twenty-eight patients (12%) were receiving more than 1 antipsychotic drug before randomization; 13 (11%) of these were randomized to the FGA arm and 15 (14%) to the SGA arm (Table 1).

Table 2 displays the drugs prescribed in each treatment arm after randomization and those used at 52 weeks together with the mean doses. The average period from randomization to initiation of the assigned drug was 8.5 days (median, 1 day).

FOLLOW-UP

We interviewed 185 patients (81%) at 1 year: 100 randomized to the FGA arm and 85 to the SGA arm (85% vs 78%; $P = .2$). There were 3 deaths in each arm. In the FGA arm, 2 deaths were due to cardiac failure and 1 was considered to be suicide or accidental death (open verdict). In the SGA arm, 2 deaths were also due to cardiac failure and 1 to septicemia (in a quadriplegic patient). Eleven patients (5%) were categorized as lost to follow-up at 1 year, and 22 (10%) withdrew from the study. Including deaths, withdrawals, and lost to follow-ups, 39 patients (17%) dropped out of the trial.

Table 3 gives the QLS data at each assessment point. **Table 4** presents the primary ITT analysis, including imputed values for missing observations, and the secondary per-protocol explorations. For the primary analysis, Table 4 shows parameter estimates for the effect of treatment arm (randomization) common to all 3 outcome times (12, 26, and 52 weeks). A negative parameter estimate means that patients in the

Table 3. Primary Outcome: QLS Score*

Assessment Point	FGA Arm		SGA Arm	
	Patients, No.	QLS Total Score, Mean (SD)	Patients, No.	QLS Total Score, Mean (SD)
Baseline	118	43.3 (21.7)	108	43.5 (20.3)
12 wk	100	49.2 (19.9)	87	46.6 (19.0)
26 wk	93	49.2 (20.5)	87	50.4 (18.8)
52 wk	100	53.2 (21.2)	85	51.3 (19.6)

Abbreviations: FGA, first-generation antipsychotic; QLS, Quality of Life Scale; SGA, second-generation antipsychotic.
*Values for occasional missing items were imputed using the median of observed responses within other subscales for that patient. Higher scores mean higher quality of life.

Table 4. Primary ITT and Secondary Per-Protocol Analyses of QLS Scores

Data	Estimate*	SE (95% CI)	*P* Value
Primary ITT analysis across 3 points to 52 wk			
QLS total scores*	−1.7	1.4 (−4.5 to 1.1)	.24
QLS total scores after multiple imputations of missing data	−2.5	1.9 (−6.2 to 1.2)	
Secondary analyses of 12-wk QLS total scores			
ITT analysis	−2.0	1.7 (−5.3 to 1.4)	.31
Per-protocol analysis, excluding data from patients who had switched within the first 3 mo of the study	−3.0	1.9 (−6.7 to 0.6)	

Abbreviations: CI, confidence interval; ITT, intention-to-treat; QLS, Quality of Life Scale.
*A negative parameter estimate means that participants in the first-generation antipsychotic arm are doing better. The hypothesis was plus 5 in favor of second-generation antipsychotics.

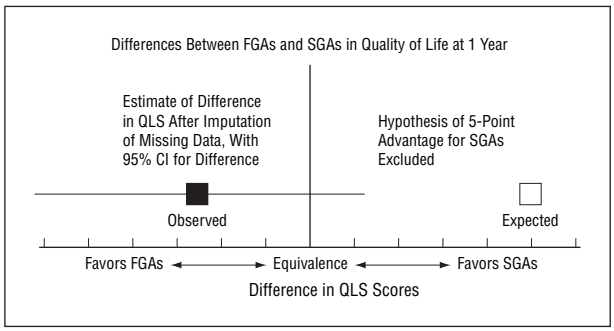

Figure 2. Differences in Quality of Life Scale (QLS) scores at 1 year between patients taking first-generation antipsychotics (FGAs) and second-generation antipsychotics (SGAs). CI indicates confidence interval.

