

Number Needed to Treat (NNT) and Number Needed to Harm (NNH) in Randomized, Blinded Trials Comparing Olanzapine to Other Atypical Antipsychotics for Treatment of Schizophrenia

Virginia Stauffer¹, Jamie Karagianis², Virginia Sutton³, Haya Ascher-Svanum¹, Tamas Treuer⁴, Mauricio Silva de Lima⁵, Tamara Ball⁶, Vicki Poole-Hoffmann¹, Mauricio Tohen^{1,7}

Abstract

Objective: All-cause discontinuation is considered a proxy for a medication's effectiveness. We examined the number needed to treat (NNT) to avoid all-cause medication discontinuation in head-to-head clinical trials of olanzapine versus other atypical antipsychotics. **Methods:** This was a post hoc analysis of five randomized trials of olanzapine versus risperidone (n=2), ziprasidone (n=2) and quetiapine (n=1) for treatment of patients with schizophrenia. All trials were of at least six months' duration. The NNT or number needed to harm (NNH) was determined for all-cause discontinuation and other efficacy and safety parameters. NNT and NNH are calculated as the reciprocal of attributable risk. Desirable treatments are characterized as having low NNTs and relatively high NNHs. These measures are useful for ranking treatments when the same outcome measure is assessed over the same amount of time in similar patients. In this analysis, positive values indicated the superiority of olanzapine and negative values indicated the superiority of the comparator treatment. **Results:** Statistically significant NNTs (95% confidence intervals) to avoid all-cause medication discontinuation were 6 (4, 12) and 7 (5, 19) for olanzapine versus ziprasidone; and 7 (4, 22) for olanzapine versus quetiapine. The NNHs indicated greater likelihood of weight gain with olanzapine versus all comparators except quetiapine. Statistically significant NNHs indicated greater likelihood of weight gain with olanzapine in one of the risperidone studies (-7 [-16, -4]) and both studies in which olanzapine was compared to ziprasidone (-4 [-5, -3] and -5 [-7, -4]). **Conclusions:** In this post hoc analysis of five studies in which olanzapine was compared to other atypical antipsychotics using the evidence-based medicine tools of NNT and NNH, olanzapine was superior to ziprasidone and quetiapine for prevention of treatment discontinuation for any cause. Ziprasidone was least associated with potentially clinically significant weight gain, followed by risperidone, with olanzapine and quetiapine ranked last.

Key Words: All-Cause Discontinuation, Atypical Antipsychotics, Number Needed to Treat, Number Needed to Harm, NNT, NNH, Schizophrenia, Olanzapine

¹ Lilly Research Laboratories, Indianapolis, IN, USA

² Eli Lilly Canada Inc., Toronto, Ontario, Canada

³ i3 Research, Memphis, TN, USA

⁴ Eli Lilly and Company, Vienna, Austria

⁵ Eli Lilly and Company, United Kingdom

⁶ i3 Statprobe, Ann Arbor, MI, USA

⁷ McLean Hospital, Harvard Medical School, Cambridge, MA, USA

Address for correspondence: Virginia L. Stauffer, PharmD,
Eli Lilly and Company, Lilly Corporate Center, Drop Code 4133,
Indianapolis, IN 46285

Phone: 317-277-2904; Fax: 317-276-7100;

E-mail: VStauffer@Lilly.com

Submitted: September 28, 2007; Revised: March 11, 2008;

Accepted: April 9, 2008

Introduction

Discontinuation of antipsychotic medication affects a sizeable proportion of patients with schizophrenia and is an important factor in their clinical management. In a recent integrated analysis of sixteen published, double-blind, randomized trials of ≥ 12 weeks' duration in which olanzapine was compared to other antipsychotic drugs, discontinuation rates ranged from 30 to 88% (1). Interruption or discontinuation of antipsychotic therapy due to poor adherence is expected to be associated with increased rates of relapse and psychiatric hospitalization, decreased functional outcome

and quality of life, and increased treatment costs (2-6).

The primary outcome of this study—time until discontinuation of treatment for any cause—is recognized as a valid measure of treatment effectiveness, an index that incorporates efficacy, safety and tolerability as evaluated by both patient and physician (7). Time to all-cause medication discontinuation and rates of discontinuation were key outcome measures of treatment effectiveness in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) (8), a large, randomized, double-blind, eighteen-month, National Institute of Mental Health (NIMH)-sponsored trial.

The number needed to treat (NNT) is a tool of evidence-based medicine designed to translate research findings into readily usable information for the practicing clinician. NNT is a measure of effect size: a number that indicates how many patients would need to be treated using intervention A instead of intervention B to see one additional success. When the comparison involves an adverse outcome, the measure is referred to as the number needed to harm (NNH). NNT (or NNH) is calculated by taking the reciprocal of attributable risk (AR): the difference in rates for the outcome of interest between two interventions. A low NNT indicates that the treatments being compared are substantially different, whereas a high NNT suggests very little difference. Desirable treatments have small NNTs and relatively large NNHs (9, 10).

An example of the clinical application of NNT and NNH is the use of influenza vaccine versus no treatment in healthy children between ages one and six. Vaccination reduces the risk of a culture-confirmed case of influenza with an NNT of 6, which means that one case of influenza can be expected to be prevented for every six children vaccinated, regardless of their exposure status. The vaccine has an NNH of 72 for low-grade fever, meaning that one additional low-grade fever can be expected to occur for every seventy-two children vaccinated (11). On balance, the vaccine offers a high likelihood of protection from a potentially dangerous infection and a low likelihood of occurrence of a relatively minor adverse event (AE); therefore, most pediatricians advise vaccination for healthy children in this age group.

Both the NNT and NNH are expressed as single numbers referred to as point estimates. As with any statistical result based on experimental data, the true value of the NNT or NNH can be higher or lower than the point estimate, and the range of possibility for the true value is indicated by confidence intervals (CIs). Whenever the NNT or NNH value is not statistically significant, the endpoints of the 95% CIs are opposite in sign and the interval includes infinity (12, 13). This occurs when the range of possible AR values includes zero; that is, there may not be any difference in event rates for patients treated with olanzapine versus the comparator. Since the inverse of zero is undefined, the range of possible

point estimates has to include infinity, suggesting that, based on our data sample, an infinite number of patients might need to be treated with olanzapine rather than the comparator to expect one additional success or harm. The point estimate may still provide guidance in clinical decision making, but should be used with caution until further data permit determination of a finite CI.

