Published in final edited form as:

Health Aff (Millwood). 2009 ; 28(5): w794–w808. doi:10.1377/hlthaff.28.5.w794.

Comparative Effectiveness Research For Antipsychotic Medications: How Much Is Enough?:

If comparative effectiveness research is to guide policy, then it must include evaluations of changes in policy

David O. Meltzer[Chief],

Section of Hospital Medicine, at the University of Chicago in Illinois

Anirban Basu[Assistant professor of medicine], and

Section of Hospital Medicine, at the University of Chicago in Illinois

Herbert Y. Meltzer[Professor of psychiatry]

Vanderbilt University in Nashville, Tennessee

David O. Meltzer: dmeltzer@medicine.bsd.uchicago.edu

Abstract

Second-generation antipsychotics have attracted practitioners' and policy-makers' attention, because of concerns over their health effects and costs. Comparative effectiveness data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE)—a high-profile National Institutes of Health (NIH)—funded study—have been used to argue for restricting coverage for these costly drugs. But concerns about the design of CATIE and its associated cost-effectiveness analysis and uncertainty about the precision of these findings raise questions about this interpretation. Our work suggests that additional research to increase the precision of comparisons of the effectiveness of antipsychotics would be well worth the cost.

Antipsychotic medications are the primary drug treatments for patients with schizophrenia because of the drugs' marked efficacy in the treatment of delusions and hallucinations. However, in most patients they have less ability to improve other symptoms in schizophrenia, including cognitive impairment.¹ First-generation antipsychotic medications were introduced in 1954 (chlorpromazine). A major issue especially with these older drugs is the risk of "extrapyramidal" symptoms, which include several serious movement disorders. The approval in 1990 for limited use of clozapine, the first antipsychotic drug to almost never produce extrapyramidal symptoms, initiated what is commonly called the second generation of antipsychotic drugs.²

The fundamental clinical difference between the first- and second-generation antipsychotics is the diminished risk of extrapyramidal symptoms in the latter.³ This paper examines the following set of second-generation antipsychotics: olan-zapine, quetiapine, and risperidone.⁴

Despite the lower rate of extrapyramidal symptoms with second-generation antipsychotics, the optimal choice of drug treatments for patients with schizophrenia has been disputed for several reasons: (1) uncertainties surrounding effectiveness in controlling psychotic symptoms, (2) the tendency of some second-generation drugs to produce weight gain and blood-lipid abnormalities, (3) some evidence that second-generation antipsychotics improve

cognition, and (4) the costliness of second-generation antipsychotics.⁵ Crucial questions have gone unanswered regarding the value of second-generation compared to first-generation antipsychotics.

The CATIE study

In 1999 the National Institute of Mental Health (NIMH) funded the \$42.6 million Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study to compare the effectiveness of one first-generation antipsychotic, perphenazine, and all second-generation antipsychotics available in the United States. CATIE's design as a randomized trial that compared these drugs head to head, its large size compared to prior studies, and federal sponsorship made its results highly anticipated.

The initial results from CATIE, published in 2005, reported that olanzapine but not the other second-generation antipsychotics (quetiapine, risperidone, and ziprasidone) was superior to perphenazine in having greater time to discontinuation of the study medication for any cause (primary outcome), and greater time to discontinuation for lack of efficacy (secondary outcome) (time to discontinuation has been suggested as an index of antipsychotic effectiveness, with longer discontinuation times considered to be evidence of superiority).⁶ Olanzapine was, however, associated with a shorter time to discontinuation for weight gain, hyper-glycemia, and other adverse effects, while perphenazine was associated with faster discontinuation because of extrapyramidal symptoms. The article concluded that perphenazine could not be rejected as an inferior treatment.

In 2006, a cost-effectiveness analysis (CEA) based on CATIE concluded that "treatment with perphenazine was less costly than treatment with second-generation antipsychotics with no significant differences in measures of effectiveness."⁷ In light of this analysis, some CATIE investigators have questioned the results of previous studies that favor second- over first-generation antipsychotics.⁸ Some have suggested that this discrepancy might reflect the influence of industry sponsorship of these trials, although recent meta-analyses of these studies have failed to support this hypothesis.⁹

Disagreement over implications of CATIE results

Regardless, investigators disagree about the implications that CATIE results should have for access to the second-generation antipsychotics. Several CATIE investigators, including the authors of the CATIE cost-effectiveness study, have stated that the study was not designed to directly answer questions of policy regarding access to antipsychotic drugs; others have suggested that the CATIE results establish that it is wasteful to use public funds to pay for second-generation antipsychotics, a perspective that has been adopted by some influential media outlets and pharmacy benefit managers (PBMs).¹⁰

Despite these differences in interpretation, there are no efforts under way to perform large-scale comparative effectiveness studies of second-generation anti-psychotics that would address CATIE's major limitations.¹¹

CATIE's usefulness for policy decisions

These issues raise questions about how comparative cost-effectiveness results should be designed, interpreted, and used to make public policy decisions. These questions are also timely because, since the introduction of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA), there have been greater incentives for comparative effectiveness research and CEA to play a larger role in U.S. public policy.¹² Federal funding for comparative effectiveness research was dramatically increased as part of the American Recovery and Reinvestment Act (ARRA) (the so-called stimulus bill) of 2009.

