

---

# Mental Health Care

---

## Multiple Outcome Assessment in a Study of the Cost-Effectiveness of Clozapine in the Treatment of Refractory Schizophrenia

*Robert Rosenheck, Joyce Cramer, Weichun Xu, John Grabowski, Richard Douyon, Jonathan Thomas, William Henderson, and Dennis Charney, for the Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia*

---

**Objective.** To develop new methods for combining results from multiple outcome domains and to demonstrate their application in a study of the cost-effectiveness of clozapine in treating hospitalized patients with refractory schizophrenia.

**Data Sources/Study Setting.** Interview assessments, and administrative utilization and cost data, concerning 423 patients with refractory schizophrenia who had been hospitalized for 30–364 days during the year before study entry, at 15 VA medical centers.

**Study Design.** A 12-month double-blind trial compared clozapine ( $n = 205$ ) and haloperidol ( $n = 218$ ) in the treatment of refractory schizophrenia.

**Data Collection/Extraction Methods.** Data from standard assessment instruments, gathered at baseline and at 6 weeks, and at 3, 6, 9, and 12 months, were used to develop a Composite Health Index for Schizophrenia, a measure that addresses outcome in six domains, weighted by patient or provider preferences. Cumulative improvement was estimated by computing the area under the improvement curve. This measure was then combined with cost data, reflecting consumption of societal resources to estimate incremental cost-effectiveness ratios.

**Principal Findings.** Clozapine was significantly more effective than haloperidol on measures of symptoms ( $p = .02$ ) and side effects ( $p < .0001$ ), with nonsignificant trends in the positive direction on community role functioning ( $p = .06$ ), family relationships ( $p = .23$ ), social relationships ( $p = .30$ ), and daily activities ( $p = .20$ ). Clozapine was also more effective than haloperidol on the one-year cumulative Composite Health Index for Schizophrenia ( $p < .0001$  for all weighting schemes). After converting this measure to a 0–1 Worst Health–Good Health Scale analogous to Quality Adjusted Life Years, clozapine was found to yield a small improvement of .049 Worst Health–Good Health Units as compared to an improvement of only .027 Units for haloperidol ( $p < .0001$ ). Average annual costs were \$2,733 lower for clozapine (95% C.I. =  $-\$9,220$  to  $\$3,754$ ). Although clozapine was significantly more effective than haloperidol, the

summary cost-effectiveness ratio had a wide 95 percent confidence interval ranging from -\$431,585 to \$177,352.

**Conclusions.** Methods demonstrate an approach to using conventional disease-specific measures to evaluate the cumulative effectiveness of novel treatments for psychotic disorders and for expressing their economic effect as cost-effectiveness ratios. Among high hospital users with refractory schizophrenia, clozapine is more cost-effective than standard treatment, although the magnitude of its effect is small and there is considerable uncertainty about the cost estimates.

**Key Words.** Cost-effectiveness, clozapine, schizophrenia, clinical trial, pharmacoeconomics

---

Schizophrenia is severe chronic illness accounting for 2.5 percent of U.S. healthcare expenditures and 10 percent of all permanent disability in the United States (Rupp and Keith 1993). Almost one-third of patients fail to respond to standard antipsychotic medications (Kane and Marder 1993). In a recent report, from a 12-month randomized controlled trial conducted at 15 Department of Veterans Affairs (VA) medical centers, we showed that the new antipsychotic agent, clozapine, is an effective treatment for the symptoms of schizophrenia and for improving quality of life (Rosenheck, Cramer, Xu, et al. 1997). This report served to extend the findings of previous short-term studies (Kane et al. 1988; Carpenter, Conley, Buchanan et al. 1996).

Clozapine is over ten times more expensive than conventional antipsychotic medications, in part because it is associated with potentially fatal agranulocytosis whose consequences can be almost completely avoided with costly blood monitoring on a weekly basis (Alvir, Lieberman, Safferman,

---

Supported by the Department of Veterans Affairs Health Services Research Service. Clozapine placebo were provided by Sandoz Pharmaceuticals, a division of Novartis Corp.

Address correspondence and requests for reprints to Robert Rosenheck, M.D. (182), Director, VA Northeast Program Evaluation Center, VA Connecticut Healthcare System, 950 Campbell Avenue, West Haven, CT 06516. Dr. Rosenheck is also Clinical Professor, Yale University School of Medicine, Department of Psychiatry. Joyce Cramer and Dennis Charney, M.D. are Project Director and Service Chief, respectively at the VA Connecticut Healthcare System and the Yale School of Medicine, Department of Psychiatry. Weichun Xu, Ph.D. and Jonathan Thomas, M.S. are Research Scientists, and William Henderson, Ph.D. is Chief at the Center for Cooperative Studies in Health Services, Hines VA Medical Center, Hines, IL; John Grabowski, M.D. is Chief of Psychiatry, the Detroit VA Medical Center and Wayne State University School of Medicine; and Richard Douyon, M.D. is a Psychiatrist with the Miami VA Medical Center and the University of Miami School of Medicine. This article, submitted to *Health Services Research* on August 21, 1997, was revised and accepted for publication on March 16, 1998.

et al. 1993). Although nonexperimental studies suggest that reductions in hospital costs with clozapine therapy can offset the high cost of drug treatment (Meltzer, Cola, and Way 1993; Reid, Mason, and Toprac 1994), those studies have lacked adequate control groups. A recent controlled trial of clozapine in long-stay state hospital patients, however, showed a reduction in hospital days (Essock et al. 1996), as did our multisite VA trial (Rosenheck, Cramer, Xu, et al. 1997). In the VA trial, clozapine reduced total societal costs, but group differences did not reach statistical significance. Our previous report presented findings on the statistical significance of two main outcome measures and data on associated costs. However, in the absence of established methods for integrating data from such studies, it did not (1) present a synthesis of clinical improvement combining results in all relevant life domains; (2) weight the results for patient preferences; (3) address the magnitude and importance of the observed improvement beyond their statistical significance; or (4) present cost-effectiveness ratios with appropriate 95 percent confidence intervals. In the current article we describe methods for accomplishing these objectives and apply them to our study of the cost-effectiveness of clozapine.

The preferred measure of cost-effectiveness is the cost-effectiveness ratio (Gold et al. 1996; Spilker 1996), the ratio of (1) the incremental cost of the treatment in comparison to that of the next best alternative, to (2) the incremental effectiveness of the treatment, also in comparison to the next best alternative. Computation of such ratios requires the calculation of a single measure of effectiveness or improvement that can be scaled in meaningful units. This is a major challenge in illnesses such as schizophrenia that severely affect many dimensions of life functioning and experience; each of which is assessed with its own instruments; and each of which may be affected to a different degree in each particular patient.

This article thus illustrates an approach to applying standard methods of cost-effectiveness analysis to the study of interventions for serious mental illnesses using instruments specific to the multiple outcomes that are affected by these illnesses.

## METHODS

This study was a prospective, double-blind clinical trial in which patients at 15 VA medical centers were randomly assigned to clozapine or haloperidol and were treated for 12 months.

### *Entry Criteria*

The study focused on patients with refractory schizophrenia and with 30–364 days of hospitalization during the previous year. Entry criteria and details of both pharmacological and psychosocial treatment have been presented elsewhere (Rosenheck, Cramer, Xu, et al. 1997).