The relative benefits and risks of treatment with different antipsychotic medications has long been a topic of debate within the psychiatric field. Phases 1 (8) and 2 (14, 15) of the CATIE trial provided a vast amount of data regarding the safety and efficacy of antipsychotic agents. In several recent publications (16, 17), NNT and NNH have been used to place data from the CATIE trial into a clinically meaningful context. Citrome and Stroup reported that in Phase 1 of CATIE, the NNTs (95% CIs) to avoid all-cause discontinuation for olanzapine compared to quetiapine, risperidone and ziprasidone were 6 (4, 9), 11 (6, 35) and 7 (5, 13), respectively. Compared to olanzapine, NNHs (95% CIs) for weight gain $\geq 7\%$ above baseline were -8 (-14, -5), -7 (-11, 5) and -5 (-7, -4), respectively.

Whether NNTs and NNHs derived from the independently funded CATIE study are comparable to those derived from industry-sponsored, long-term, randomized clinical trials is unknown. In this analysis, we use the clinically useful measure, NNT, to present data on all-cause discontinuation from five randomized, double-blind, comparative clinical trials of olanzapine versus other atypical antipsychotics from the Eli Lilly and Company Clinical Trial Database. The NNT and NNH for other secondary efficacy and safety measures are also presented, along with relative rankings based on these data.

Methods

Study Selection and Patient Characteristics

This was a post hoc analysis using data from five clinical trials within the Eli Lilly and Company Clinical Trial Database. Each study met the following criteria: randomized, double-blind clinical trial; head-to-head comparison of olanzapine versus at least one other atypical antipsychotic; study duration of at least six months; and, participants meeting *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)* (18) criteria for schizophrenia, schizophreniform disorder or schizoaffective disorder.

Five studies met inclusion criteria, including two trials comparing olanzapine to risperidone (RISP 1 [19], RISP 2 [20]), two trials comparing olanzapine to ziprasidone (ZIP 1 [21], ZIP 2 [22]), and one trial comparing olanzapine to quetiapine (QUET [23]). None of the studies included a pla-

Table 1 Characteristics of the Studies Used to Calculate NNTs and NNHs for Olanzapine versus other Atypical Antipsychotics

Primary Reference (Author, Year)	Primary Outcomes	Study Drugs	N	Mean Modal Dose (mg/day [SD])	Study Duration (weeks)	Diagnoses	Other Baseline Inclusion Criteria
Tran, 1997 (RISP 1)	Efficacy, Safety	Olanzapine Risperidone	172 167	17.2 (3.6) 7.2 (2.7)	28	Schz, Schzfm, Schzaff	Inpatient and outpatient Age 18 to 65 BPRS (ext) score ≥ 42
Keefe, 2006 (RISP 2)	Neurocognitive, Psychosocial, Efficacy, Safety	Olanzapine Haloperidol* Risperidone	159 97 158	13.1 (3.9) 8.3 (3.7) 5.3 (2.1)	52	Schz, Schzaff	Inpatient and outpatient Age 18 to 65 Score ≥ 18 on the BPRS (ext) and ≥ 4 on at least 2 positive items of the PANSS
Breier, 2005 (ZIP 1)	Efficacy, Safety	Olanzapine Ziprasidone	277 271	15.3 (4.5) 116.0 (39.9)	28	Schz	Inpatient and outpatient Age 18 to 75 Scores ≥ 42 on the BPRS (ext), ≥ 4 on at least one positive symptom item of the PANSS, and ≥ 4 on the severity of illness subscale of the CGI
Kinon, 2006 (ZIP 2)	Depressive symptoms, Efficacy, Safety	Olanzapine (10, 15, 20 mg) Ziprasidone (80, 120, 160 mg)	202 192	14.2 [†] 110.2 [†]	24	Schz, Schzaff	Inpatient and outpatient Age 18 to 60 Scores ≥ 16 (mild depression) on the MADRS and ≥ 4 (pervasive feelings of sadness or gloominess) on item 2 (reported sadness) of the MADRS
Kinon, 2006 (QUET)	Negative symptoms, Functional outcome, Efficacy, Safety	Olanzapine Quetiapine	171 175	15.6 (4.3) 455.8 (156.3)	24	Schz, Schzaff	Outpatients Age 18 to 65 Score ≥ 4 on at least 3, or ≥ 5 on at least 2, of the 7 negative symptom items of the PANSS, and ≤ 60 (moderate difficulties) on the GAF

*52 weeks after enrollment began, this arm was incompletely populated, and randomization to haloperidol was discontinued per protocol guidelines.

†This study had multiple fixed doses; therefore, SD is not given.

Schz=Schizophrenia; Schzfm=Schizophreniform Disorder; Schzaff=Schizoaffective Disorder; N=number; NNTs=numbers needed to treat; NNHs=numbers needed to harm; BPRS (ext)=Brief Psychiatric Rating Scale (scored 0–6) extracted from the Positive and Negative Syndrome Scale (30); PANSS=Positive and Negative Syndrome Scale (scored 1–7) (24); CGI=Clinical Global Impression Scale (29); MADRS=Montgomery-Asberg Depression Rating Scale (35); GAF=Global Assessment of Functioning Scale (18); SD=standard deviation

cebo arm. One study included randomization to haloperidol treatment but, 52 weeks after enrollment began, this arm was incompletely populated and randomization to haloperidol was discontinued per protocol guidelines. Data from the 97 patients assigned to receive haloperidol were not included in the present analysis (19). Four of the five trials lasted either 24 or 28 weeks. One study lasted 52 weeks (19) and, to allow comparability with the other trials, only data collected through week 24 were included in this analysis.

Each study examined efficacy and safety outcomes of patients diagnosed with schizophrenia. Three of the five studies had an additional focus: one included evaluation of neurocognitive outcomes (19), another enrolled patients with prominent depressive symptoms (22) and a third enrolled patients with prominent negative symptoms and poor functioning at baseline (23). The five studies are summarized in Table 1. Detailed descriptions are available in their respective published reports.