FGA arm were doing better (see the observed means in Table 3).

Contrary to the primary hypothesis, the estimate of 5 points in favor of the SGA arm was excluded at the 95% confidence level. The apparent advantage for FGAs, an effect opposite to the hypothesis, did not reach statistical significance ($P=.24$). These effects, together with our primary hypothesis, are summarized graphically in **Figure 2**. The secondary, per-protocol analysis of 12-week outcomes gave a similar estimate (Table 4).

Table 5 gives the results of the secondary outcomes. The PANSS total scores include imputed values obtained by multiple imputations; all available data were used for the other outcomes. There was a trend for the mean (SD) costs for the 52 weeks of the trial to be lower for people allocated to the FGA arm ($34 750 [$48 100] or £18 800 [£26 000]) than the SGA arm ($37 185 [$46 250] or £20 100 [£25 000]). The major cost in both groups was psychiatric hospital inpatient admissions: 93.2% of total costs in the FGA arm and 81.5% in the SGA arm. Antipsychotic drug costs accounted for a small proportion of total costs (2.1% in the FGA arm and 3.8% in the SGA arm).

Polypharmacy before randomization (Table 1) and at the end of the study (**Table 6**) were similar in the 2 groups. More patients randomized to receive an SGA than an FGA remained in their allocated treatment arm for the whole year, but this difference was not significant (65% [71/109] vs 54% [64/118]; $P=.1$) (Figure 1). Twenty-eight (48%) of the 58 patients randomized to the FGA arm and prescribed sulpiride were still taking that drug at the end of the study, although 3 were receiving another antipsychotic drug in addition. Thirty-seven (74%) of the 50 patients randomized to receive SGAs and who were prescribed olanzapine were still taking the drug at the end of the study. Participants reported no clear preference for either class of drug at any stage.

COMMENT

The results of this pragmatic randomized trial refute the hypothesis that the use of SGAs is superior to the use of FGAs in terms of quality of life at 1 year. Clinical superiority had been defined a priori as a 5-point difference in the QLS score. Statistical precision was limited, but the ITT analysis indicated that true effects may have been in the opposite direction for this primary outcome and for the main symptom assessments. The confidence intervals for this effect in the opposite direction were wide, including the possibility of a small benefit for SGAs but much smaller than we had hypothesized.

Why did the trial fail to find a clinical advantage for SGAs? The first possibility is that the proposed effect size