### *Assessment of Outcomes*

Outcomes were assessed with six measures. (1) Symptoms were measured with the Structured Clinical Interview for Positive and Negative Syndrome Scale (PANSS) for schizophrenia (Kay, Fiszbein, and Opler 1987). The Quality of Life Scale (Heinrichs, Hanlon, and Carpenter 1984), a clinician-rated scale, was used to assess (2) social relationships (frequency, intensity and quality of social interaction); (3) role functioning (employment and community adaptation); and (4) general daily activity and recreation. (5) Family relationships were evaluated using a section of the Quality of Life Interview (Lehman 1988); and (6) medication side effects were assessed with standard scales for akathisia, tardive dyskinesia, and extrapyramidal symptoms (Barnes 1989; Guy 1976; Simpson and Angus 1970).

### *Assessment of Costs*

*Healthcare Costs.* Healthcare costs were estimated by multiplying the number of units of specific services by their estimated unit costs for each patient.

*Unit Costs.* Unit costs for VA general psychiatry, substance abuse and medical-surgical inpatient and outpatient care were estimated for each participating medical center, using cost data from VA's site-specific Cost Distribution Report (CDR) and computerized workload data. Non-VA healthcare costs were minimal (less than 2 percent of all costs) and were estimated on the basis of a recent study that compared VA and non-VA costs in various communities (Office of the Inspector General 1992).

*Service Utilization.* VA health service use by study patients was measured using data from VA's national computerized workload systems. Utilization of non-VA services was evaluated through monthly patient interviews and validated using treatment records from non-VA providers.

*Drug and Specific Clozapine-Related Costs.* The costs of study medication were estimated at VA pharmacy costs. Since the double-blind design of the study artificially inflated the cost of haloperidol treatment by requiring weekly clinic visits and blood draws, outpatient costs for the haloperidol group were deflated to 73 percent of their actual value.

*Non-healthcare Costs.* Interview data were used to measure non-healthcare services use. Unit costs were derived both from interview data and from published literature (Rosenheck, Cramer, Xu, et al. 1997). These costs included (1) the administrative costs of transfer payments (e.g., disability, welfare, etc.); (2) criminal justice system costs; (3) productivity (estimated by employment earnings, included as negative costs); and (4) family burden (days lost by family members from work and from unpaid domestic activity).

*Summary Cost Estimates.* Cost data were summarized from the perspective of society (healthcare plus all non-healthcare costs, less productivity) at the individual patient level.

### *Composite Health Index for Schizophrenia*

A composite index of health status was constructed using the six rating scales described earlier. In combining these scales we weighted them by their importance to patients and providers using results from a study by Kleinman (1995) involving 40 severely mentally ill clients and 40 providers from a university-based case management program. In that study (1) focus groups were conducted with patients, providers, and family members, to identify principal outcome domains for assessment of utility in severe mental illness; and (2) a categorical rating scale questionnaire was constructed and used to evaluate the relative importance of outcomes to both patients and providers in each of the six domains. The six domains identified as most important to patients, providers, and families through this procedure are symptoms; side effects; family relationships; social relationships; daily activities (e.g. recreation); and community (role) functioning (e.g. employment). The specific measures used to evaluate each domain are presented in Table 1 (column 2).

Second, we calculated standardized ( $Z$ ) scores for the measures in each domain by dividing the difference between each observation on each measure and the baseline group mean of that measure by the standard deviation of the baseline mean (Table 1, columns 3 and 4) (Cohen 1969). Third, we combined the standardized scores from each of the six domains by calculating their weighted average using preference weights derived from the Multi-Attribute Utility Functions estimated by Kleinman (1995) (Table 1, columns 5 and 6). The units of the Composite Health Index for Schizophrenia are thus average, weighted, standardized ( $Z$ ) scores.

### *Cumulative Improvement*

To assess cumulative improvement over time we computed the area under the curve defined for each patient by his or her sequential measures on the

Table 1: Domains, Measures, and Derivation of the Composite Health Index for Schizophrenia (N = 423)

| 1-Outcome Domain   | 2-Specific Measure                   | 3-Baseline Mean | 4-Baseline Standard Deviation | 5-Patient Preference Weight** | 6-Provider Preference Weight** | 7-Full Scale Range | 8-Baseline Standard Deviations from in Full Range†† |
|--|--------------------------------------|-----------------|-------------------------------|-------------------------------|--------------------------------|--------------------|---|
| Symptoms   | PANSS Total Score*                   | 91.6            | 14.7                          | 0.14                          | 0.08                           | 0-180              | 12.24   |
| Side effects   | Side effects composite†              | 5.5             | 5.4                           | 0.14                          | 0.19                           | 0-40               | 7.48  |
| Family relationships   | Lehman QOLI Family Subscale‡         | 18.1            | 5.5                           | 0.19                          | 0.19                           | 0-24               | 4.36  |
| Social relationships   | QOL Relationship subscale§           | 14.1            | 7.8                           | 0.19                          | 0.14                           | 0-48               | 6.14  |
| Daily activities   | QOL Objects and Activities subscale§ | 4.8             | 2.5                           | 0.22                          | 0.25                           | 0-12               | 4.83  |
| Community (role) function  | QOL Role Function subscale§          | 2.2             | 3.4                           | 0.11                          | 0.14                           | 0-24               | 7.04  |
| Weighted sum (Column 8 weighted by column 5)                         |                                      |                 |                               |                               |                                |                    | 6.59  |
| 80% of weighted sum (Average s.d.s from worst health to good health) |                                      |                 |                               |                               |                                |                    | 5.27  |

\* Positive and Negative Syndrome Scale (Kay, Fiszbein, and Opler 1987).

† Combined score on the AIMs (Guy 1976), Simpson (Simpson and Angus 1970), and Barnes (Barnes 1989) Side Effects Scales.

‡ Lehman Quality of Life Interview (Lehman 1988).

§ Heinrichs-Carpenter Quality of Life Scale (Heinrichs, Hanlon, and Carpenter 1984).

\*\* From Kleinman (1995).

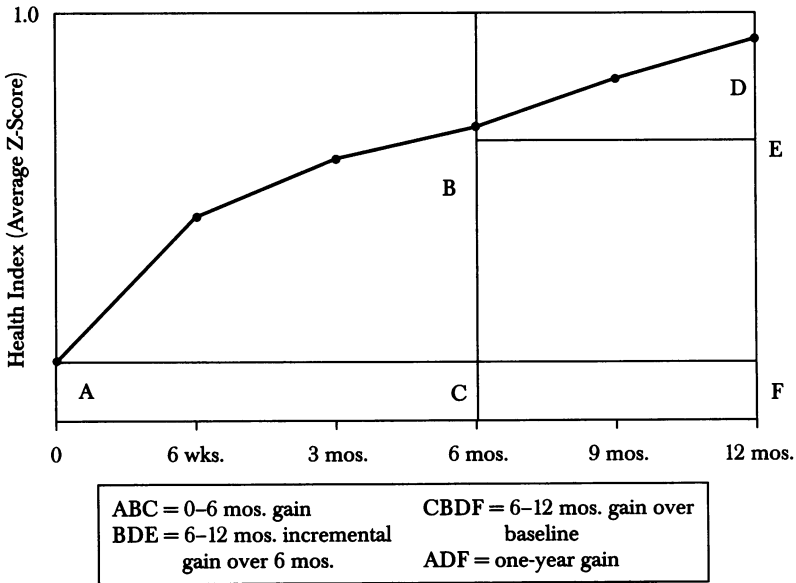
†† Baseline s.d.s in the complete range of each scale (column 7/column 4).

Composite Health Index for Schizophrenia. Rather than evaluating overall pre-post change across the year, this cumulative measure reflects the person's total time in good health, taking into consideration whether improvement occurred early or late in the follow-up period (see Figure 1 and the section on Statistical Analyses, further on).