All were multisite studies and were carried out internationally (19-21) or within the United States (22, 23). A total of 2,041 men and women aged 18 to 70 years participated in the trials. All protocols were approved by the ethical review boards responsible for individual study sites, and all patients or their legal guardians provided written, informed consent consistent with all applicable regulations prior to receiving any study therapy or undergoing any study procedure. Antipsychotics were dosed within a specified range at clinician discretion, except in one study in which a multiple, fixed-dose design was used (22). A limited number of concomitant psychotropic medications were permitted: benzodiazepines/hypnotics; approved antiparkinson medications (for treatment, but not prevention, of extrapyramidal symptoms [EPS]); and, in two studies (22, 23), fixed doses of antidepressants if the patient had used them within thirty days prior to enrollment.

Assessments

The dichotomous variables included in this analysis were chosen *a priori*. Throughout this report, NNTs and NNHs are followed by 95% CIs. The primary outcome in this analysis was the NNT to avoid all-cause discontinuation for olanzapine versus the comparator in each study. Secondary outcomes included NNT to avoid discontinuation due to lack of efficacy and NNT to avoid discontinuation due to medication intolerability. For this analysis, lack of efficacy included discontinuation due to any psychiatric AE or due to patient and/or physician perception of lack of efficacy. Medication intolerability was defined as discontinuation due to any nonpsychiatric AE.

Secondary efficacy outcomes included NNTs to achieve response, to achieve remission, and to avoid worsening psychosis. The measure of psychopathology common to all

studies was the Positive and Negative Syndrome Scale score (PANSS) (scored 1–7) (24). Response was defined as a $\geq 20\%$ improvement at endpoint in the PANSS total score. Remission was defined as an endpoint score of mild or better (≤ 3) for each of the following PANSS items: delusions, conceptual disorganization, hallucinatory behavior, excitement, blunted affect, lack of spontaneity and flow of conversation, mannerisms and posturing, and unusual thought content (25). Worsening psychosis was defined as an increase in any PANSS positive item to >4 and an absolute increase of at least 4 on the PANSS positive subscale, or an increase in any PANSS positive item to >4 and an absolute increase of at least 2 on that specific item (26).

Adverse outcomes included treatment-emergent EPS (including tardive dyskinesia), weight gain and metabolic changes. Patients were identified as having EPS if item 4 was ≥ 2 on the Barnes Rating Scale for Drug-Induced Akathisia (27), or if their total score on the Simpson-Angus Rating Scale (28) was >3 . Patients were identified as having developed tardive dyskinesia if, at any time during the study, they had a score ≥ 2 on two items or a score ≥ 3 on any one item on the Abnormal Involuntary Movement Scale (AIMS) (29).

Potentially clinically significant weight gain, defined as a $\geq 7\%$ increase at any time over baseline, was evaluated for participants in all studies. Metabolic outcomes were evaluated at 24 weeks for patients in the two studies in which fasting laboratory values were available (21, 22). Abnormal values were defined as: total cholesterol ≥ 200 mg/dl; low-density lipoprotein cholesterol (LDL-C) ≥ 100 mg/dl; high-density lipoprotein cholesterol (HDL-C) ≤ 40 mg/dl in male patients and <50 mg/dl in female patients; fasting triglycerides ≥ 150 mg/dl; fasting glucose ≥ 100 mg/dl; and, prolactin ≥ 18.77 ng/ml in male patients and ≥ 24.2 ng/ml in female patients. (To convert values from mg/dl to mmol/l, multiply by 0.02586 for cholesterol, by 0.01129 for triglycerides and by 0.05551 for glucose. To convert prolactin from ng/ml to pmol, multiply by 0.42478.)

Patients were assessed at baseline and intermittently through week 24 or week 28. The timing of assessments varied by measure according to the schedule of events in the individual protocol. Patients who did not complete the study were also assessed at the time of discontinuation. Endpoint was determined as the last nonmissing observation during the 24- to 28-week period. In addition, investigators completed a clinical report form indicating the reason for treatment discontinuation.

Statistical Analysis

Data were extracted from the original studies found in the Eli Lilly and Company Clinical Trial Database. By convention, calculations of NNT and NNH were structured such that olanzapine was superior to the comparator when

the value was positive, and the comparator was superior to olanzapine when the value was negative. Equations used to calculate NNT and NNH and their upper and lower CIs, are shown in Table 2.

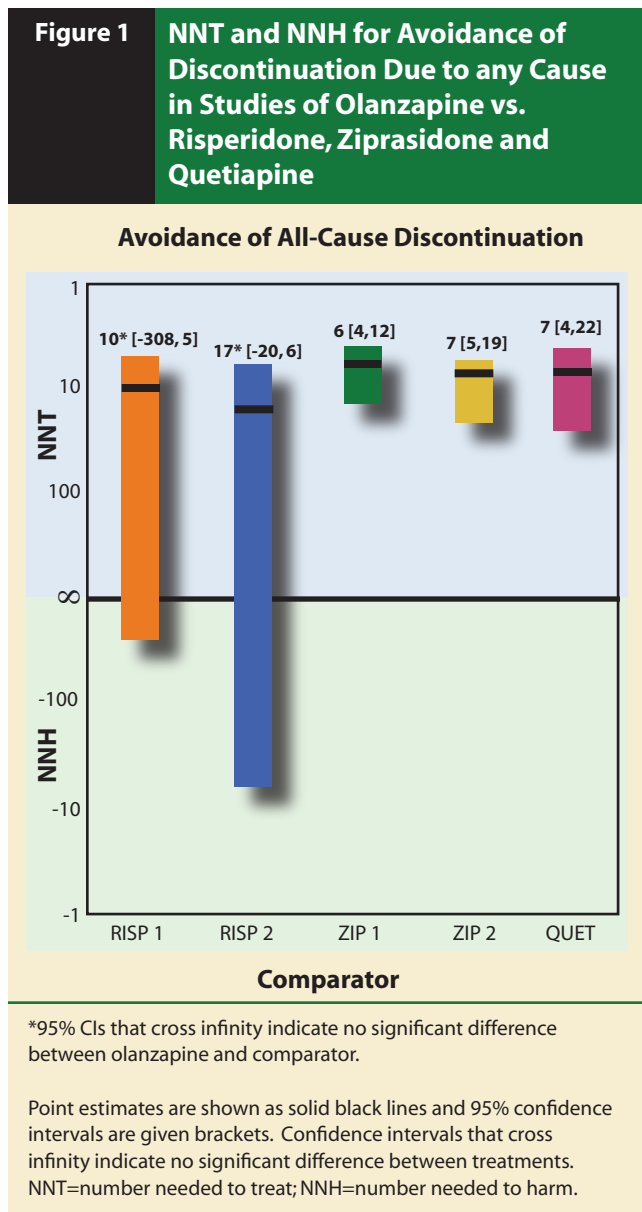
Table 2 Calculation of NNT (or NNH) and 95% Confidence Intervals	
Calculation of NNT (or NNH)	
NNT (or NNH) = 1/Attributable Risk (AR), where $AR = \text{event rate}_{\text{olanzapine}} - \text{event rate}_{\text{comparator}}$ -OR- $AR = \text{event rate}_{\text{comparator}} - \text{event rate}_{\text{olanzapine}}$	
By convention, calculations were structured so that olanzapine was superior to comparator when the NNT or NNH was positive, and the comparator was superior when the NNT or NNH was negative.	
Calculation of 95% Confidence Intervals	
$\text{Lower} = \frac{1}{AR + Z_{.95} * \sqrt{\frac{pct_1(1-pct_1)}{n_1} + \frac{pct_2(1-pct_2)}{n_2}}}$	
$\text{Upper} = \frac{1}{AR - Z_{.95} * \sqrt{\frac{pct_1(1-pct_1)}{n_1} + \frac{pct_2(1-pct_2)}{n_2}}}$	