Value-of-information calculation

Here we focus on the design of CATIE and the precision of its findings that could influence its potential policy implications, such as coverage of second-generation antipsychotics and the value of additional public spending on comparative effectiveness research in this area.¹³

The main focus of our quantitative analysis is a “value-of-information” calculation that builds on CATIE data to calculate the expected value of research to further reduce the uncertainty about the costs and benefits of the first- and second-generation antipsychotics and implications for the design of future studies in this area. Although the focus is on CATIE, some of the issues that arise are clearly of more general relevance, including challenges in making coverage decisions in the presence of substantial uncertainty about costs and benefits, and the use of “value-of-information” techniques to assess the value of reducing such uncertainty and to identify optimal designs to carry out such research.

Design And Analysis Issues In CATIE

The relatively long duration (eighteen months) and number of patients (1,460) in CATIE have caused it to be described as the most comprehensive randomized clinical trial (RCT) ever conducted in patients with schizophrenia.¹⁴ However, CATIE’s design has also been criticized from a clinical perspective, and these and other concerns are relevant in considering its public policy implications.¹⁵

Time to discontinuation

Probably the most important issue with CATIE from a clinical perspective concerns its primary outcome measure: time to discontinuation of the study medication. In clinical practice for many conditions, time to discontinuation provides a measure that integrates information on both effects on symptoms and tolerability, and perhaps cost, from the perspectives of the patient and the prescriber. However, in the treatment of schizophrenia, decisionmakers might not have much data to assess when medication changes will benefit a given patient, which raises doubts about the selection of this measure as a primary outcome. Similarly, because patients or their families could choose different medications in CATIE at no cost, switching behavior would not reflect cost concerns.

Also, because patients who dropped out of CATIE might have had the cost of treatment fall on themselves or their family, and the prescriber would have lost income from study participation, CATIE may have produced incentives to switch medications more often than otherwise, to keep the patient in the study protocol—again raising concerns about the use of time to discontinuation as an outcome measure.

Inadequate evidence to justify departures from accepted treatments

Other important concerns about the clinical implications of CATIE include (1) assignment of 231 patients with tardive dyskinesia (a movement disorder sometimes caused by first-generation antipsychotics) to treatment with the second-generation antipsychotics but not to the first-generation antipsychotic (perphenazine); (2) failure to explore a range of fixed doses for the drugs; (3) a peak dose of one second-generation antipsychotic, olanzapine (30 mg per day) that is three times the optimally effective dose in non-treatment-resistant patients; (4) inclusion of up to one-third treatment-resistant patients, for whom the higher doses of olanzapine may have increased efficacy; and (4) failure to define and categorize the largest cause for discontinuation (“patient choice”), which need not imply lack of efficacy or intolerability.¹⁶ These concerns have led some to argue that the CATIE results are insufficient to justify what would be radical departures from accepted practice for the treatment of schizophrenia, as reflected in multiple published algorithms.¹⁷

Policy-relevant concerns

Several other concerns are especially relevant from a policy perspective. First, CATIE was designed to compare rates of discontinuation from the primary assigned medication, not coverage policies that restrict access to specific medications, which, unfortunately, is how its results have sometimes been used. Second, given the primary objective of studying discontinuation and the high frequency of switching in clinical practice because of the widespread belief that patients who do not respond to one antipsychotic drug may respond better to another, CATIE had to allow patients to switch medications whenever they and their clinicians chose. This implies that CATIE's findings apply only to the outcome of initial medication assignment if patients can readily switch and not if the assigned medication was the only one available or if coverage policies created barriers to switching. Indeed, switching occurred especially frequently in CATIE: more than 75 percent of patients on perphenazine changed medications over the eighteen-month course of the study, a rate far higher than in other effectiveness studies.¹⁸ This reinforces the notion that the effects of initial medication assignment in CATIE might not reflect those in routine clinical care.

For our analysis, another important issue concerning CATIE is that the outcomes needed for cost-effectiveness, including quality-adjusted life-years (QALYs) and costs, must be estimated. Exhibit 1 reports these cost-effectiveness results from CATIE.