*Scaling from Worst Health to Good Health*

Although several measures have been developed in recent years to assess utility as measured in standard Quality Adjusted Life Year units (QALYs) (Gold et al. 1996; Spilker 1996), the primary focus of these measures is on problems such as physical pain and motor functioning that are not typically affected by psychotic illnesses. These instruments cannot be used in evalu-

Figure 1: Cumulative Gain Measured by the Area Under the Change Curve. Diagram of Area-Under-the-Curve Estimates of Cumulative Improvement—Baseline to Six Months (Area ABC); Incremental Improvement During the Second Six Months, Beyond That Achieved During the First Six Months (Area BDE); Improvement Between Six and Twelve Months over the Baseline Levels (Area CBDF); and Cumulative Improvement over the Entire Year Calculated (Area ADF)



ating outcomes in schizophrenia. Instead, we developed a method of scaling our Composite Health Index for Schizophrenia on a 0–1, Worst Health to Good Health scale that is analogous to, but not the equivalent of, conventional QALY measures. Although we recognize that using this scale as a metric does not mean that it can be interpreted in terms of utilities, our thinking is that the procedure we describe provides an informative way of scaling outcomes more meaningfully for patients with schizophrenia. First we divided the full range of possible scores for each of the six domain measures (Table 1, column 7) by the standard deviation of the mean baseline score (Table 1, column 1). This transformed the range from worst possible health to best possible health on each scale into standardized *z*-score units of the type used for the Composite Health Index for Schizophrenia (Table 1, column 8). We then used patient preference weights to compute the weighted average of these scores (Table 1, second-to-lowest row of column 8). This generated an estimate of the average number of baseline standard deviations, or *z*-score units on our Composite Health Index that represent the full range of health states, from the worst possible health on all domains to the best possible health on all domains (6.59 standard score units). Assuming that the lower 10 percent of the range represents worst-case quality of life, and that scores at 90 percent or higher represent the equivalent of good health, we took 80 percent of this range as representing a scale from worst health to good health, an analogue of the death to perfect health range represented in standard QALY measures ( $0.8 \times 6.59 = 5.27$  average standard deviations; see bottom row of column 8 in Table 1).

### *Statistical Analyses*

To maximize statistical power for testing hypotheses for longitudinal data, primary outcomes were analyzed using Random Effects Regression models (Gibbons, Hedeker, and Elkin 1993), conducted using PROC MIXED of SAS, version 6.12. These models accommodate correlations among the repeated observations and therefore allow the inclusion of individuals with missing observations. The significance of differences between treatment conditions in random regression analyses was tested using the likelihood ratio chi-square test. We thus compared models for each outcome that included the effects of time, time-squared, and treatment condition, to a model that added group by time interactions. The group by time interaction is the hypothesis of interest.

As noted earlier, the area under each regression curve was used as a summary measure of cumulative outcome. The area under the improvement



curve was determined by integrating the fixed-effect terms and the individual's own random-effects terms of the regression line between targeted assessment periods, after adjusting for the initial effects. The first period of improvement covered the time from baseline to six months (area ABC in Figure 1). The second period consisted of the area representing new improvement beyond that achieved during the first six months (BDE). A third period represented sustained improvement between 6 and 12 months over the baseline levels (CBDF). Finally cumulative improvement over the entire year was calculated (ADF), with *t*-tests used to compare group means on this measure of cumulative change.

Incremental cost-effectiveness ratios were calculated as the ratio of (1) average costs for the clozapine group less average costs for the haloperidol group to (2) the average effectiveness, measured in worst health–good health units for the clozapine group less the average effectiveness for the haloperidol group. The 95 percent confidence interval of the cost-effectiveness ratio was estimated using the method of O'Brien et al (1994).

### *Secondary Analysis with Crossovers Excluded*

Some patients discontinued the assigned study medication because of lack of efficacy or adverse effects and received open treatment with clozapine, haloperidol, or other standard medications, as clinically indicated. Since these crossovers could affect study comparisons, the groups were also compared with crossover cases excluded, that is, with all data from patients who crossed over at any time during the trial excluded from the analysis. This analysis allows evaluation of clozapine's cost-effectiveness among patients who fully comply with prescribed treatment (an upper-bound analysis).

## RESULTS

### *Sample and Treatment*

Comparison of patients randomized to clozapine ( $n = 205$ ) and to haloperidol ( $n = 218$ ) on sociodemographic and clinical characteristics showed that the randomization procedure successfully created groups that were well balanced at baseline (Rosenheck, Cramer, Xu, et al. 1997).

### *Effectiveness*

*Specific Domains.* Altogether, 82 percent of all assessments across all time points were completed. Random-effects regression analysis of subjects as

randomized (intention to treat analysis) showed that clozapine was significantly more effective than haloperidol on two single-domain measures: symptoms ( $p = .02$ ) and side effects ( $p < .0001$ ). Nonsignificant trends favored clozapine on the other four measures: family relationships ( $p = .23$ ); social relationships ( $p = .30$ ), daily activities ( $p = .20$ ), and community (role) functioning ( $p = .06$ ).

*The Composite Health Index for Schizophrenia.* Clozapine was also more effective than haloperidol on random regression analysis of the unweighted cumulative Composite Health Index ( $p = .009$ ) as well as on the patient-weighted version of this utilities index ( $p = .01$ ) (see illustration in Figure 2) and the provider-weighted version of the utilities index ( $p = .02$ ).

A correlation matrix was used to examine the intercorrelation of one-year cumulative improvement measures (i.e., the area under the improvement curve) across the six domains. Among the 15 ( $n*[n - 1]/2$ ) paired correlations, the average correlation coefficient across all six measures was only moderate strength (0.24, range 0.00–0.46), indicating that improvements in the six domains are relatively independent of each other.

Figure 2: Composite Health Index for Schizophrenia: Patient Weights. Comparison of Patients Assigned to Clozapine and Haloperidol on the Composite Health Index for Schizophrenia (Intention to Treat) Weighted by Patient Preferences ( $p < .01$ )

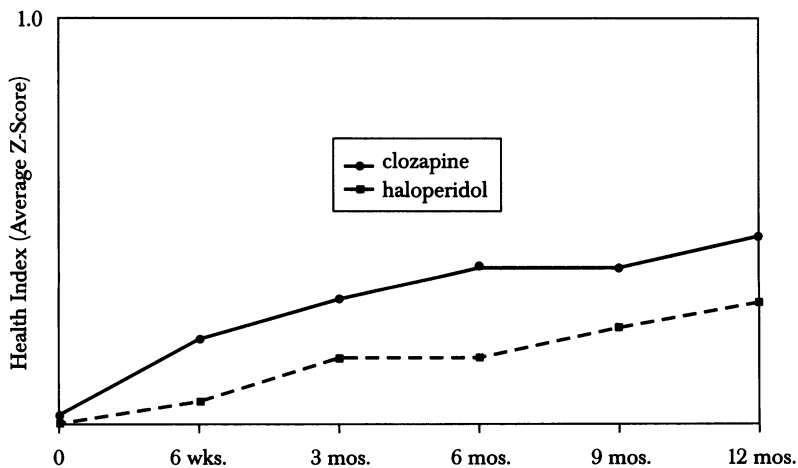


Table 2 presents cumulative gain (area-under-the-curve) comparisons for the intention to treat analysis for the Composite Health Index (i.e., one unit = an average of one standard deviation ( $z$ -score) improvement across all six measures for a full year). Clozapine-haloperidol comparisons are summarized in the seventh column to the right, showing the percentage difference in improvement between groups. On the unweighted measure of composite effectiveness (upper panel in Table 2), the clozapine group showed 49 percent greater improvement than the haloperidol group during the first six months, but 43 percent less improvement beyond that achieved during the first six months, during the second six months. This reversal is attributable to crossover patients who switched from clozapine to standard antipsychotic medication, or from haloperidol to clozapine. Relative to baseline levels, the clozapine group still showed 33 percent greater improvement than the haloperidol group in composite improvement during the second six months (see third row in Table 2) and a 37 percent greater improvement across the entire year. Differences were statistically significant for all time periods.