Results

Avoidance of Discontinuation

Treatment with olanzapine led to lower likelihoods of discontinuation for any cause compared to treatment with ziprasidone (2 studies) or quetiapine (1 study) (see Figure 1). The NNTs were low and CIs were narrow (6 [4, 12], 7 [5, 19] and 7 [4, 22], respectively), indicating that the estimates were clinically meaningful and fairly precise. For olanzapine versus risperidone, the NNTs to avoid all-cause discontinuation suggested no statistically significant difference between treatments.

As previously reported by Liu-Seifert et al. (3), the most common reasons for discontinuation in these five studies were lack of efficacy/psychiatric AE. The NNTs to avoid these outcomes are shown in Figures 2 and 3. The NNTs to avoid discontinuation due to lack of efficacy were small with narrow CIs in the study comparing olanzapine to quetiapine (NNT=6 [4, 10]). A significant treatment difference was also observed in one of the two trials comparing olanzapine with ziprasidone (NNT=10 [7, 26]). Point estimates for this outcome also favored olanzapine over risperidone, but were not statistically significant (see Figure 2).

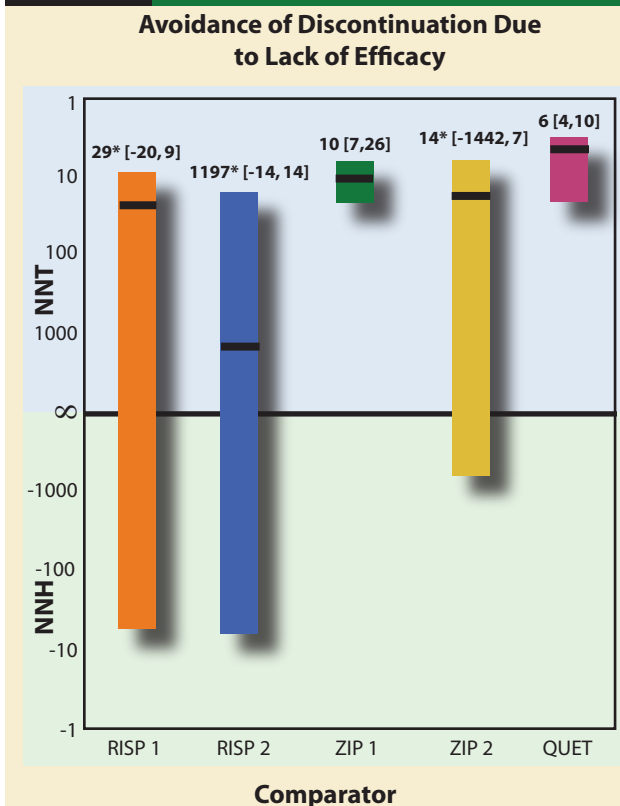


The NNTs to avoid discontinuation due to medication intolerance were somewhat inconsistent, as shown in Figure 3. The CIs of these point estimates suggest a significant difference between treatments for one of each of the studies comparing olanzapine to risperidone (16 [9, 97]) and to ziprasidone (16 [8, 158]). The NNTs favored the comparator in each of the other studies of olanzapine versus risperidone or ziprasidone, but CIs were large and included infinity. The NNT also favored olanzapine versus quetiapine, but failed to demonstrate statistical significance.

Efficacy

The NNTs for response, remission and avoidance of worsening psychosis are shown in Table 3. In one study in which ziprasidone was the comparator, olanzapine showed significantly greater likelihood of all three efficacy outcomes,

Figure 2 NNT and NNH for Avoidance of Discontinuation Due to Lack of Efficacy in Studies of Olanzapine vs. Risperidone, Ziprasidone and Quetiapine



*95% CIs that cross infinity indicate no significant difference between olanzapine and comparator.

Point estimates are shown as solid black lines and 95% confidence intervals are given in brackets. Confidence intervals that cross infinity indicate no significant difference between treatments. NNT=number needed to treat; NNH=number needed to harm.

with an NNT of 7 (5, 14) for response, 7 (5, 13) for remission and 24 (13, 323) for avoidance of worsening psychosis. In the other study where ziprasidone was the comparator, there was a significant difference compared with olanzapine for remission (NNT=10 [5, 122]). Olanzapine also showed a significant advantage over quetiapine for response (NNT=10 [5, 744]) and avoidance of worsening psychosis (12 [7, 108]).