Table 5. Results of the Secondary Outcomes

Variable	FGA Arm		SGA Arm		Estimate	SE (95% CI)	*P* Value
	Patients, No.	Score, Mean (SD)	Patients, No.	Score, Mean (SD)			
PANSS*							
Total							
Baseline	118	72.9 (17.2)	109	71.3 (16.5)	2.3	1.5 (–0.6 to 5.2)	.13
52 wk	99	64.6 (15.1)	86	66.2 (17.5)			
After multiple imputations*					2.4	1.7 (–0.2 to 5.4)	.11
Positive subscale							
Baseline	118	15.9 (5.9)	109	15.5 (5.4)			.90
52 wk	99	13.9 (4.5)	86	14.0 (5.3)			
Negative subscale							
Baseline	118	20.6 (6.9)	109	20.0 (6.5)			.84
52 wk	99	17.3 (5.8)	86	18.2 (6.1)			
General subscale							
Baseline	118	36.4 (8.8)	109	35.8 (9.5)			.92
52 wk	99	33.3 (8.1)	86	34.0 (9.4)			
Global Assessment of Functioning scale†							
Total							
Baseline	118	45.6 (14.9)	108	42.7 (13.6)	0.1	1.4 (–2.6 to 2.7)	.96
52 wk	100	52.4 (13.3)	85	52.3 (13.9)			
Symptoms							
Baseline	103	44.42 (15.2)	96	43.1 (14.9)			.89
52 wk	100	50.96 (14.3)	85	51.6 (14.6)			
Disability							
Baseline	103	44.6 (14.1)	96	41.9 (13.3)			.94
52 wk	100	52.8 (13.0)	85	52.6 (14.4)			
Calgary depression scale*							
Baseline	118	6.6 (5.0)	108	6.9 (5.2)	0.5	0.4 (–0.2 to 1.3)	.16
52 wk	98	4.2 (3.8)	83	5.0 (3.9)			
Drug attitude inventory*							
Baseline	117	8.2 (11.5)	108	10.6 (10.6)	1.3	1.0 (–0.7 to 3.4)	.20
52 wk	96	10.9 (11.6)	81	14.4 (10.1)			
Compliance scale*							
Baseline	116	5.1 (1.3)	108	5.1 (1.4)			.20
52 wk	104	5.0 (1.5)	96	5.2 (1.6)			
Simpson-Angus extrapyramidal adverse effects scale*							
Baseline	115	4.4 (5.2)	104	4.2 (4.6)	0.2	0.4 (–0.6 to 0.9)	.66
52 wk	94	3.0 (3.9)	80	3.2 (3.7)			
Barnes Akathisia Rating Scale*							
Baseline	118	2.4 (3.0)	107	3.2 (2.6)	0.01	0.3 (–0.5 to 0.5)	.98
52 wk	95	1.5 (2.4)	81	2.0 (2.7)			
Abnormal Involuntary Movement Scale*							
Baseline	118	1.7 (2.9)	107	1.8 (3.3)	–0.2	0.4 (–1.0 to 0.6)	.63
52 wk	95	2.3 (4.5)	81	1.8 (3.3)			
Total of previous 3 scales*							
Baseline	115	8.5 (7.3)	104	9.0 (7.8)	–0.2	0.7 (–1.5 to 1.2)	.08
52 wk	94	6.8 (6.7)	80	7.1 (6.3)			
ANNSERS*							
Baseline	117	14.6 (9.3)	103	15.6 (9.6)	1.1	0.8 (–0.4 to 2.6)	.14
52 wk	95	10.8 (7.7)	82	12.5 (8.4)			

Abbreviations: ANNSERS, Antipsychotic Non-Neurological Side-Effects Rating Scale; CI, confidence interval; FGA, first-generation antipsychotic; PANSS, Positive and Negative Syndrome Scale; SGA, second-generation antipsychotic.

*High scores on this scale mean a worse outcome. A positive parameter estimate means that participants in the FGA arm are doing better.

†High scores on this scale mean a better outcome. A negative parameter estimate means that participants in the FGA arm are doing better.

of 5 points on the QLS was unrealistically large. Designing the trial to show equivalence between FGAs and SGAs would not have tested a clinically meaningful question given the observed migration of prescription toward the more expensive class since its introduction. An improvement of 5 points in the QLS score resulting from a change in treatment because of adverse effects or lack of effect from previous treatment is a reasonable clinical aim that has also been used in similar trials.[24]

The second, related possibility is the limited sample size and statistical power. Clinical equipoise shifted in favor of SGAs during the trial, pressurizing recruitment. However, good follow-up and a close correlation between QLS score at baseline and follow-up

meant that the recruited sample gave 75% power to detect the hypothesized difference in QLS scores. Participants in the FGA arm tended to have greater improvements in QLS and symptom measures than those in the SGA arm, suggesting that the failure to find an advantage for SGAs was not due to the sample simply being too small. We emphasize that we do not present a null result; the hypothesis that SGAs are superior was clearly rejected.

Third, quality of life is difficult to assess in schizophrenia, and insensitivity and imprecision of the QLS would have reduced power, although similarly in both trial arms. Furthermore, there is a striking consistency of findings across the primary and secondary outcomes and between interviewer ratings and self-report. Nevertheless, the choice of the QLS deserves scrutiny.