Methodological concerns

Comparing health states—The methods used to develop these estimates raise several concerns. First, to compute QALYs, the authors had to estimate the “utilities” of the patients' health at all times. Such utilities assign values to various health states based on patients' preferences, to allow comparison of these health states. This was done by defining a series of health states described by psychotic symptoms and potential side effects of treatment with antipsychotic drugs, including tardive dyskinesia and weight gain. Psychotic symptoms were defined by severity scores based on the Positive and Negative Syndrome Scale (PANSS), which is the primary rating scale used to evaluate psychopathology in clinical trials in schizophrenia.

These symptoms and side-effect profiles were presented to members of the general public for rating using the “standard gamble technique,” a reasonably common and theoretically grounded, yet still controversial, approach for measuring preferences for health states.¹⁹ Whether this accurately reflects the welfare of patients experiencing these health states is impossible to know. Another concern arises because when psychotic symptoms and drug side effects occurred at the same time, the authors estimated utilities by multiplying utilities for the two separate health states—an approach that has been shown to underpredict utility of health states where conditions co-occur.²⁰ Bias for any of these reasons could alter the conclusions from the CATIE cost-effectiveness study, given the small differences in outcomes between treatment arms.

Length of treatment—Another issue is that the CATIE results are based on average outcomes over the eighteen-month duration of the study. A longer-term investigation could have resulted in different conclusions—for example, first-generation antipsychotics could have been found more likely than second-generation anti-psychotics to cause tardive dyskinesia and thus be discontinued.

By contrast, if patients in all treatment arms were to switch over time to treatments that were more effective for them, differences between treatment arms would decrease. Even the nature of switching behavior itself must be considered. For example, if one drug was thought to be more effective or more tolerable than another, switching would be less likely

with that drug, increasing its performance with respect to this primary outcome measure. These issues make generalizing about long-term outcomes from CATIE extremely complex and raise caution about drawing policy conclusions from CATIE.

Long-term value of research—The complexity of these issues also makes it challenging to accurately estimate the long-term value of research using the value-of-information approach we describe below. Some of these issues tend to make our estimates of the value of research too high (for example, the tendency of people to find the best treatments for themselves), while others (such as unclear long-term outcomes) tend to make estimates of the value of research too low.

Cost issues—On the cost side, the CATIE CEA found significant differences in total costs across arms, with costs associated with perphenazine less than with quetiapine, olanzapine, and risperidone (all $p < 0.001$).²¹

Two major methodological concerns preclude the acceptance of this conclusion. First, the adjusted differences in “raw-scale” costs (dollars) across treatments might not have been calculated correctly.²² Second, the finding of significant differences in costs between perphenazine and the atypical (second-generation) antipsychotic arms appears to be based on estimates from the transformed (log-scale) data, rather than the raw-scale cost data (in dollars).²³

Additionally, examination of the time trend in monthly costs estimates suggests that the conclusions of Robert Rosenheck and colleagues regarding the overall average monthly costs may be sensitive to the duration of the study.²⁴ Figure 3 in that paper reveals that monthly costs were rising almost continually for the perphenazine group from the eight-month mark until the end of the study, at least partially because of these patients’ switching to more expensive drugs, while the monthly costs in all of the other second-generation antipsychotic groups were declining slightly or stable. This suggests that had the study been continued longer, costs associated with perphenazine may have further approached or even exceeded those of the second-generation antipsychotics.

Evaluations of similar studies

Investigators examining the results of long-term studies such as CATIE and the U.K.-based Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS) study, which found that the overall effectiveness (in terms of QALYs) is similar between first- and second-generation antipsychotics, have noted the differences in findings from prior short-term trials that found evidence of benefits for the second-generation antipsychotics.²⁵ However, like CATIE, CUtLASS also has important design concerns that may explain why these studies find that overall effectiveness (in QALYs) is similar for first- and second-generation antipsychotics, while prior short-term trials found benefits for the second-generation antipsychotics.²⁶

Some of the reasons discussed for such discrepancies are that many of these prior studies were industry-funded, examined highly selected patient populations, and had high dropout rates. Another potential explanation is that these short-term studies are more likely to identify the effects of specific drugs than studies like CATIE that allow patients to switch medications, reducing the effect of initial assignment over time. Indeed, even patients started on placebo might have long-term outcomes similar to those of patients receiving an effective treatment if switching to effective treatments were allowed.

Using CATIE findings for coverage policy

All of these concerns suggest that there is great uncertainty about the extent to which CATIE findings should be used to make coverage policy. If existing data are not sufficient to address this need, it is useful to ask whether additional studies to address this question would be justified.