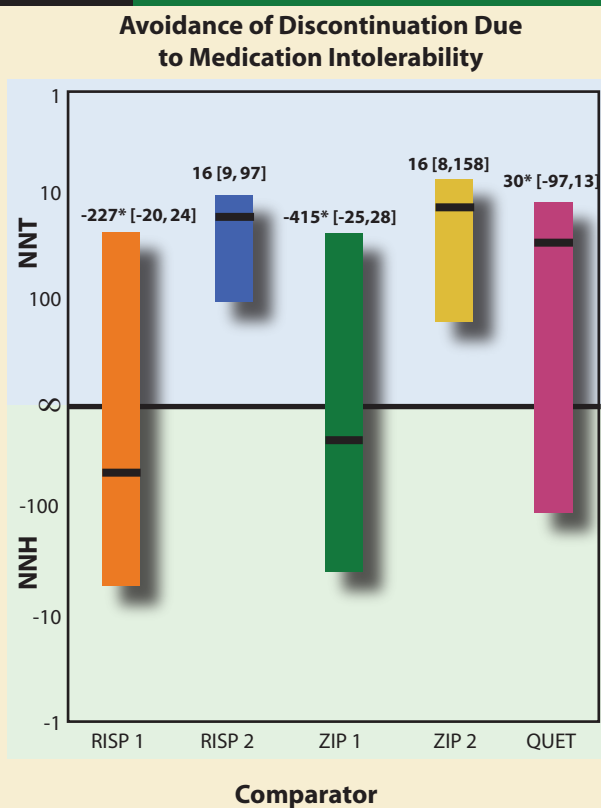
Safety

The NNHs derived from treatment-emergent EPS data are shown in Table 4. In one study comparing olanzapine with risperidone, clinically meaningful NNHs were seen on the Barnes Rating Scale for Drug-Induced Akathisia (10 [5, 333]) and on the Simpson-Angus Rating Scale (8 [5, 28]). Other comparisons yielded NNHs lacking statistical significance.

The safety analysis revealed a consistent pattern of potentially clinically significant weight gain $\geq 7\%$ above baseline for olanzapine versus all comparators, but was not statistically significant for one of the studies involving risperidone and for the study involving quetiapine (see Figure 4). The NNHs were small and CIs narrow for three of the five studies, and point estimates favored the comparator in all trials. The NNHs were -10 (-5 to ∞ to 242) and -7 (-16, -4) versus risperidone, -4 (-5, -3) and -5 (-7, -4) versus ziprasidone and -17 (-8 to ∞ to 54) versus quetiapine.

Prolactin levels were available for all studies. Treatment with olanzapine was favorable to treatment with risperidone for avoidance of prolactin abnormalities. The NNHs were small, with very narrow CIs (3 [2, 3] for RISP 1 and 2 [2, 2] for RISP 2). Comparisons of olanzapine to ziprasidone and quetiapine showed large NNHs with CIs that included infin-

Figure 3 NNT and NNH for Avoidance of Discontinuation Due to Medication Intolerability in Studies of Olanzapine vs. Risperidone, Ziprasidone and Quetiapine



*95% CIs that cross infinity indicate no significant difference between olanzapine and comparator.

Point estimates are shown as solid black lines and 95% confidence intervals are given in brackets. Confidence intervals that cross infinity indicate no significant difference between treatments. NNT=number needed to treat; NNH=number needed to harm.

Table 3	NNT or NNH with (95% CI) for Efficacy Outcomes: Olanzapine versus Risperidone, Ziprasidone and Quetiapine*				
	Study				
	RISP 1	RISP 2	ZIP 1	ZIP 2	QUET
Response	-64 (-9 to ∞ to 12)	21 (-18 to ∞ to 7)	7 (5, 14)	11 (-147 to ∞ to 6)	10 (5, 744)
Remission	16 (-23 to ∞ to 6)	-60 (-9 to ∞ to 12)	7 (5, 13)	10 (5, 122)	28 (-16 to ∞ to 8)
Avoidance of Worsening Psychosis	-85 (-18 to ∞ to 30)	25 (-110 to ∞ to 11)	24 (13, 323)	38 (-35 to ∞ to 13)	12 (7, 108)

*When the NNT or NNH point estimate is not statistically significant, the endpoints of the 95% CI will be opposite in sign, and the interval itself will include infinity.

NNT=number needed to treat; NNH=number needed to harm; CI=confidence interval

Table 4	NNT or NNH with (95% CI) for Avoidance of Treatment-Emergent EPS: Olanzapine versus other Atypical Antipsychotics*				
	Study				
	RISP 1	RISP 2	ZIP 1	ZIP 2	QUET
AIMS	19 (-203 to ∞ to 9)	44 (-21 to ∞ to 11)	-107 (-22 to ∞ to 36)	25 (-71 to ∞ to 11)	112 (-21 to ∞ to 16)
Barnes	10 (5, 333)	19 (-18 to ∞ to 7)	32 (-34 to ∞ to 11)	18 (-142 to ∞ to 9)	-738 (-12 to ∞ to 12)
Simpson-Angus	8 (5, 28)	15 (-35 to ∞ to 7)	56 (-29 to ∞ to 15)	72 (-19 to ∞ to 13)	52 (-18 to ∞ to 11)

*When the NNT or NNH point estimate is not statistically significant, the endpoints of the 95% CI will be opposite in sign, and the interval itself will include infinity.

EPS=extrapyramidal symptoms; NNT=number needed to treat; NNH=number needed to harm; CI=confidence intervals; AIMS=Abnormal Involuntary Movement Scale; Barnes=Barnes Akathisia Scale; Simpson-Angus=Modified Simpson-Angus Scale