A good quality-of-life scale should be appropriate to the study population, the clinical condition, and the illness phase; have established psychometric properties; and measure several dimensions.[25] There is no perfect scale for schizophrenia, but the QLS fares well on these criteria and is widely used in schizophrenia studies.[12] One of several quality-of-life measures in the Veterans Affairs Cooperative Study in Health Services No. 17 comparing clozapine and haloperidol in refractory schizophrenia,[26] the QLS was sensitive to subtle change and treatment effect.[24] Criticisms include its administration by an external assessor (although self-report has problems) and being affected by symptoms.[27,28] Regarding the latter point, the PANSS total score in CUtLASS 1 accounted for only 30% of the variance in QLS scores at baseline. Overall, the QLS seems to be a reasonable choice.

Finally, we have to consider the participants: patients, psychiatrists, and researchers. Regarding the study sample, PANSS total and other scores were similar at baseline to those of other treatment trials in schizophrenia. Randomization was satisfactory, although more participants were referred owing to adverse effects in the FGA arm. Given that any disadvantage for this class may have been due to adverse effects, any resulting bias would have operated against, not for, these drugs. This factor was included as a covariate, and there was no evidence of differential outcomes for patients referred to the trial because of treatment intolerance compared with those entering because of inadequate response.

Overall, the patients had fairly long-term illness, and treatment effects were not large. The results may have been clearer in subgroups of patients with certain clinical features or shorter duration of illness, for example. However, the trial was designed to mimic the clinical situation, including problematic differential diagnoses, such as delusional disorder vs schizophrenia, and choice of drug from within the class. This selection will have been driven by psychiatrist and patient choice, and it supports the applicability of the results to routine clinical practice. Nevertheless, the trial was biased toward schizophrenia that had shown an inadequate response to treatment, an area in which it is most difficult to achieve and demonstrate major change. The patient sample was not skewed toward those who had previously failed to respond to an SGA; most participants were being treated with an FGA before randomization.

Table 6. Polypharmacy at 52 Weeks

Pattern of Polypharmacy	Patients, No.	
	Randomized to FGA Arm	Randomized to SGA Arm
FGA + FGA	8	2
FGA + SGA	3	8
SGA + SGA	0	0
FGA + clozapine	0	0
SGA + clozapine	1	0
Total	**12**	**10**

Abbreviations: FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.

This trial was independent of industry, being funded by the National Health Service. This organization also has interests in treatment costs, although its Health Technology Assessment Program is charged with providing objective evidence on interventions. If the investigators, themselves, had any bias or previous expectation it was in favor of SGAs; we were surprised to refute the hypothesis. Participating psychiatrists used appropriate drug doses in both classes (Table 2) but may have been less ready to change from SGAs in the face of nonresponse during the trial compared with FGAs. However, the data did not indicate that this was the case. Many psychiatrists who took part in the trial were, inevitably, particularly interested in schizophrenia management, and their ability to individualize treatments within the randomized arms may have minimized rather than emphasized differences in outcome.

Two recent systematic reviews[3,4] provided evidence that some SGAs are more efficacious than others, so our comparison of the 2 groups of drugs may have masked the effects of individual drugs that have particular efficacy or tolerability advantages (or disadvantages) for subgroups of patients or between individuals. We do not think that this was a problem in the present trial; Lewis and colleagues[29] used the same pragmatic design, ITT analysis, and primary outcome to demonstrate the superiority of clozapine over SGAs as a group in treatment-resistant schizophrenia (CUtLASS 2). This suggests that the present trial design was sensitive enough to show the effect we hypothesized, had it been present. Although we note the considerable pharmacologic heterogeneity within and between the FGA and SGA groups, we consider the comparison between groups to have been clinically useful.

In contrast to published efficacy trials, sulpiride was the FGA chosen most often by psychiatrists, whereas haloperidol, the standard industry comparator, was selected infrequently; of 8 patients prescribed haloperidol at baseline, only 2 were still using it at 52 weeks (albeit at high doses). The point has been made that haloperidol carries a considerable adverse effect burden, particularly at the relatively high doses often selected for its role as comparator in efficacy trials.[2] The fact that so few psychiatrists opted for haloperidol in the present trial reflects current clinical practice and inevitably hinders interpretation of the results in the context of exist-

ing systematic reviews[3,4] of studies using this drug as standard treatment.