The Value Of Future Research

“Value-of-research” calculations have recently begun to be applied in both academic and policy settings (for example, by the U.K. National Institute for Health and Clinical Effectiveness [NICE]) to determine the value of additional studies on the value of treatments, including diverse clinical applications such as the treatment of Alzheimer’s disease and the routine removal of wisdom teeth.²⁷ The approach uses data on the uncertainty around the effects of treatment and the likelihood that new information would change the recommended treatment to calculate the expected value of the improvement in outcomes that would result from additional research.²⁸

To take a simple example, a study comparing a currently accepted treatment (A) to a potential new treatment (B) that had a 75 percent chance of showing that A was better than B, so that no change in practice was indicated, and a 25 percent chance of showing that B increased QALYs by four years, would have an expected value of one year per patient (0.25×4 years). This could then be multiplied by the size of the relevant population to estimate the population-level benefits. The approach can also incorporate effects on the costs of treatment or research, or both, to calculate measures of benefits net of costs. The value-of-information approach can estimate the expected benefits of a study that could eliminate all uncertainty surrounding the outcomes of a treatment decision (the expected value of perfect information) or the benefits of less powerful studies with a specific sample size or more limited outcome measures that incompletely characterize the uncertainty in outcomes.²⁹

Value of performing additional studies

We began our analysis by assessing the value of performing additional studies that would tell us exactly how treatment with typical and atypical antipsychotics would affect costs and outcomes. To do this, we used data from CATIE and other sources on the uncertainty in the lifetime survival, costs, and QALYs resulting from initial schizophrenia treatment with perphenazine, olanzapine, risperidone, and quetiapine to estimate the average changes in outcomes and costs expected from perfect knowledge of the effects of these atypicals. This expected value-of-research calculation was done assuming that without additional research, we accept the CATIE results and use perphenazine as the starting treatment. The value of future research, therefore, depends on the probability that this decision would change because of the additional research and the expected net benefits of improving that decision.³⁰

Expected value-of-research calculation

To perform this expected value-of-research calculation, we first used data from the literature on schizophrenia prevalence and mortality and data from CATIE on the effects of treatments on quality of life and costs to characterize the uncertainty in lifetime incremental costs (change in costs) and incremental benefits (change in QALYs) between perphenazine and each of the second-generation antipsychotics. We then used varying assumptions about the monetary value of a QALY to calculate the net monetary benefit (NMB) of choosing each second-generation antipsychotic over perphenazine, given these effects on QALYs and costs, by taking the monetary value of the gain in QALYs and subtracting the increase in costs. When the resulting NMB was positive, this provided an estimate of the value of

perfect information about that second-generation antipsychotic compared to perphenazine, given these effects on QALYs and costs, because the default treatment choice of perphenazine would have been incorrect.

When the NMB was zero or negative, the research was considered to have no value, given those effects on QALYs and costs, because it did not change the treatment decision. For each second-generation antipsychotic, we then calculated the expected value of perfect information by multiplying each level of positive NMB by its probability calculated based on the distribution of effects on costs and QALYs for that second-generation antipsychotic and then adding up over all possible values of positive NMB. The overall value of future research was calculated by taking the maximum value of perfect information over all second-generation antipsychotics.

Study Findings

Assuming that schizophrenia incidence occurs at age twenty, we estimated the incidence for schizophrenia to be 13 per 1,000 twenty-year-olds, producing 52,620 incident cases annually. Exhibit 2 shows simulated means and variances for each drug.³¹ The means and variances match well with the statistics reported in CATIE. The variances in QALYs may seem small when considered on a scale of 0 to 1, but even if a QALY is valued at only \$50,000, which is below most recent estimates, a difference of even 0.03 would be valued at \$1,500 annually, which is roughly similar in magnitude to the costs of even the more costly second-generation antipsychotics.³² The distribution of costs is also notable for having a relatively large proportion of observations that exceed the mean by a large margin (known as a “long right tail”), which is a common feature of many health spending distributions, including this one, because of costly hospitalizations.

CATIE’s effectiveness results versus ours

The CATIE results about the mean effects on outcomes across treatment arms imply that first-generation anti-psychotics are an effective first-line treatment. However, our analysis suggests that there is a 37 percent probability that this current decision based on the CATIE cost-effectiveness results will be wrong. Moreover, at \$50,000 per QALY, we estimate that the expected value of more precisely determining the effectiveness (ignoring costs) of atypical/typical antipsychotics in the United States is \$17.3 billion: \$11.7 billion accrues to patients who already have schizophrenia and the remainder to future cohorts.

When we also incorporate the costs of treatment, we estimate that the CATIE finding that initial assignment to typical antipsychotics is cost-effective has a 55 percent probability of being wrong, and the value of more precisely determining the cost-effectiveness of the initial assignment of antipsychotics in the United States increases to \$308 billion: \$207 billion accrues to the prevalent cohort of schizophrenia patients and \$6.6 billion, to each cohort of persons expected to develop schizophrenia each year over the next twenty years (Exhibit 3).³³

Variance of results with value of a QALY

Exhibit 4 illustrates how this expected value of future research and the probability that the CATIE findings are wrong (that is, that perphenazine is not the most cost-effective treatment) vary with the value of a QALY. It is notable that the expected value of research varies little even as the value of a QALY varies widely. This is because the major determinant of the value of research for these drugs is uncertainty about their effects on total costs.