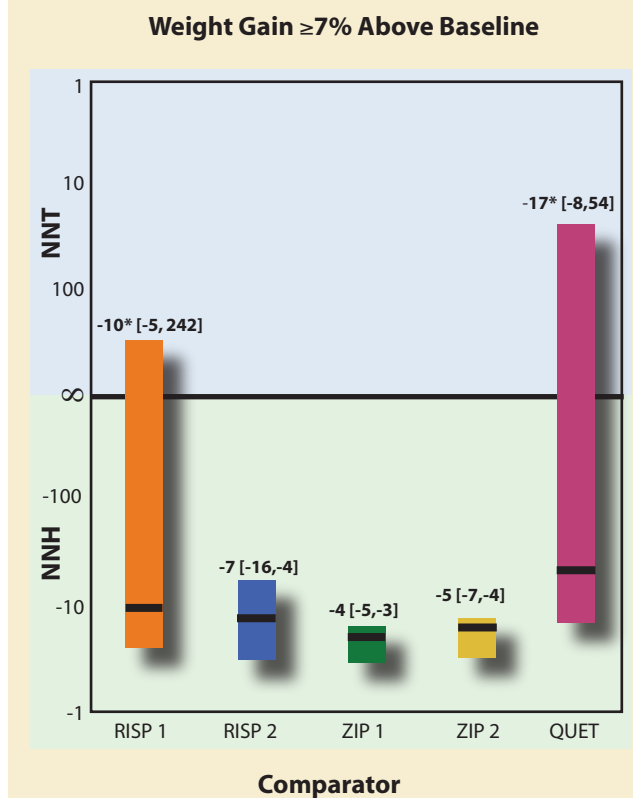
ity and, thus, were not statistically significant. Two studies in which ziprasidone was the comparator provided laboratory results obtained in the fasting state that were used to generate NNHs for metabolic AEs. Ziprasidone demonstrated a significant advantage over olanzapine with regard to treatment-emergent hypertriglyceridemia (NNHs=-4 [-7, -3] and -6 [-69, -3]). Data for triglyceride levels and other metabolic measures are shown in Table 5. In one of the two ziprasidone studies there was a greater likelihood of elevated total cholesterol associated with olanzapine (NNH=-7 [-22, -4]). Point estimates generally favored ziprasidone for avoidance

of elevated LDL-C, reduced HDL-C and fasting hyperglycemia, but CIs were wide and included infinity and, thus, were not statistically significant.

Discussion

In this analysis, data from five randomized trials in which olanzapine was compared to another atypical antipsychotic are presented in the clinically meaningful format of NNT and NNH. These values can be calculated easily and kept as a single numerical reminder of a particular therapy's potential for effectiveness or harm with regard to specific

Figure 4 NNT and NNH for Olanzapine vs. Comparators for Weight Gain $\geq 7\%$ Above Baseline



*95% CIs that cross infinity indicate no significant difference between olanzapine and comparator.

Point estimates are shown as solid black lines and 95% confidence intervals are given brackets. Confidence intervals that cross infinity indicate no significant difference between treatments.

NNT=number needed to treat; NNH=number needed to harm.

outcomes. These statistics are particularly useful for converting the sizable data output of systematic reviews or large clinical trials such as CATIE into directly applicable clinical information. Also, in studies of similar patients measuring the same outcome over the same duration of time, relative rankings of efficacy or harm can be made across studies.

Based on our analysis, over a six-month period an estimated six to seven patients would need to be treated with olanzapine versus ziprasidone or quetiapine to avoid one medication discontinuation for any cause. Favorable NNTs with olanzapine were also seen for efficacy compared to ziprasidone and quetiapine, and for avoidance of treatment-emergent EPS compared to risperidone. However, point estimates of NNH for potentially clinically meaningful weight gain demonstrated a disadvantage for olanzapine observed for every seventh patient in one of the studies involving risperidone and for every fourth to fifth patient in the studies involving ziprasidone. For the two studies in which blood

was obtained in the fasting state, treatment-emergent hypertriglyceridemia with olanzapine compared to ziprasidone was encountered in every fourth to sixth patient.

Our data provide a relative ranking of atypical antipsychotics for prevention of all-cause discontinuation as follows: olanzapine and risperidone with similar effectiveness, followed by ziprasidone and quetiapine, also with similar effectiveness. This ranking corresponds to that described in a 2003 efficacy meta-analysis by Davis et al. (30).

The NNTs and NNHs reported here for five comparator trials sponsored by Eli Lilly and Company can be compared to those observed in the NIMH-sponsored CATIE trial (16). Similar relative rankings can be seen for all-cause discontinuation and increases in body weight $>7\%$ from baseline. However, the CATIE trial differed from our studies in several respects, including the maximum length of participation (six months for studies included in the present analysis versus eighteen months for CATIE), sample size and medi-

Table 5 NNT or NNH with (95% CI) for Metabolic Parameters* Obtained from Patients in the Fasting State for Two Trials of Olanzapine versus Ziprasidone[†]

	Study	
	ZIP 1	ZIP 2
Total cholesterol ≥ 200 mg/dl [‡]	-7 (-22, -4)	-8 (-4 to ∞ to 35)
LDL-C ≥ 100 mg/dl [‡]	-19 (-5 to ∞ to 10)	-5 (-3 to ∞ to 26)
HDL-C ≤ 40 mg/dl [‡] in males ≤ 50 mg/dl [‡] in females	-11 (-5 to ∞ to 40)	-11 (-4 to ∞ to 17)
Fasting triglycerides ≥ 150 mg/dl [§]	-4 (-7, -3)	-6 (-69, -3)
Fasting glucose ≥ 100 mg/dl	35 (-14 to ∞ to 8)	-11 (-4 to ∞ to 14)

*Other than weight gain (see Figure 4).

[†]When the NNT or NNH point estimate is not statistically significant, the endpoints of the 95% CI will be opposite in sign, and the interval itself will include infinity.

[‡]To convert total cholesterol, HDL or LDL from mg/dl to mmol/l, multiply by 0.02596.

[§]To convert triglycerides from mg/dl to mmol/l, multiply by 0.0113.

^{||}To convert glucose from mg/dl to mmol/l, multiply by 0.0555.

NNT=number needed to treat; NNH=number needed to harm; CI=confidence interval; LDL-C=low-density lipoprotein cholesterol; HDL-C=high-density lipoprotein cholesterol

cation dosing regimens. Trends in relative benefit or harm may change over time, and our results should be generalized with caution to populations treated for longer or shorter time periods.

Other studies support the validity of our findings. In the Schizophrenia Care and Assessment Program (SCAP), a three-year, longitudinal, observational study of schizophrenia, NNTs to avoid all-cause discontinuation with olanzapine compared to treatment with risperidone, ziprasidone and quetiapine were 13 (7, 379), 6 (-5 to ∞ to 2) and 7 (4, 20), respectively (31). Similarly, data from the ten-country European Schizophrenia Outpatient Health Outcomes (SOHO) trial, a prospective, three-year, observational study in outpatients with schizophrenia, demonstrated after eighteen months that NNTs to avoid discontinuation were 17 (11, 29), 5 (4, 6), and 7 (6, 11) for olanzapine versus risperidone, quetiapine and oral atypical antipsychotics, respectively (32).