Results using patient preference weights are similar to those using the unweighted measures during the first six months of treatment (middle panel in Table 2). Gains over baseline for the clozapine group as compared to the haloperidol group, during the second six months and over the entire year, were somewhat greater with patient preference weights than with unweighted measures.

Results for provider-weighted measures also favored the clozapine group and were virtually identical to the results using the other two weighting schemes for the first six months. Outcomes with provider-weighted measures fell between the results using the other two measures, for the remaining time periods (Table 2).

### *Secondary Analysis Excluding Crossover Patients*

Altogether 83 (40 percent) of the clozapine patients switched to standard antipsychotic medication (including haloperidol) during the follow-up period, and 49 (22 percent) of the haloperidol patients received clozapine. Secondary analyses excluded all data from these crossover cases.

With crossover cases excluded, results for the first six months were similar for all three weighting schemes (Table 3) and were virtually the same as in the intention to treat analysis. Results for the other time periods, as expected, showed greater advantages for the clozapine group than in the intention to treat analysis. With crossovers excluded, both incremental improvement over

Table 2: Cumulative Gain in Composite Health Index for Schizophrenia—Intention-to-Treat Sample ( $N = 423$ ) (Area-Under-the-Curve Analysis)

|                         | <i>Clozapine</i> | <i>Haloperidol</i> | <i>Difference (Cloz-Hal)</i> | <i>Percent Difference</i> | <i>t</i> | <i>df</i> | <i>p</i> |
|-------------------------|------------------|--------------------|------------------------------|---------------------------|----------|-----------|----------|
| <i>Unweighted</i>       |                  |                    |                              |                           |          |           |          |
| 0–6 months              | 0.10             | 0.05               | 0.05                         | 49                        | 12.42    | 406.2     | .0001    |
| 6–12 months*            | 0.02             | 0.03               | –0.01                        | –43                       | 2.39     | 406.2     | .0169    |
| 6–12 months†            | 0.18             | 0.12               | 0.06                         | 33                        | 5.83     | 406.2     | .0001    |
| 0–12 months             | 0.27             | 0.17               | 0.10                         | 37                        | 7.48     | 406.2     | .0001    |
| <i>Patient Weights</i>  |                  |                    |                              |                           |          |           |          |
| 0–6 months              | 0.09             | 0.04               | 0.04                         | 50                        | 11.23    | 395.0     | .0001    |
| 6–12 months*            | 0.02             | 0.02               | –0.00                        | –5                        | 0.31     | 395.0     | .7512    |
| 6–12 months†            | 0.17             | 0.10               | 0.07                         | 41                        | 6.10     | 395.0     | .0001    |
| 0–12 months             | 0.26             | 0.14               | 0.11                         | 44                        | 7.38     | 395.0     | .0001    |
| <i>Provider Weights</i> |                  |                    |                              |                           |          |           |          |
| 0–6 months              | 0.09             | 0.04               | 0.04                         | 49                        | 11.16    | 395.2     | .0001    |
| 6–12 months*            | 0.02             | 0.03               | –0.01                        | –31                       | 1.53     | 395.2     | .1262    |
| 6–12 months†            | 0.17             | 0.11               | 0.06                         | 37                        | 5.52     | 395.2     | .0001    |
| 0–12 months             | 0.26             | 0.15               | 0.11                         | 41                        | 6.93     | 395.2     | .0001    |

*Note:* Score represents average improvement in Composite Health Index (e.g., a score of 1.0 would represent an average improvement of one standard deviation on all six outcomes for an entire year).

\* Incremental gain over six-month level.

† Sustained gain over baseline level (time 0).

the first six months and improvement over baseline levels during the second six months were greater for the clozapine group. These results favored clozapine and were statistically significant ( $p < .0001$ ). Cumulative effectiveness results for the full year show consistently greater cumulative benefit for the group assigned to clozapine. The clozapine advantage was greatest with the patient preference weights.

### *Scaling from Worst Health to Good Health*

Although the benefits of clozapine are statistically significant, data presented thus far do not allow evaluation of their magnitude on a more meaningful scale. Table 4 presents group differences from the middle panels of Tables 2 and 3, but with the metric converted to a 0–1 scale representing worst health–good health units (analogous to QALYs), as described earlier. Intention to treat analysis shows a one-year gain of .049 worst health–good health units

Table 3: Cumulative Gain in Composite Health Index for Schizophrenia—Crossovers-excluded Sample ( $N = 423$ ) (Area-Under-the-Curve Analysis)

|                         | <i>Clozapine</i> | <i>Haloperidol</i> | <i>Difference (Cloz-Hal)</i> | <i>Percent Difference</i> | <i>t</i> | <i>df</i> | <i>p</i> |
|-------------------------|------------------|--------------------|------------------------------|---------------------------|----------|-----------|----------|
| <i>Unweighted</i>       |                  |                    |                              |                           |          |           |          |
| 0-6 months              | 0.10             | 0.05               | 0.05                         | 49                        | 12.76    | 214.2     | .0001    |
| 6-12 months*            | 0.03             | 0.02               | 0.01                         | 37                        | 2.67     | 214.2     | .008     |
| 6-12 months†            | 0.21             | 0.11               | 0.10                         | 47                        | 8.28     | 214.2     | .0001    |
| 0-12 months             | 0.32             | 0.17               | 0.15                         | 47                        | 9.40     | 214.2     | .0001    |
| <i>Patient Weights</i>  |                  |                    |                              |                           |          |           |          |
| 0-6 months              | 0.09             | 0.05               | 0.04                         | 48                        | 9.68     | 190.5     | .0001    |
| 6-12 months*            | 0.03             | 0.01               | 0.02                         | 70                        | 4.51     | 190.5     | .0001    |
| 6-12 months†            | 0.19             | 0.09               | 0.10                         | 53                        | 7.38     | 190.5     | .0001    |
| 0-12 months             | 0.28             | 0.14               | 0.14                         | 51                        | 7.95     | 190.5     | .0001    |
| <i>Provider Weights</i> |                  |                    |                              |                           |          |           |          |
| 0-6 months              | 0.09             | 0.05               | 0.04                         | 47                        | 9.60     | 191.4     | .0001    |
| 6-12 months*            | 0.03             | 0.01               | 0.02                         | 60                        | 3.62     | 191.4     | .0001    |
| 6-12 months†            | 0.19             | 0.10               | 0.09                         | 50                        | 6.95     | 191.4     | .0001    |
| 0-12 months             | 0.28             | 0.14               | 0.14                         | 49                        | 7.61     | 191.4     | .0001    |

Note: Score represents average improvement in Composite Health Index (e.g., a score of 1.0 would represent an average improvement of one standard deviation on all six outcomes for an entire year).

\* Incremental gain over six-month level.

† Sustained gain over baseline level (time 0).

for the clozapine group versus .027 units for the haloperidol group—a small incremental gain of .021 units. With crossovers excluded, worst health-good health improvement increases to .053 for the clozapine group versus .026 for haloperidol, a greater (but still small) incremental gain of .027 units favoring clozapine.