Optimal Research Designs For Future Studies

Having determined that the potential value of perfect information concerning the comparative costs and effectiveness of first- and second-generation anti-psychotics is large, similar analyses can identify the optimal sample size for future studies that try to more precisely address the same questions that CATIE sought to address. As suggested earlier, CATIE was designed to focus on discontinuation of the assigned drug as a primary outcome; it was not powered to be able to show differences in cost-effectiveness. To explore this further, we performed a traditional deterministic power calculation to determine the number of subjects needed to identify a statistically significant effect on NMB based on the largest average effect size seen in CATIE between a second-generation antipsychotic (ziprasidone) and perphenazine. With estimated NMB of the two treatments of \$15,680 (standard deviation: \$315,000) and \$26,296 (SD: \$140,000), we found that a trial with 80 percent power to find statistically significant differences at the 5 percent level would require 8,300 patients in each arm. Similar calculations reveal that CATIE, with only 400 patients per arm, had only 10 percent power to detect differences at this level. Although these findings might be surprising to those who have put much faith in the CATIE CEA results, they are consistent with many previous findings that cost-effectiveness studies require larger sample sizes than clinical studies require, because of high variances in costs.

Limitations of power calculations

An important limitation of these power calculations is that they require assuming an effect size of a given magnitude, and they do not explicitly incorporate the value of the information obtained from a study of a given size. To address these concerns, we used the CATIE data to calculate the expected value of information for trials of varying sizes in each arm net of the estimated cost of performing these trials.³⁴

In making this estimate, we first carried out these calculations ignoring costs and only focusing on QALYs. We found that the expected value of information is maximized for a future trial of 4,000–4,500 subjects per arm. Such research is expected to produce a value of about \$13.8 billion, with health benefits valued at \$50,000 per QALY. Optimal sample sizes were not sensitive to the threshold value for a QALY.

Factoring in costs

When the effects of costs are taken into account, optimal sample sizes and the expected value they generate change considerably (Exhibit 5). First, the value of these trials of finite size is all much higher, and very close to the \$308 billion estimated value of perfect information noted above. Second, the optimal sample size is exceptionally large—more than 20,000 subjects per arm.

Discussion

This value-of-research analysis suggests that CATIE has important limitations for guiding public policy concerning coverage of antipsychotic medications. It also suggests that future research on this important class of medications is likely to be of immense value—exceeding \$300 billion—to people with schizophrenia today and those who will develop it over the next twenty years. Much more of the value of research appears to be generated by uncertainty about the effects of treatments on costs than by uncertainty in outcomes, and very large sample sizes will be needed to address such uncertainty.

CATIE not intended to determine coverage policy

Although we identified a number of important limitations of CATIE from the technical and clinical perspectives, among the most important from a policy perspective is that it was never intended to determine coverage policy. Indeed, the high rate of medication switching in CATIE implies that policy decisions based on CATIE's findings are relevant only in contexts that allow for effective switching. This is important because if comparative effectiveness research is motivated by attempts to inform payment policy, then it must be able to address these sorts of concerns. This suggests that if comparative effectiveness research is to be used to guide policy decisions, such as changes in coverage policy, then such research should include direct evaluations of policy changes.

Such analyses could include observational studies of policy changes or large-scale social experiments in which patients are randomly assigned to different forms of coverage, of which there are some examples in health care (such as the RAND Health Insurance Experiment). Social experiments could also be designed to work at higher levels—for example, if states are assigned different treatments.

The importance of considering social experiments as an important part of comparative effectiveness research is increased by our findings about the remarkably large sample sizes needed for cost-effectiveness studies, whether assessed by traditional deterministic sample-size calculations or by value-of-information approaches. Traditional clinical trials with samples of the size our results suggest are needed seem quite unlikely to be funded in the current funding environment, even if policymakers accept our conclusion that such studies would produce expected benefits that justify their expense, and they might not be feasible even if funding were not a concern.

Influence of switching behavior on value of research

The idea that switching behavior itself influences the value of research if patients tend to gravitate toward (or away from) preferred treatments over time and the idea that switching behavior may be influenced by research that suggests some treatment is dominant show that merely increasing the size of comparative effectiveness studies such as CATIE is unlikely to provide definitive evidence on the best use of antipsychotic agents in schizophrenia. This point is reinforced further by the need to study the potential for more carefully specified treatment algorithms involving multiple medications (at a point in time or over time) to alter cost and outcomes.