Limitations of this analysis encompass those inherent in the five source studies, those related to the design of this analysis, and those associated with the use of NNT and NNH. Discussions of study-specific limitations can be found in the individual published reports (19-23, 33). An obvious limitation of the current analysis is that all data came from studies within the Eli Lilly and Company Clinical Trial Database. This provided access to studies of longer duration in which all available atypicals were evaluated, and in which data for all dichotomous variables of interest were available. Also, though NNT and NNH lend themselves to ranking treatments relative to one another, the design of this analysis limits such comparisons because data were extracted from five different studies comparing olanzapine to a single atypical agent, rather than from a single study involving multiple comparators. Nonetheless, rankings based on studies of the same duration and measuring the same outcomes in patients with similar disease severity may have sufficient validity to be of use to clinicians. The five studies in this analysis shared similar inclusion and exclusion criteria, although two studies included specific subgroups of patients: those with comorbid depression and those with prominent negative symptoms and low functioning. It is possible that the relative efficacies of olanzapine, ziprasidone and quetiapine are different in these subpopulations compared to a more generalized population of patients with schizophrenia. This might partly explain instances where the NNT or NNH was significant for one of the ziprasidone studies, but not for the other, such as was observed for discontinuation due to lack of efficacy, response, avoidance of worsening psychosis and total cholesterol.

Though NNT and NNH are powerful tools rooted in evidence-based medicine, they have limitations. One limitation is that NNT and NNH can be appropriately applied only to dichotomous variables. Consideration of continuous

variables requires establishing cutoff points which may or may not already be well established. Also, NNT and NNH are often presented as a single number, the point estimate. The true value can be higher or lower, and knowledge of the associated CI is crucial in guiding interpretation of reported values. Also, point estimates are calculations based on data representing groups of patients; values cannot be applied to individual patients who differ in their underlying medical risks and responses to medication. When applied to individual patients, NNT and NNH may need to be adjusted for a host of factors including genetic predisposition, prior disease and treatment history, patient preference, caregiver experience and judgment, and resource constraints. Additionally, NNT and NNH are numerical values measuring frequency. Their clinical value depends on the disease being treated, the pain, cost and danger associated with treatment, and likely outcomes with and without treatment. Calculation of NNT and NNH may be simple, but their interpretation and patient-by-patient application is complex.

No single number can accurately characterize the clinical usefulness of a medication. An attempt to base clinical decision making on a mathematical balance between benefit and risk can yield divergent results depending on which NNT and NNH are considered. Some have advocated comparing the NNT for the greatest possible benefit to the NNH for the worst possible harm. However, the possibility for error is amplified when clinical decisions for individual patients are based on a ratio of two group-based figures. A medication's benefit-risk ratio is best approximated through thoughtful consideration of multiple potential benefits and harms in light of individual patient characteristics.

In this analysis, we have shown that atypical antipsychotics differ in their relative likelihood of leading to specific outcomes. The many consequences that flow from these outcomes are not well characterized. Clinicians and other decision makers can benefit from knowing how many patients who discontinue treatment are subsequently hospitalized, and at what personal and economic cost; how many experience future treatment resistance or job loss; and, in what ways and to what degree are caretakers affected. Similarly, questions remain regarding the consequences of treatment-emergent metabolic abnormalities. Future research should establish the frequency with which affected patients develop subsequent diabetes or coronary artery disease, whether pharmacologic intervention can ameliorate these outcomes, and whether this patient population can participate in, and benefit from, nutritional coaching, structured exercise programs or other lifestyle interventions.

Conclusions

Treatment discontinuation is a common problem in the management of schizophrenia, and one with potentially

dangerous consequences. The evidence-based medicine tools of NNT and NNH provide a clinically meaningful assessment of the degree to which, in this post hoc analysis of five antipsychotic comparator trials, olanzapine was superior to ziprasidone and quetiapine, but not risperidone, for prevention of discontinuation for any cause. Ziprasidone was least associated with potentially clinically significant weight gain, followed by risperidone, with olanzapine and quetiapine ranked last.

Acknowledgments

This study was supported by Eli Lilly and Company, Indianapolis, IN, USA. We would like to acknowledge Caron Modeas for editorial assistance and Jennifer Lambert for statistical programming in the preparation of this manuscript.

Competing Interests

Dr. Ball is a scientific writer employed full-time by i3 Statprobe, a division of Ingenix, which is a subsidiary of UnitedHealth Group. Eli Lilly contracted with i3 Statprobe for assistance with this manuscript. All other authors were employees of Eli Lilly and Company at the time of study submission.

Authors Contributions

Conception and Design: Stauffer, Karagianis, Sutton, Ascher-Svanum, Treuer, Silva de Lima, Poole-Hoffmann and Tohen.

Acquisition of Data: Stauffer, Karagianis, Sutton, Treuer and Silva de Lima.

Data Analysis: Sutton.

Data Interpretation: Stauffer, Karagianis, Ascher-Svanum, Treuer, Silva de Lima, Ball, Poole-Hoffmann and Tohen.

Draft of Manuscript: Ball.

All authors were involved in critical revision of this manuscript, and provided final approval prior to submission.