Sulpiride, a more selective dopamine D_2 receptor blocker than haloperidol, is a low-potency FGA that has been licensed in England since the 1960s. Despite its name, its pharmacologic features have little in common with amisulpride, an SGA. Anecdotally, sulpiride is sometimes thought to pose a lower risk of EPSs than other FGAs; this may have been the reason it was chosen relatively frequently. If the preference for sulpiride in the FGA arm explained the results, this drug would have to have remarkably superior efficacy and relative atypicality to negate a real advantage of SGAs, particularly when any such effect would be diluted among other FGAs. Neither property has been supported by a systematic review of trials of sulpiride.[30]

There were slightly more patients receiving depot preparations before randomization among those allocated to the FGA compared with the SGA group. Any residual benefit was unlikely to have persisted during the 1-year follow-up, although it could have been operating in the earlier months of the trial. The decision to use depot preparations at randomization was not common (12 in the FGA arm) compared with previous treatment. Similar numbers of patients in each arm were being treated with depot FGAs at 1 year (18 in the FGA arm and 17 in the SGA arm). Improved adherence to treatment in patients considered to be in the SGA arm according to the ITT design but who were receiving a depot FGA 1 year later may have given a spurious advantage to the SGA group in terms of efficacy but at a cost in terms of adverse effects. Again, the effect would have needed to be unrealistically large to have generated our results. The CUtLASS 1 and 2 studies predated the availability of any depot SGA preparations, and trials including these are required.

The per-protocol estimate of the treatment effect of the randomized class for the first 3 months of the trial should be interpreted with care because it may be subject to selection biases. However, the effect estimate was similar to the primary ITT analysis, suggesting that switching between classes during the first 3 months had little impact on the result (Table 4).

Much evidence concerning the relative efficacy of FGAs and SGAs comes from relatively short-term trials; dropout rates are high, and effects are assessed using symptom ratings rather than broader outcomes.[7,31] It is reasonable to speculate that the superior tolerability and possible benefits in efficacy in these studies might translate into better treatment adherence, improved clinical effectiveness, and enhanced quality of life, but, as yet, few data support such a view. The doses of some SGAs have become higher in routine clinical practice than those used in the original preregistration trials. These trials provide benchmark data on adverse effect burden, but this may represent an underestimate. Furthermore, a range of adverse effects of FGAs and SGAs is emerging. Serious weight gain,[32] diabetes mellitus,[33] and hyperlipidemia[34] may all adversely affect quality of life.

One observational study[35] supports our result, but there have been few pragmatic, long-term, randomized studies of the clinical effectiveness of FGAs vs SGAs. Two such studies stand out. In a study by Rosenheck and colleagues,[36] 309 patients were randomized to receive olanzapine, an SGA, and the classic FGA haloperidol, with flexible dosing and the use of prophylactic anticholinergic drugs. This double-blind comparison did not reveal any advantages at 1 year for olanzapine in treatment adherence, symptoms, EPSs, or overall quality of life as measured using the QLS. Benefits in terms of a reduction in observed akathisia and improved cognition were weighed against the problems of weight gain and higher costs.

Lieberman and colleagues[37] reported an 18-month double-blind trial in which 1493 patients with chronic schizophrenia were randomized to receive olanzapine, quetiapine, risperidone, ziprasidone, or perphenazine, a low-potency FGA. Despite circumventing the haloperidol comparator problem, most patients in each group discontinued the assigned treatment because of lack of effect or intolerability. Olanzapine treatment was associated with the lowest risk of discontinuation and a different adverse effect profile. The remaining SGAs differed neither from each other in terms of effectiveness nor from perphenazine.

Overall, the results of these US studies are in line with the data we present from England. All the data suggest that careful prescribing of FGAs, at least in the context of a trial, is not associated with poorer efficacy or a greater adverse effect burden, both of which would translate into lower quality of life in the medium term. This suggests that despite recent policy statements and prescribing patterns, further randomized and other evaluations of SGAs would still be useful in establishing their role in the long-term management of schizophrenia and, likewise, the continued role of older drugs.