*Societal Costs*

Table 5 shows that total societal costs were high (approximately \$60,000 per year) for both groups, but that one-year costs for the clozapine group were \$2,773 (4.5 percent) lower than costs for the haloperidol group ( $p = .41$ ). The difference between the groups increased from \$450 (1.3 percent) during the first six months to \$2,283 (9.2 percent) in the second six months.

Looking at specific cost components (Table 5), outpatient treatment costs were substantially greater for clozapine than for haloperidol patients (\$5,000

Table 4: Comparison of Gains in 0–1 Worst Health–Good Health Units (QALY Analogue Measure)

|              | <i>Intention to Treat</i> |                    |                              | <i>Crossovers Excluded</i> |                    |                              |
|--------------|---------------------------|--------------------|------------------------------|----------------------------|--------------------|------------------------------|
|              | <i>Clozapin</i>           | <i>Haloperidol</i> | <i>Difference (Cloz-Hal)</i> | <i>Clozapine</i>           | <i>Haloperidol</i> | <i>Difference (Cloz-Hal)</i> |
| 0–6 months   | 0.016                     | 0.008              | 0.008                        | 0.017                      | 0.009              | 0.008                        |
| 6–12 months* | 0.004                     | 0.004              | –0.000                       | 0.005                      | 0.002              | 0.004                        |
| 6–12 months† | 0.032                     | 0.019              | 0.013                        | 0.036                      | 0.017              | 0.019                        |
| 0–12 months  | 0.049                     | 0.027              | 0.021                        | 0.053                      | 0.026              | 0.027                        |

*Note:* Incremental gain in Worst Health–Good Health Units, analogous to QALYs, were estimated by dividing cumulative Composite Health Index gain weighted by patient preference, by the full worst health–good health range of 5.27 and standard deviation units (see Table 1).

\* Incremental gain over six-month level.

† Sustained gain over baseline level (time 0).

or 144 percent), and the difference increased from the first to the second six-month period. However, clozapine was associated with 21 fewer psychiatric hospital days than haloperidol ( $p = .02$ ) resulting in a total inpatient savings of \$7,440. Differences in inpatient costs were substantially greater during the second six months of the trial (15 fewer days with \$5,015 lower costs) than during the first six months (six fewer days with \$2,424 lower costs). Clozapine treatment resulted in a total reduction of \$8,684 (16 percent) in all total psychiatric hospital costs ( $p = .01$ ).

Overall, the greater costs of medication and outpatient treatment for clozapine patients were offset by inpatient savings during both six-month periods, but the magnitude of the offset was not sufficient to result in statistically significant total cost savings.

*Cost Analyses with Crossovers Excluded.* Results of secondary cost analyses excluding crossovers were similar to intention to treat analyses in that no significant differences occurred between the groups in total healthcare cost or total societal costs. With crossover cases excluded, however, total societal costs were \$3,295 greater for clozapine patients than for the controls (\$60,028 versus \$56,733;  $p = .41$ ), primarily because they had continued on the more costly outpatient pharmacotherapy regime for a longer period of time. While inpatient costs were \$5,060 lower in the clozapine group (\$47,835 versus \$52,895;  $p = .23$ ), outpatient and medication costs were \$8,495 greater (\$11,779 versus \$3,284;  $p = .0001$ ).

Table 5: One-Year Cost Comparison of Patients Assigned to Clozapine (205) and Haloperidol (218)

| Cost Category                   | 1-6 months |          |       | 6-12 months |           |          | Total |        |           |          |       |        |
|---------------------------------|------------|----------|-------|-------------|-----------|----------|-------|--------|-----------|----------|-------|--------|
|                                 | Clozapine  | Halldol  | t     | p           | Clozapine | Halldol  | t     | p      | Clozapine | Halldol  | t     | p      |
| <i>Utilization</i>              |            |          |       |             |           |          |       |        |           |          |       |        |
| Inpatient/Residential Care Days | 100.7      | 106.2    | 0.96  | .34         | 58.0      | 73.6     | 2.3   | .02    | 158.7     | 179.8    | 1.8   | .07    |
| Psychiatric                     | 93.8       | 101.9    | 1.44  | .15         | 50.0      | 66.2     | 2.5   | .01    | 143.8     | 168.1    | 2.2   | .03    |
| Medical/Surgical/Other          | 6.9        | 4.3      | 1.24  | .21         | 8.0       | 7.4      | 0.2   | .84    | 14.9      | 11.7     | 0.7   | .50    |
| Outpatient Services†            | 48.8       | 39.5     | 1.39  | .17         | 84.8      | 58.4     | 2.5   | .01    | 133.6     | 97.9     | 2.2   | .03    |
| <i>Total Cost to Society</i>    | \$35,614   | \$36,064 | 0.26  | .80         | \$22,537  | \$24,820 | 1.18  | .24    | \$58,151  | \$60,884 | 0.83  | .409   |
| Health Care Costs               | 35,404     | 35,779   | 0.22  | .83         | 22,380    | 24,446   | 1.07  | .29    | 57,785    | 60,225   | 0.74  | .395   |
| Inpatient/Residential Care      | 31,848     | 34,273   | 1.34  | .18         | 17,463    | 22,479   | 2.45  | .01    | 49,311    | 56,751   | 2.14  | .033   |
| Psychiatric                     | 29,510     | 33,011   | 2.01  | .04         | 15,736    | 20,920   | 2.60  | .01    | 45,247    | 53,931   | 2.59  | .010   |
| Medical/Surgical                | 2,337      | 1,262    | 1.98  | .05         | 1,727     | 1,559    | 0.28  | .78    | 4,064     | 2,821    | 1.28  | .198   |
| Outpatient Care                 | 3,557      | 1,507    | 8.90  | <.0001      | 4,917     | 1,967    | 8.13  | <.0001 | 8,473     | 3,474    | 9.17  | <.0001 |
| Patient care†                   | 1,999      | 1,314    | 3.25  | .001        | 3,275     | 1,793    | 4.61  | <.0001 | 5,274     | 3,107    | 4.50  | <.0001 |
| Anti-psychotic Meds             | 1,558      | 193      | 20.12 | <.0001      | 1,642     | 174      | 14.40 | <.0001 | 3,199     | 367      | 17.51 | <.0001 |
| Non-Health Care Costs           | 210        | 285      | 1.39  | .16         | 157       | 374      | 2.45  | .02    | 367       | 659      | 2.68  | .008   |
| Lost productivity               | -35        | -2       | 1.64  | .06         | -48       | -4       | 1.72  | .10    | -83       | -5       | 2.27  | .024   |
| Transfer payments‡              | 133        | 135      | 0.17  | .86         | 132       | 118      | 1.13  | .26    | 265       | 253      | 0.52  | .605   |
| Criminal justice                | 69         | 73       | 0.10  | .92         | 36        | 185      | 1.89  | .06    | 106       | 258      | 1.72  | .080   |
| Family burden (lost income)     | 42         | 79       | 1.17  | .24         | 36        | 74       | 0.97  | .32    | 78        | 153      | 1.37  | .162   |

\* Outpatient services include professional services and laboratory visits including those required by the study protocol.

† Costs for the haloperidol group have been adjusted for services exclusively required for the study protocol to 73 percent of their actual value.

‡ Only administrative costs of transfer payments are included (see Gold et al. 1996).