References

1. Beasley CM Jr, Stauffer VL, Liu-Seifert H, Taylor CC, Dunayevich E, Davis JM. All-cause treatment discontinuation in schizophrenia during treatment with olanzapine relative to other antipsychotics: an integrated analysis. *J Clin Psychopharmacol* 2007;27(3):252-258.
2. Dunayevich E, Ascher-Svanum H, Zhao F, Jacobson JG, Phillips GA, Dellva MA, et al. Longer time to antipsychotic treatment discontinuation for any cause is associated with better functional outcomes for patients with schizophrenia, schizophreniform disorder, or schizoaffective disorder. *J Clin Psychiatry* 2007;68(8):1163-1171.
3. Liu-Seifert H, Adams DH, Kinon BJ. Discontinuation of treatment of schizophrenic patients is driven by poor symptom response: a pooled post-hoc analysis of four atypical antipsychotic drugs. *BMC Med* 2005;3:21.
4. Herz MI, Glazer WM, Mostert MA, Sheard MA, Szymanski HV, Hafez H, et al. Intermittent vs maintenance medication in schizophrenia. Two-year results. *Arch Gen Psychiatry* 1991;48(4):333-339.
5. Gilmer TP, Dolder CR, Lacro JP, Folsom DP, Lindamer L, Garcia P, et al. Adherence to treatment with antipsychotic medication and health care costs among Medicaid beneficiaries with schizophrenia. *Am J Psychiatry* 2004;161(4):692-699.
6. Ascher-Svanum H, Zhu B, Faries D, Lacro JP, Dolder CR. A prospective study of risk factors for nonadherence with antipsychotic medication in the treatment of schizophrenia. *J Clin Psychiatry* 2006;67(7):1114-1123.
7. Stroup TS, McEvoy JP, Swartz MS, Byerly MJ, Glick ID, Caine JM. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr Bull* 2003;29(1):15-31.
8. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, et al.; Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 2005;353(12):1209-1223.
9. Gray GE. Concise guide to evidence-based psychiatry. 1st ed. Washington (DC): American Psychiatric Publishing, Inc.; 2004. p. 67-68.
10. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318(26):1728-1733.
11. Belshe RB, Mendelman PM, Treanor J, King J, Gruber WC, Piedra P, et al. The efficacy of live attenuated, cold-adapted, trivalent, intranasal influenza virus vaccine in children. *N Engl J Med* 1998;338(20):1405-1412.
12. Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments: a randomized trial. *Ann Intern Med* 2007;146(12):848-856.
13. Muthu V. The number needed to treat: problems describing non-significant results. *Evid Based Ment Health* 2003;6(3):72.
14. McEvoy JP, Lieberman JA, Stroup TS, Davis SM, Meltzer HY, Rosenheck RA; CATIE Investigators. Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *Am J Psychiatry* 2006;163(4):600-610.
15. Stroup TS, Lieberman JA, McEvoy JP, Swartz MS, Davis SM, Rosenheck RA; CATIE Investigators. Effectiveness of olanzapine, quetiapine, risperidone, and ziprasidone in patients with chronic schizophrenia following discontinuation of a previous atypical antipsychotic. *Am J Psychiatry*

- 2006;163(4):611-622.
16. Citrome L, Stroup TS. Schizophrenia, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and number needed to treat: how can CATIE inform clinicians? *Int J Clin Pract* 2006;60(8):933-940.
 17. Karagianis J, Rosenbluth M, Tohen M, Ascher-Svanum H, Treuer T, de Lima MS, et al. Reviewing CATIE for clinicians: balancing benefit and risk using evidence-based medicine tools. *Curr Med Res Opin* 2007;23(10):2551-2557.
 18. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000.
 19. Keefe RS, Young CA, Rock SL, Purdon SE, Gold JM, Breier A; HGGN Study Group. One-year double-blind study of the neurocognitive efficacy of olanzapine, risperidone, and haloperidol in schizophrenia. *Schizophr Res* 2006;81(1):1-15.
 20. Tran PV, Hamilton SH, Kuntz AJ, Potvin JH, Andersen SW, Beasley C Jr, et al. Double-blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *J Clin Psychopharmacol* 1997;17(5):407-418.
 21. Breier A, Berg PH, Thakore JH, Naber D, Gattaz WF, Cavazoni P, et al. Olanzapine versus ziprasidone: results of a 28-week double-blind study in patients with schizophrenia. *Am J Psychiatry* 2005;162(10):1879-1887.
 22. Kinon BJ, Lipkovich I, Edwards SB, Adams DH, Ascher-Svanum H, Siris SG. A 24-week randomized study of olanzapine versus ziprasidone in the treatment of schizophrenia or schizoaffective disorder in patients with prominent depressive symptoms. *J Clin Psychopharmacol* 2006;26(2):157-162.
 23. Kinon BJ, Noordsy DL, Liu-Seifert H, Gulliver AH, Ascher-Svanum H, Kollack-Walker S. Randomized, double-blind 6-month comparison of olanzapine and quetiapine in patients with schizophrenia or schizoaffective disorder with prominent negative symptoms and poor functioning. *J Clin Psychopharmacol* 2006;26(5):453-461.
 24. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;13(2):261-276.
 25. Andreasen NC, Carpenter WT Jr, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry* 2005;162(3):441-449.
 26. Beasley CM Jr, Sutton VK, Hamilton SH, Walker DJ, Dossenbach M, Taylor CC; Olanzapine Relapse Prevention Study Group. A double-blind, randomized, placebo-controlled trial of olanzapine in the prevention of psychotic relapse. *J Clin Psychopharmacol* 2003;23(6):582-594.
 27. Barnes TR. A rating scale for drug-induced akathisia. *Br J Psychiatry* 1989;154:672-676.
 28. Simpson GM, Angus JW. A rating scale for extrapyramidal side effects. *Acta Psychiatr Scand Suppl* 1970;212:11-19.
 29. Guy W. ECDEU assessment manual for psychopharmacology. DHEW publication, no. (ADM) 76-338. Rockville (MD): U. S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs, 1976. p. 534-537.
 30. Davis JM, Chen N, Glick ID. A meta-analysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry* 2003;60(6):553-564.
 31. Ascher-Svanum H, Zhu B, Faries D, Landbloom R, Swartz M, Swanson J. Time to discontinuation of atypical versus typical antipsychotics in the naturalistic treatment of schizophrenia. *BMC Psychiatry* 2006;6:8.
 32. Novick D, Suarez D, Ascher-Svanum H, Karagianis J, Perrin E, Haro JM, et al. Number needed to treat (NNT) for all-cause medication discontinuation in CATIE compared to the Schizophrenia Outpatients Health Outcomes (SOHO) study. *Eur Psychiatry* 2007;22(Suppl 1):S130.
 33. Schooler NR, Tollefson GD, Tran PV, Gheuens J, Greb JA. Letter on article by Tran and Associates, "Double-blind comparison of olanzapine versus risperidone in treatment of schizophrenia and other psychotic disorders." *J Clin Psychopharmacol* 1998;18(2):174-179