In conclusion, there is no disadvantage in terms of quality of life, symptoms, or associated costs of care across 1 year in commencing treatment with FGAs rather than atypical SGAs in people with schizophrenia whose medication is being changed because of intolerance or inadequate response and who are treated in the context of a pragmatic trial.

Submitted for Publication: February 24, 2005; final revision received February 17, 2006; accepted February 23, 2006.

Correspondence: Peter B. Jones, MD, PhD, Department of Psychiatry, University of Cambridge, Box 189 Addenbrooke's Hospital, Cambridge CB2 2QQ, England (pbj21@cam.ac.uk).

Financial Disclosure: Dr Jones has acted as a consultant to Bristol Myers Squibb, Otsuka, Eli Lilly, and Janssen Cilag. Dr Barnes has acted as a consultant to Servier and Johnson & Johnson Pharmaceutical Services.

Funding/Support: This study was supported by project grant 96/19/06 from the Secretary of State for Health under the United Kingdom National Health Service Health Technology Assessment Program. Drs Jones, Murray, and Lewis gratefully acknowledge research support from the Stanley Medical Research Institute.

REFERENCES

1. Barnes TRE, Edwards JG. The side-effects of antipsychotic drugs, I: neuropsychiatric effects. In: Barnes TRE, ed. *Antipsychotic Drugs and Their Side-Effects*. London, England: Academic Press; 1993:213-247.
2. Geddes J, Freemantle N, Harrison P, Bebbington P. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ*. 2000;321:1371-1376.
3. Bagnall AM, Jones L, Ginnelly L, Lewis R, Glanville J, Gilbody S, Davies L, Torgerson D, Kleijnen J. A systematic review of atypical antipsychotic drugs in schizophrenia. *Health Technol Assess*. 2003;7:1-193.
4. Davis JM, Chen N, Glick ID. A meta-analysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry*. 2003;60:553-564.
5. Voruganti L, Cortese L, Oyewumi L, Cernovsky Z, Awad A. Comparative evaluation of conventional and novel antipsychotic drugs with reference to their subjective tolerability, side-effect profile and impact on quality of life. *Schizophr Res*. 2000;43:135-145.
6. Leucht S, Wahlbeck K, Hamann J, Kissling W. New generation antipsychotics versus low-potency conventional antipsychotics: a systematic review and meta-analysis. *Lancet*. 2003;361:1581-1589.
7. Leucht S, Barnes TR, Kissling W, Engel RR, Correll C, Kane JM. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. *Am J Psychiatry*. 2003;160:1209-1222.
8. National Collaborating Centre for Mental Health. *Guidance on the Use of Newer (Atypical) Antipsychotic Drugs for the Treatment of Schizophrenia*. London, England: National Institute for Clinical Excellence; 2002.
9. National Collaborating Centre for Mental Health. *Core Interventions in the Treatment and Management of Schizophrenia in Primary and Secondary Care*. London, England: National Institute for Clinical Excellence; 2002.
10. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 1994.
11. Heinrichs DW, Hanlon TE, Carpenter WT. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull*. 1984;10:388-398.
12. Lehman AF. Measures of quality of life among persons with severe and persistent mental disorders. *Soc Psychiatry Psychiatr Epidemiol*. 1996;31:78-88.
13. Lehman AF. *Measures of Quality of Life for People With Severe Mental Disorders: Mental Health Outcome Measures*. London, England: Gaskell; 2001.
14. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13:261-276.
15. Addington D, Addington J, Schissel B. A depression rating scale for schizophrenics. *Schizophr Res*. 1990;3:247-251.
16. Hogan TP, Awad AG, Eastwood R. A self-report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity. *Psychol Med*. 1983;13:177-183.