*Summary Measures and Cost-Effectiveness (CE) Ratios*

Summary measures of cost-effectiveness were calculated in two ways. First, we examined the 95 percent confidence intervals of the measures of effectiveness and cost followed by the 95 percent confidence interval of the combined incremental CE ratio. The incremental effectiveness of clozapine over haloperidol was small in magnitude (.022 worst health–good health units over one year), with tight confidence intervals. Differential effectiveness was greater during the second six months than in the first six months (.013 worst health–good health units versus .008 units), and was greater with crossovers excluded. The clozapine-haloperidol cost differences, as noted previously, showed lower costs for clozapine in the intention to treat analysis, but higher costs with crossovers excluded—with wide 95 percent confidence intervals (Table 6).

The negative CE ratios observed in the intention to treat analyses are presented here only to illustrate the wide calculated confidence intervals. They are not substantively interpretable, since large negative ratios could reflect either greater cost savings for clozapine (a desirable objective) or smaller effectiveness (an undesirable objective) (Weinstein 1996). The positive CE ratios observed on analysis with crossovers excluded give a less desirable picture of clozapine's cost-effectiveness, although the CE ratio for the 6–12 month interval is more hopeful at \$7,143/worst health–good health unit. One could hope that this result for patients who stay on clozapine would be sustained beyond the 12-month duration of this study, but a longer trial would be needed to verify this possibility.

The final analysis of the joint uncertainty about the CE ratio revealed substantial uncertainty, with the full range of the 95 percent confidence interval (from the lower limit to the upper limit) coming to \$608,000 per worst health–good health unit (the QALY analogue) in the one-year intention to treat analysis, and \$287,000 per worst health–good health unit with crossovers excluded. Again, these calculations are presented only to illustrate the extreme width of the 95 percent confidence intervals for one unit change in our QALY analogue measure, since values in the negative range cannot be unambiguously interpreted.

## DISCUSSION

Schizophrenia is a chronic, relapsing illness that adversely affects multiple domains of life and is associated with the consumption of substantial societal



Table 6: Incremental Effectiveness (in Worst Health-Good Health Units) of Clozapine vs. Haloperidol; Incremental Cost; and Incremental Cost-Effectiveness Ratios (Cost/Worst Health-Good Health Unit) with 95% C.I.

|                            | Incremental Effectiveness |       | 95% C.I. |       | Incremental Cost | 95% C.I.  |    | CE Ratio | 95% C.I.    |             |   |           |
|----------------------------|---------------------------|-------|----------|-------|------------------|-----------|----|----------|-------------|-------------|---|-----------|
|                            | LL                        | UL    | LL       | UL    |                  | LL        | UL |          | LL          | UL          |   |           |
| <i>Intention to Treat</i>  |                           |       |          |       |                  |           |    |          |             |             |   |           |
| 0-6 months                 | 0.008                     | 0.007 | -        | 0.010 | (\$450)          | (\$3,878) | -  | \$2,978  | (\$55,556)  | (\$478,688) | - | \$367,576 |
| 6-12 months*               | 0.013                     | 0.009 | -        | 0.017 | (\$2,283)        | (\$6,099) | -  | \$1,494  | (\$172,955) | (\$466,953) | - | \$120,892 |
| 0-12 months                | 0.022                     | 0.016 | -        | 0.027 | (\$2,733)        | (\$9,220) | -  | \$3,754  | (\$127,116) | (\$431,585) | - | \$177,352 |
| <i>Crossovers Excluded</i> |                           |       |          |       |                  |           |    |          |             |             |   |           |
| 0-6 months                 | 0.008                     | 0.007 | -        | 0.010 | \$3,161          | (\$1,005) | -  | \$7,327  | \$385,488   | (\$118,018) | - | \$888,993 |
| 6-12 months*               | 0.019                     | 0.014 | -        | 0.024 | \$135            | (\$4,359) | -  | \$4,630  | \$7,143     | (\$230,254) | - | \$244,540 |
| 0-12 months                | 0.027                     | 0.022 | -        | 0.032 | \$3,295          | (\$4,449) | -  | \$11,041 | \$121,587   | (\$165,745) | - | \$121,624 |

Note: C.I.s calculated using method of O'Brien et al. 1994).

\* Sustained gain over baseline level.

resources. The evaluation of treatments for schizophrenia requires an assessment of diverse outcomes, including societal costs, over extensive periods of time. Although treatment outcome studies of severe and persistent mental illness typically collect assessment data from multiple domains, methods for synthesizing the resulting information have yet to be developed. In this study we present an approach that addresses six methodological challenges: (1) imperfect protocol compliance (crossovers); (2) integration of data from multiple outcome domains; (3) weighting of outcomes for preferences; (4) assessment of cumulative benefits; (5) conversion of effect-size scores to a worst health-good health measure analogous to the conventional Quality Adjusted Life Year; and (6) estimation of cost-effectiveness ratios and their uncertainty. Before addressing the substantive conclusions of this study we will review critically our approach in each of the areas.

### *Crossovers*

It is inevitable that in a long-term trial, as in actual clinical practice, not all patients will adhere faithfully to the treatment protocol. Recognition has been growing in recent years about the need to differentiate the valuation of treatment efficacy in clinical trials from the evaluation of treatment effectiveness (Detsky 1996). Whereas efficacy studies assess the specific effects of treatment under controlled conditions, effectiveness studies assess the effects of a treatment under conditions that more closely approximate circumstances in the "real world." Because current expert consensus suggests that patients should continue with clozapine treatment for a full year to maximize their chance for delayed benefits (Meltzer 1992), our study was designed to treat all patients with the assigned drug for 12 months. In actual practice, however, patients take medications for variable amounts of time, depending on their physician's assessment of their progress and on their own experience of pharmacologic benefit. Although the main reason for giving primacy to the intention to treat analysis is that it provides the most unbiased evaluation of drug efficacy, it also provides insight into the effectiveness of medication under less controlled conditions, incorporating the consequences of the fact that only a subset of clozapine patients actually continue treatment for a full year. Although it is more vulnerable to selection biases, the analysis with crossovers excluded presents a clearer picture of drug efficacy because it approximates more closely a pure comparison of continuous treatment with the assigned medications for a full year. This analysis, however, is less representative of typical practice.

The greater cost of treatment for clozapine patients in the analysis with crossovers excluded is especially notable and suggests that, although benefits are greatest when treatment is provided for a full year, cost savings with clozapine may be more limited under such circumstances. Together these two analytic approaches provide a "binocular" perspective on cost-effectiveness and suggest that the optimal treatment strategy may be to start persons with refractory schizophrenia on clozapine, but to use clozapine for a limited period of time, only continuing long-term prescription if the clozapine regimen passes a preset threshold for effectiveness (Rosenheck, Evans, Herz, et al. under review). It is important to note, as well, that cost data from analyses of both intention to treat and crossovers excluded showed no statistically significant cost differences between treatment groups.

*A Single Measure of Effectiveness.* Although a guiding principle of controlled trials is to focus outcome assessment on a single primary outcome measure, this principle may be overly restrictive in studies of chronic mental illness in which several disease-specific outcomes are important. In such illnesses no single outcome or utility measure can adequately represent clinical status, dysphoric side effects, and various aspects of quality of life. Such studies must use multiple specialized measures.

The primary advantage of a single outcome measure, such as the average  $z$ -score used here, is that it is simple to generate from obtainable data and can be combined with cost data to generate a cost-effectiveness ratio (Hargreaves et al. 1998). Its major drawback is that the resulting outcome measure is not clinically meaningful and cannot be related directly to patient utility without undergoing additional validating empirical study. Although Cohen (1969) has suggested conventions for interpreting effect sizes measured by standardized ( $z$ ) scores (classifying changes of 0.2, 0.5, and 0.8 as small, medium, and large effects, respectively), this metric also lacks either direct clinical meaning or a valid utility interpretation.