Value of CATIE study; limitations of ours

Although the CATIE study had major limitations, it should be emphasized that analysis performed here would not have been possible without the data it provided. Comparative effectiveness, like most research, is inherently iterative; prior studies inform future ones. Accordingly, our study has a variety of limitations, including being dependent on mortality data collected before the atypicals were available, being limited by the duration of and outcomes data collected by CATIE, and the distributional assumptions we made to analyze those data. Finally, our analysis cannot address the value of algorithms that may improve decisions about which sequences or combinations of drugs should be used when initial treatment with a single drug does not produce a desirable outcome.

Answering these questions will require further investment in research that will need to be justified relative to other areas of research. Building on the tradition of CATIE in establishing the potential of comparative effectiveness research to inform policy making, value-of-research calculations such as those presented here can play an important role in designing such trials, prioritizing them relative to other studies, and building the evidence

base to assess the appropriate level of overall spending on comparative effectiveness research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was sponsored by Best Practice Inc., with partial and unrestricted grant support from the Foundation for Education and Research on Mental Illness, Janssen Pharmaceuticals, and the Center for Medicine in the Public Interest. David Meltzer also acknowledges salary support from the Agency for Healthcare Research and Quality through the Hospital Medicine and Economics Center for Education and Research in Therapeutics (U18HS016967) and the National Institute on Aging through a mid-career career development award (K24 AG031326). Herbert Meltzer acknowledges salary support from and participation in the CATIE study as a consultant and investigator. A part of Anirban Basu's time was supported by a research grant from the National Institute of Mental Health (1R01MH083706-01).

NOTES

1. Tandon R, Fleischhacker WW. Comparative Efficacy of Antipsychotics in the Treatment of Schizophrenia: A Critical Assessment. *Schizophrenia Research*. 2005; 79(2–3):145–155. [PubMed: 16139989]
2. This paper uses the commonly used terms “first-generation antipsychotics” and “second-generation anti-psychotics,” which are sometimes called “typical” and “atypical” antipsychotic drugs H.Y Meltzer, “Atypical Antipsychotic Drugs Have Their Merits. *Lancet*. 2009; 373(9668):1007. [PubMed: 19304014]
3. Kane JM. Tardive Dyskinesia Rates with Atypical Antipsychotics in Adults: Prevalence and Incidence. *Journal of Clinical Psychiatry*. 2004; 65(9 Supp):16–20. [PubMed: 15189107]
4. We excluded clozapine because of its risk of causing dangerously low white blood cell counts (agranulocytosis), which has limited its approved use for patients who do not respond sufficiently to treatment with other antipsychotic drugs or who cannot tolerate the drug's greater potential to cause movement disorders.
5. Kahn RS, et al. Effectiveness of Antipsychotic Drugs in First-Episode Schizophrenia and Schizophrenia-form Disorder: An Open Randomised Clinical Trial. *Lancet*. 2008; 371(9618):1085–1097. [PubMed: 18374841] Meltzer HY. Focus on the Metabolic Consequences of Long-Term Treatment with Olanzapine, Quetiapine and Risperidone: Are There Differences? *International Journal of Neuropsychopharmacology*. 2005; 8(2):153–156. [PubMed: 15780147] Woodward ND, et al. A Meta-Analysis of Neuropsychological Change to Clozapine, Olanzapine, Quetiapine, and Risperidone in Schizophrenia. *International Journal of Neuropsychopharmacology*. 2005; 8(3):457–472. [PubMed: 15784157] Lagnado L. Prescription Abuse Seen in U.S. Nursing Homes. *Wall Street Journal*. December.2007 4
6. Lieberman J, et al. Effectiveness of Antipsychotic Drugs in Patients with Chronic Schizophrenia. *New England Journal of Medicine*. 2005; 353(12):1209–1223. [PubMed: 16172203]
7. Rosenheck RA, et al. Cost-Effectiveness of Second Generation Antipsychotics and Perphenazine in a Randomized Trial of Treatment for Chronic Schizophrenia. *American Journal of Psychiatry*. 2006; 163(12):2080–2089. [PubMed: 17151158]
8. Lewis S, Lieberman J. CATIE and CUtLASS: Can We Handle the Truth? *British Journal of Psychiatry*. 2008; 192(3):161–163. [PubMed: 18310570] Rosenheck RA, et al. Rethinking Antipsychotic Formulary Policy. *Schizophrenia Bulletin*. 2008; 34(2):375–380. [PubMed: 17634413] Lieberman JA, Hsiao JK. Interpreting the Results of the CATIE Study. *Psychiatric Services*. 2006; 57(1):139. [PubMed: 16399980] Lieberman JA. What the CATIE Study Means for Clinical Practice. *Psychiatric Services*. 2006; 57(8):1075. [PubMed: 16870955]
9. Rosenheck, et al. Cost-Effectiveness. Lieberman, J. Plenary Lecture, International Congress of Schizophrenia Research. Colorado Springs; April 2007; Davis J, Chen N, Glick ID. Issues That May Determine the Outcome of Antipsychotic Trials: Industry Sponsorship and Extrapyramidal