17. Hayward P, Chan N, Kemp R, Youle S, David A. Medication self-management: a preliminary report on an intervention to improve medication compliance. *J Ment Health*. 1995;4:511-518.
18. Simpson GM, Angus JW. A rating scale for extrapyramidal side effects. *Acta Psychiatr Scand Suppl*. 1970;212:11-19.
19. Barnes TR. The Barnes Akathisia Rating Scale: revisited. *J Psychopharmacol*. 2003;17:365-370.
20. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Rev ed. Rockville, Md: National Institutes of Health; 1976.
21. Yusufi BZ, Mukherjee S, Aitchison K, Dunn G, Page E, Barnes TR. Inter-rater reliability of the Antipsychotic Non-Neurological Side-Effects Rating Scale (ANNSERS). *Schizophr Bull*. 2005;31:574.
22. Little RJA, Rubin DB. *Statistical Analysis With Missing Data*. New York, NY: John Wiley & Sons; 2002.
23. Rubin DB, Schenker N. Multiple imputation in health-care data bases: an overview and some applications. *Stat Med*. 1991;10:585-598.
24. Cramer JA, Rosenheck R, Xu W, Thomas J, Henderson J, Charney D. Quality of life in schizophrenia: a comparison of instruments *Schizophr Bull*. 2000;26:659-666.
25. Awad AG, Voruganti LNP. Intervention research in psychosis: issues related to the assessment of quality of life. *Schizophr Bull*. 2000;26:557-564.
26. Rosenheck R, Cramer J, Xu W, Thomas J, Henderson W, Frisman L, Fye C, Charney D; Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. *N Engl J Med*. 1997;337:809-815.
27. Voruganti L, Heslegrave R, Awad AG, Seeman MV. Quality of life measurement in schizophrenia: reconciling the quest for subjectivity with the question of reliability. *Psychol Med*. 1998;28:165-172.
28. Atkinson M, Zibin S, Chuang H. Characterising quality of life among patients with chronic mental illness: a critical examination of the self-report methodology. *Am J Psychiatry*. 1997;154:99-105.
29. Lewis SW, Barnes TRE, Davies L, Murray RM, Dunn G, Hayhurst KP, Markwick A, Lloyd H, Jones PB. Randomized controlled trial of effect of prescription of clozapine versus other second-generation antipsychotic drugs in resistant schizophrenia [epub ahead of print]. *Schizophr Bull*. 2006. doi:10.1093/schbul/sbj067. Accessed June 12, 2006.
30. Soares BG, Fenton F, Chue P. Sulpiride for schizophrenia [Cochrane Review on CD-ROM]. Oxford, England: Cochrane Library, Update Software; 2002;issue 3.
31. Duggan L, Fenton M, Dardennes RM, El-Dosoky A, Indran S. Olanzapine for schizophrenia [Cochrane Review on CD-ROM]. Oxford, England: Cochrane Library, Update Software; 2001:issue 1.
32. Casey DE, Zorn SH. The pharmacology of weight gain with antipsychotics. *J Clin Psychiatry*. 2001;62:4-10.
33. American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, North American Association for the Study of Obesity. Consensus development conference on antipsychotic drugs and obesity and diabetes. *J Clin Psychiatry*. 2004;65:267-272.
34. Melkersson KI, Hulting AL, Brismar KE. Elevated levels of insulin, leptin, and blood lipids in olanzapine-treated patients with schizophrenia or related psychoses. *J Clin Psychiatry*. 2000;61:742-749.
35. Kilian R, Dietrich M, Toumi M, Angermeyer C. Quality of life in persons with schizophrenia in out-patient treatment with first- or second-generation antipsychotics *Acta Psychiatr Scand*. 2004;110:108-118.
36. 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; Department of Veterans Affairs Cooperative Study Group on the Cost-Effectiveness of Olanzapine. Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: a randomized controlled trial. *JAMA*. 2003;290:2693-2702.
37. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RS, Davis SM, Davis CE, Lebowitz BD, Severe J, Hsiao JK; Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with schizophrenia. *N Engl J Med*. 2005;353:1209-1223.