*Preference Weights.* Recognizing that some outcomes are likely to be more important than others, we modified the average effect size, or  $z$ -score, by weighting the six components of the Composite Health Index on the basis of preferences elicited by Kleinman from patients and providers using a categorical rating scale questionnaire (1995). Although the relative effectiveness of clozapine was greater using the patient preference weights, preference weights had little effect on the overall results.

Although the application of preference ratings in this study is the first published effort to address differential outcome preferences in a

cost-effectiveness study of psychotic illness, the study suffers from several potential methodological limitations.

First, Kleinman's test sample was not drawn from the current study sample and is not broadly representative of schizophrenic patients in general. In contrast to patients in the VA study—in which all met the criteria for refractoriness, 2 percent were female, and 29 percent African American—the patients in Kleinman's sample were not necessarily refractory to treatment (although all were severely ill); 60 percent were female and 87.5 percent were African American. Kaplan (1996), however, has presented evidence that the given consistent instrumentation preference ratings are generally consistent across samples and between patients and informed nonpatients.

Second, by simply summing the weighted  $z$ -scores, we have opted, for the sake of simplicity, for an additive model of improvement in which the importance of outcomes in any given domain are assumed not to be influenced by the outcomes in other domains. Although the magnitude of the correlations of outcome among the six domains was generally modest (averaging .24 across all domains), the outcomes in most domains were significantly correlated with outcomes in others. Our weighted average thus may include some degree of "double counting."

Finally, it is important to acknowledge that we have not demonstrated the interval properties of our Composite Health Index for Schizophrenia (i.e., increments of change have the same meaning on all segments of the scale). Following the law of diminishing marginal utility (i.e., the more of a commodity a person consumes, the less utility or satisfaction he or she obtains for the last increment of that commodity), it seems likely that the utility realized by an improvement in the Composite Health Index from a score of 0.0 to 1.0 would be greater than the utility realized by improvement from 1.0 to 2.0 (Gold et al. 1996). To address this limitation we conducted a sensitivity analysis in which we differentially weighted improvement measures by baseline status. Improvement for patients who were in perfect health at baseline was given a weight of zero (since they would have been minimally ill to begin with), and improvement for those in the poorest health was weighted by a factor of three, with linear increases in the weighting between these extremes. This adjustment did not substantially alter the findings. Patients in the clozapine group showed a gain of .091 worst health–good health units, versus .055 for those in the comparison group.

*Cumulative Gains.* Because schizophrenia has a highly variable course, we quantified the cumulative effect of treatment by calculating the area under the improvement curve, not just its slope. This allows consideration of the

timing of improvement in addition to its final magnitude, taking into account, for example, that a patient who gains symptom relief during the first month of treatment realizes greater benefit than one who gains the same symptom relief six months later.

*Worst Health–Good Health Scale.* The major limitation of our Composite Health Index for Schizophrenia, as noted earlier, is that the resulting metric is not expressed in clinically meaningful units. To address this problem we transformed the average standard score measure into a 0–1 worst health–good health scale analogous to the Quality Adjusted Life Year (QALY) units used in conventional utility measurement. The basis for this transformation was a calculation of the magnitude of the full range of the average effect sizes across all six outcome measures, with a 20 percent reduction to allow for states worse than death and an upper range for good health. In developing this method, we reasoned that being in a completely psychotic state, consumed by hallucinations and delusions, without any capacity for clear thought, social relatedness, or community functioning, and suffering severe side effects and involuntary movements, as one would if one scored at the bottom of all six assessment scales—is, by convention, a state worse than death (one that is rarely encountered in practice). It is also reasonable to think that being entirely without either symptoms of schizophrenia or medication side effects, but with some dissatisfactions in social relationships and employment performance, approximates good health. The anchors at the extremities of the QALY analogue measure thus have face value.

The average score on this scale at the time of study enrollment was .47 on the 0–1 death–perfect health scale, just below the level of .56 identified by Revicki, Shakespeare, and Kind (1996) for hospitalized schizophrenic patients using a standard gamble methodology with physicians as key informants. Since by definition all of the patients in this study were treatment refractory, we would expect their health status to be lower than that of the standard hospitalized patients in Revicki's vignettes. In addition, the degree of improvement observed for clozapine patients represents a small improvement using Cohen's (1969) effect-size conventions (effect size for clozapine compared to haloperidol = 0.14 (s.d.) for patient weights with crossovers excluded; see Table 3, eighth row, sixth column). This improvement can be plausibly equated with the .027 improvement in worst health–good health units (Table 4, fourth row, sixth column). In the absence of a "gold standard" rating of schizophrenic states, our Composite Health Index for Schizophrenia has face value and is consistent with data from the single published study that presents standard gamble ratings for schizophrenia (Revicki, Shakespeare,

and Kind 1996), and with published conventions for evaluating effect sizes (Cohen 1969). Thus, the method illustrated here correctly shows clozapine's effect to be small in magnitude and provides a practical and accessible approximation of a measure of QALYs based on disease-specific outcome measures.

*Cost-Effectiveness (CE) Ratios and Uncertainty Estimates.* Finally, we combined effectiveness and cost data into a single measure of cost-effectiveness and estimated the uncertainty of (1) our effectiveness measures, (2) our cost estimates, and (3) the cost-effectiveness ratio. These analyses demonstrated the statistical reliability of our effectiveness estimates and the substantial uncertainty of our cost estimates. These large uncertainty estimates reflect both the characteristically large standard deviations of cost data, and the conservative evaluation of extreme combinations of high cost and low effects (and vice versa) in the O'Brien et al. (1994) approach to assessing confidence intervals for CE ratios.

## SUBSTANTIVE CONCLUSION: PHARMACOECONOMIC EVALUATION OF CLOZAPINE

This study demonstrates that (a) clozapine is somewhat more effective than conventional antipsychotic medications in the treatment of refractory schizophrenia; and that (b) although it is more expensive, its greater cost is offset by reductions in hospital utilization, at least among high hospital users in the VA system.

The patients involved in this study were all treated in the VA system, and the generalizability of our findings to other healthcare systems is somewhat uncertain. Patients treated for schizophrenia in the VA system are typically older, poorer, more disabled, and more exclusively male than those treated elsewhere, and VA lengths of stay are longer than the typical LOS in private sector hospitals. Our findings are likely to be generally applicable, however, to severely disabled, high hospital users treated in other public mental health systems, a relatively small but extremely costly and troubled population. In this population the clinical gains with clozapine are highly significant but small in magnitude, while cost savings are somewhat more substantial, but not statistically significant.

This study shows that statistical significance does not automatically equal clinical significance. We found highly statistically significant differences

favoring the effectiveness of clozapine, but on average they were of small magnitude, as was most clearly illuminated by the transformation of the outcome measures into worst health–good health units analogous to QALYs. Cost data showed average savings of almost \$3,000 per patient per year, although these savings were not statistically significant. Clozapine thus appears, on average, to be a cost-neutral treatment in hospitalized patients with a small margin of clinical effectiveness.