- Side Effects. *Neuropsychopharmacology*. 2008; 33(5):971–975. [PubMed: 17609679] Leucht S, et al. Second- Generation versus First-Generation Antipsychotic Drugs for Schizophrenia: A Meta-Analysis. *Lancet*. 2009; 373(9657):31–41. [PubMed: 19058842]
10. Rosenheck, RA., et al. CATIE Investigators' Educational Series. 2007. Rosenheck, RA. New York Times. September 21. 2005 "Comparing Schizophrenia Drugs" (editorial). Soumerai SB, Law MR. Cost-Effectiveness of Schizophrenia Pharmacotherapy. *American Journal of Psychiatry*. 2007; 164(4):678. [PubMed: 17403984]
 11. Jones PB, et al. Randomized Controlled Trial of the Effect on Quality of Life of Second- versus First-Generation Antipsychotic Drugs in Schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Archives of General Psychiatry*. 2006; 63(10):1079–1087. [PubMed: 17015810] Kahn, et al. Effectiveness of Antipsychotic Drugs.
 12. Neumann, PJ. Using Cost-Effectiveness to Improve Health Care. New York: Oxford University Press; 2004.
 13. The CATIE Investigators' Educational Series recommends that states' public health officials consider restricting coverage of second-generation antipsychotics.
 14. Tandon R, et al. World Psychiatric Association Pharmacopsychiatry Section Statement on Comparative Effectiveness of Antipsychotics in the Treatment of Schizophrenia. *Schizophrenia Research*. 2008; 100(1–3):20–38. [PubMed: 18243663]
 15. Möller HJ. Are the New Antipsychotics No Better Than the Classical Neuroleptics? The Problematic Answer from the CATIE Study. *European Archives of Psychiatry and Clinical Neuroscience*. 2005; 255(6):371–372. [PubMed: 16382376] Soumerai and Law. Cost-Effectiveness. Covell NH, Finnerty MT, Essock SM. Implications of CATIE for Mental Health Services Researchers. *Psychiatric Services*. 2008; 59(5):526–529. [PubMed: 18451010]
 16. Kinon BJ, et al. Standard and Higher Dose of Olanzapine in Patients with Schizophrenia or Schizoaffective Disorder: A Randomized, Double-Blind, Fixed-Dose Study. *Journal of Clinical Psychopharmacology*. 2008; 28(4):392–400. [PubMed: 18626265] Meltzer HY, et al. A Randomized, Double-Blind Comparison of Clozapine and High-Dose Olanzapine in Treatment-Resistant Patients with Schizophrenia. *Journal of Clinical Psychiatry*. 2008; 69(2):274–285. [PubMed: 18232726] Meltzer HY, Bobo WV. Interpreting the Efficacy Findings in the CATIE Study: What Clinicians Should Know. *CNS Spectrum*. 2006; 11(7 supp 7):14–24. Weiden PJ. Discontinuing and Switching Antipsychotic Medications: Understanding the CATIE Schizophrenia Trial. *Journal of Clinical Psychiatry*. 2007; 68(1 Supp):12–19. [PubMed: 17286523]
 17. Soumerai and Law. Cost-Effectiveness. An example of algorithms is the International Psychopharmacology Algorithm Project. (home page: <http://www.ipap.org>)
 18. Tiihonen J, et al. Effectiveness of Antipsychotic Treatments in a Nationwide Cohort of Patients in Community Care after First Hospitalisation due to Schizophrenia and Schizoaffective Disorder: Observational Follow-Up Study. *BMJ*. 2006; 333(7561):224–231. [PubMed: 16825203] Kahn, et al. Effectiveness of Antipsychotic Drugs.
 19. Specifically, the standard gamble technique assigns a utility weight between 0 and 1 to a health state by finding the highest risk of immediate death that people will accept in order to be restored to perfect health instead of the health state being rated.
 20. Dale W, et al. Predicting Utility Ratings for Joint Health States from Single Health States in Prostate Cancer: Empirical Testing of Three Alternative Theories. *Medical Decision Making*. 2008; 28(1):102–112. [PubMed: 18057188]
 21. Rosenheck, et al. Cost-Effectiveness.
 22. The problem has to do with the uses of transformed data. See Technical Appendix 1, "Detailed Discussion of Data Transformation, Part I," online at <http://content.healthaffairs.org/cgi/content/full/hlthaff.28.5.w794/DC2>.
 23. Further discussion of this problem is available in Technical Appendix 1, "Detailed Discussion of Data Transformation, Part II," *ibid*.
 24. Rosenheck, et al. Cost-Effectiveness.
 25. Lewis and Lieberman. CATIE and CUtLASS.