## ACKNOWLEDGMENTS

Members of the Cooperative Study Group on Clozapine are John Grabowski, M.D., Detroit, MI; Denise Evans, M.D., Augusta, GA; Lawrence Herz, M.D., Bedford, MA; George Jurjus, M.D., Brecksville, OH; Sidney Chang, M.D., Brockton, MA; Lawrence Dunn, M.D., Durham, NC; John C. Crayton, M.D., Hines, IL; William B. Lawson, M.D., Ph.D., Little Rock, AR; Yeon Choe, M.D., Lyons, NJ; Richard Douyon, M.D., Miami, FL; Edward Allen, M.D., Montrose, NY; John Lauriello, M.D., Palo Alto, CA; Michael Peszke, M.D., Perry Point, MD; Jeffrey L. Peters, M.D., Pittsburgh, PA; Janet Tekell, M.D., San Antonio, TX; and Joseph Erdos, M.D., Ph.D., West Haven, CT. We would like to thank Linda Frisman, Ph.D., Don Hedeker, Ph.D., Leah Klineman, Ph.D., David Paltiel, Ph.D., John Rizzo, Ph.D., Douglas Leslie, and Charlotte Hitchcock for their advice and suggestions; and Lois Ucas, Jennifer Cahill, and Dennis Thompson of the Chairman's Office. We are indebted to the Data Monitoring Board (Alan Breier, M.D., Howard Goldman, M.D., James Klett, Ph.D., David Pickar, M.D.); the Executive Committee (Boris Astrachan, M.D., John Crayton, M.D., Linda Frisman, Ph.D., Carol Fye, R.Ph., M.S., William Hargreaves, Ph.D., and William Lawson, M.D.) for their careful overview of the progress of the trial.

## REFERENCES

- Alvir, J. M., J. A. Lieberman, A. Z. Safferman, J. L. Schwimmer, and J. Schaaf. 1993. "Clozapine-induced Agranulocytosis: Incidence and Risk Factors in the United States." *The New England Journal of Medicine* 329 (3): 162–67.
- Barnes, T. R. E. 1989. "A Rating Scale for Drug Induced Akathisia." *British Journal of Psychiatry* 131 (3): 222–23.
- Carpenter, W. T., R. R. Conley, R. W. Buchanan, A. Breier, and C. Tamminga. 1996. "Patient Response and Resource Management: Another View of Clozapine Treatment of Schizophrenia." *American Journal of Psychiatry* 152 (6): 827–32.

- Cohen, J. 1969. *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Detsky, A. S. 1996. "Evidence of Effectiveness." In *Valuing Health Care*, edited by F. A. Sloan. New York: Cambridge University Press.
- Essock, S. M., W. A. Hargreaves, N. H. Covell, and J. Goethe. 1996. "Clozapine's Effectiveness for Patients in State Hospitals: Results from a Randomized Trial." *Psychopharmacology Bulletin* 32 (4): 683-97.
- Gibbons, R. D., D. Hedeker, and J. Elkin. 1993. "Some Conceptual and Statistical Issues in Analysis of Longitudinal Psychiatric Data." *Archives of General Psychiatry* 50 (9): 739-50.
- Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein. 1996. *Cost Effectiveness in Health and Medicine*. New York: Oxford University Press.
- Guy, W. 1976. "Abnormal Involuntary Movements." In *ECDEU Assessment Manual for Psychopharmacology*, edited by W. Guy. Department of Health, Education, and Welfare, Pub. No. (ADM) 76-338. Rockville, MD: National Institute of Mental Health.
- Hargreaves, W. A., M. Shumway, T. W. Hu, and B. Cuffel. 1998. *Cost-Outcome Methods for Mental Health*. San Diego, CA: Academic Press.
- Heinrichs, D. W., E. T. Hanlon, and W. T. Carpenter. 1984. "The Quality of Life Scale: An Instrument for Rating the Schizophrenic Deficit Syndrome." *Schizophrenia Bulletin* 10 (3): 388-98.
- Kane, J. M., and S. R. Marder. 1993. "Psychopharmacologic Treatment of Schizophrenia." *Schizophrenia Bulletin* 19 (2): 287-302.
- Kane, J. M., G. Honigfeld, J. Singer, and H. Y. Meltzer, and the Clozaril Collaborative Study Group. 1988. "Clozapine for the Treatment-Resistant Schizophrenic: A Double Blind Comparison with Chlorpromazine." *Archives of General Psychiatry* 45 (9): 789-96.
- Kaplan, R. M. 1996. "Utility Assessment for Estimating Quality-Adjusted Life Years." In *Valuing Health Care*, edited by F. A. Sloan. New York: Cambridge University Press.
- Kay, S. R., A. Fiszbein, and D. R. Opler. 1987. "The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia." *Schizophrenia Bulletin* 13 (2): 261-76.
- Kleinman, L. S. 1995. "Preferences for Outpatient Mental Health Treatment." Ph.D. Thesis, Johns Hopkins University, School of Hygiene and Public Health.
- Lehman, A. F. 1988. "A Quality of Life Interview for the Chronically Mentally Ill." *Evaluation and Program Planning* 11 (1): 51-62.
- Meltzer, H. Y. 1992. "Treatment of the Neuroleptic-Nonresponsive Schizophrenic Patient." *Schizophrenia Bulletin* 18 (3): 515-42.
- Meltzer, H. Y., P. Cola, and L. Way. 1993. "Cost-Effectiveness of Clozapine in Neuroleptic Resistant Schizophrenia." *American Journal of Psychiatry* 150 (11): 1630-38.
- O'Brien, B. J., M. F. Drummond, R. Labelle, and A. Willan. 1994. "In Search of Power and Significance: Issues in the Design and Analysis of Stochastic Cost-Effectiveness Studies in Health Care." *Medical Care* 32 (2): 150-63.
- Office of the Inspector General. 1992. *Comparison of Costs and Outcomes of Matched Pairs*



- of VAMCs and Their University Affiliates. Washington, DC: Office of the Inspector General.
- Reid, W. H., M. Mason, and M. Toprac. 1994. "Savings in Hospital Bed-Days Related to Treatment with Clozapine." *Hospital and Community Psychiatry* 45 (3): 261-68.
- Revicki, D. A., A. Shakespeare, and P. Kind. 1996. "Preferences for Schizophrenia-related Health States: A Comparison of Patients, Caregivers and Psychiatrists." *International Clinical Psychopharmacology* 11 (2): 101-108.
- Rosenheck, R. A., D. Evans, L. Herz, J. A. Cramer, W. Xu, J. Thomas, W. Henderson, and D. Charney. In press. "How Long to Wait for a Response to Clozapine? A Comparison of Time Course of Response to Clozapine and Conventional Antipsychotic Medication in Refractory Schizophrenia." *Schizophrenia Bulletin*.
- Rosenheck, R. A., J. Cramer, W. Xu, J. Thomas, W. Henderson, L. K. Frisman, C. Fye, and D. Charney, for the Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. 1997. "A Comparison of Clozapine and Haloperidol in the Treatment of Hospitalized Patients with Refractory Schizophrenia." *The New England Journal of Medicine* 337 (12): 809-13.
- Rupp, A., and S. J. Keith. 1993. "The Costs of Schizophrenia." *Psychiatric Clinics of North America* 16 (2): 413-23.
- Simpson, G. M., and J. W. S. Angus. 1970. "A Rating Scale for Extrapyrarnidal Side Effects." *Acta Psychiatry Scandanavica* 212 (Supplement): 11-19.
- Spilker, B., ed. 1996. *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2d ed. Philadelphia, PA: Lippincott-Raven.
- Weinstein, M. C. 1996. "From Cost-Effectiveness Ratios to Resource Allocation: Where to Draw the Line." In *Valuing Health Care*, edited by F. A. Sloan. New York: Cambridge University Press.