26. Ibid. Tiihonen, et al. Effectiveness of Antipsychotic Treatments. Tandon, et al. World Psychiatric Association Pharmacopsychiatry Section Statement. Jones, et al. Randomized Controlled Trial of the Effect on Quality of Life.
27. Claxton K, et al. Bayesian Value-of-Information Analysis: An Application to a Policy Model of Alzheimer's Disease. *International Journal of Technology Assessment in Health Care*. 2001; 17(1):38–55. [PubMed: 11329844] National Institute for Health and Clinical Excellence. Technology Appraisal Guidance no. 1. London: NICE; March. 2000 Guidance on the Removal of Wisdom Teeth.
28. Meltzer D. Addressing Uncertainty in Medical Cost-Effectiveness Analysis: Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research. *Journal of Health Economics*. 2001; 20(1):109–129. [PubMed: 11148867] Claxton K. The Irrelevance of Inference: A Decision-Making Approach to the Stochastic Evaluation of Health Care Technologies. *Journal of Health Economics*. 1999; 18(3):341–364. [PubMed: 10537899]
29. Claxton, et al. Bayesian Value of Information Analysis.
30. For more detail, see Technical Appendix 2, “Details on Calculation Method,” online as in Note 22.
31. These means and variances come from simulated distributions of mean QALYs and mean annual costs; the full distributions can be seen in Technical Appendix 3, online as in Note 22.
32. Braithwaite RS, et al. What Does the Value of Modern Medicine Say about the \$50,000 per Quality-Adjusted Life-Year Decision Rule? *Medical Care*. 2008; 46(4):349–356. [PubMed: 18362813]
33. It is interesting that the typical antipsychotics are less than 50 percent likely to be cost-effective yet appear to be the cost-effective choice. This is possible because the average net benefit if they are cost-effective is greater than the average negative net benefit if they are not cost-effective.
34. See Technical Appendix 2, online as in Note 22.

EXHIBIT 3.

Value Of Future Research To Prevalent And Incident Schizophrenia Cohorts At \$50,000 Per Quality-Adjusted Life-Year (QALY), 2007–2087

SOURCE: Authors' calculations.

NOTE: CATIE is Clinical Antipsychotic Trials of Intervention Effectiveness.

EXHIBIT 4.

Value Of Research And The Probability That The CATIE Conclusion Is Incorrect (That Perphenazine Is Not Cost-Effective)

SOURCE: Authors' calculations.

NOTES: CATIE is Clinical Antipsychotic Trials of Intervention Effectiveness. Maximum value of research (gray line and points) relates to the left-hand y axis. Probability that CATIE conclusion is incorrect (black line and points) relates to the right-hand y axis. QALY is quality-adjusted life-year.

EXHIBIT 5.

Net Expected Value Of Sample Information (At \$50,000, \$100,000, And \$150,000 Per Quality-Adjusted Life-Year, Or QALY)

SOURCE: Authors' calculations.

EXHIBIT 1
Cost-Effectiveness Results From The Clinical Antipsychotic Trials Of Intervention Effectiveness (CATIE) Study

Drug	Monthly costs (\$)		QALYs		Incremental cost- effectiveness ratio (\$/QALY)
	Mean	(SD)	Mean	(SD)	
Perphenazine	817	(728)	0.722	(0.0064)	^a
Olanzapine	1,619	(1,442)	0.723	(0.0063)	9,624,000
Risperidone	1,635	(1,457)	0.706	(0.0066)	Dominated
Quetiapine	1,680	(1,497)	0.721	(0.0065)	Dominated

SOURCES: R.A. Rosenheck et al., “Cost-Effectiveness of Second Generation Antipsychotics and Perphenazine in a Randomized Trial of Treatment for Chronic Schizophrenia,” *American Journal of Psychiatry* 163, no. 12 (2006): 2080–2089; and authors’ personal communications with Robert Rosenheck on the SAS mixed-models analyses of costs and quality-adjusted life-years (QALYs) for CATIE study.

NOTES: “Dominated” means that some other treatment was superior in both costs and outcomes. SD is standard deviation.

^a Other drugs are compared to perphenazine.

EXHIBIT 2

Simulated Mean Quality-Adjusted Life-Years (QALYs) And Mean Costs For Four Antipsychotic Drugs

Drug	QALYs		Costs (\$)	
	Mean	Variance	Mean	Variance
Olanzapine	0.723	0.0063	1,606	1,421
Quetiapine	0.721	0.0065	1,685	1,485
Risperidone	0.706	0.0066	1,621	1,439
Perphenazine	0.722	0.0064	810	723

SOURCE: Authors' calculations based on data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Study data; and authors' personal communication with Robert Rosenheck.