



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

STATISTICAL REVIEW AND EVALUATION

NDA: 19,758 (SE1-047)

DRUG NAME: Clozaril ® (clozapine) Tablets

INDICATION: Reduction of Suicidality in Patients with Schizophrenia or Schizoaffective disorder who are at risk for suicide

SPONSOR: Novartis Pharmaceuticals Corporation

STATISTICAL REVIEWER: Kun He (HFD-710)

DATE OF DOCUMENT: February 28, 2002

HFD-710 Kun Jin, Ph.D., Team Leader
George Chi, Ph.D., Director

HFD-120 Steven Hardeman, R. Ph., Project Manager
Gregory Dubitsky, M.D., Clinical Reviewer
Thomas Laughren, M.D., Team Leader
Russell Katz, M.D., Director

HFD-700 Charles Anello, Sc. D., Deputy Director

TABLE OF CONTENTS

0. Executive Summary 3

0.1. Brief Overview of Clinical Studies..... 3

0.2 Issues and Conclusion 4

1. Introduction..... 5

2. Study ABA 451 5

2.1. Objective 5

2.2. Study Design 5

2.3. Efficacy Measures 6

2.4. Statistical Analysis Plan..... 7

2.5. Study Population 8

2.6. Sponsor’s Efficacy Results..... 11

3. Reviewer’s Analysis 13

3.1. The Issue of Open-Label..... 13

3.2. The Issue of Censoring..... 15

3.3. Retrieved Dropout..... 16

3.4. Subgroup Analyses..... 18

4. Conclusion 21

Statistical Review and Evaluation

0. Executive Summary

0.1. Brief Overview of Clinical Studies

The current submission NDA 19-758 for Clozaril ® (clozapine HCL) Tablets includes one study to demonstrate decreased risk for suicide among schizophrenic patients treated with Clozaril as compared to those treated with Zyprexa. This was a multicenter, randomized, open-label, 24-month study, using Zyprexa 5 to 20 mg/day as active control treatment or using Clozaril 200 to 900 mg/day as the treatment, conducted in US, Canada, U.K., France, Italy, Hungary, Croatia, South Africa, Czech Republic, Argentina, and Chile. A total of 980 males or females aged 18-65 and diagnosed with schizophrenia or schizoaffective disorder using DSM-IV criteria and who were deemed at high risk for suicide were included and randomized equally to two treatment groups.

The primary endpoint is the time (in days, after randomization) to either of a Type 1 Event which is defined as occurrence of a significant suicide attempt or hospitalization due to imminent suicide risk (including increased level of surveillance) as confirmed by a Suicide Monitoring Board (SMB); or a Type 2 Event which is defined as occurrence of a worsening of severity of suicidality as demonstrated by a score of 6 or 7 (much worse or very much worse) on the change score of the Clinical Global Impression for Severity of Suicidality (CGI-SS-BP), rated by the Blinded Psychiatrist (BP), or an implicit worsening of severity of suicidality as indicated by the occurrence of a Type 1 event.

The Principal Investigator (PI) is unblinded to the randomized treatment. If a complete suicide, suicide attempt, or hospitalization for imminent risk of suicide is reported during the study, the Principal Investigator (PI) will prepare a brief narrative, and complete a Suicide Attempt Form (SAF) or a Imminent Risk Requiring Hospitalization Form (IRH), and send the information to Worldwide Clinical Trials (WCT) medical monitor within 24 hours. The WCT medical monitor will send the information to the Suicide Monitoring Board (SMB) and Blinded Psychiatrist (BP). The SMB and BP will independently evaluate and complete a Suicide Event Form, respectively, regarding whether it met a primary efficacy outcome criterion. The primary efficacy endpoint of Type 1 event is determined by the SMB.

CGI-SS will be formally evaluated at baseline, Week 2, Week 4 and every 4 weeks up to Week 52, and then at Weeks 60, 68, 80, 92, and 104 by the PI; and at baseline and at Weeks 8, 16, 24, 32, 40, 48, 52, 60, 68, 80, 92, and 104 by the BP. Part of CGI-SS by the BP is used in defining a Type 2 event.

0.2 Issues and Conclusion

The key issue is whether there is any bias caused by open-label as the Principal Investigator (PI) is unblinded. P-values on CGI-SS for the PI are always smaller than that for the BP, .0946 for the PI and .2803 for the BP in 7-point CGI-SS, and .2767 for the PI and .8708 for the BP in 5-point CGI-SS. The difference between p-values indicates that the PI might have bias favoring Clozaril.

The Type 1 event is determined by the SMB but it is the PI who determines who should be referred to the SMB. There are 122 patients in Clozaril and 157 patients in Zyprexa, respectively, who were referred to the SMB. The difference between the number of referred patients is 35. However, there are 102 patients in Clozaril and 141 patients in Zyprexa, respectively, who were judged by the SMB as the Type 1 event. The difference of the number of Type 1 event is 39. It is also noted that the correlation between the number of referred and the number of event is high. Based on the above, it is seen that the less referred will have the less Type 1 event, or the number of event is determined by the number of referred. Consequently this affects the number of censored patients.

In this study, for Type 1 event, there are 102 failed, 352 censored, and 36 completed in Clozaril, and 141 failed, 309 censored, and 40 completed in Zyprexa, respectively. The significance of analysis based on Type 1 event is due to 39, the difference of the number of event. The significance of analysis based on Type 2 event is mainly due to the contribution of Type 1 event because analysis on CGI-SS is not statistically nominally significant.

The conclusion is that although the primary analysis is statistically nominally significant with p-value .0309, one should interpret the result with caution due to the issues discussed above.

1. Introduction

The current submission NDA 19-758 for Clozaril® (clozapine HCL) Tablets includes one study to demonstrate decreased risk for suicide among schizophrenic patients treated with Clozaril as compared to that for patients treated with Zyprexa.

This was a multicenter, randomized, open-label, 24-month study, using Zyprexa 5 to 20 mg/day as active control treatment or using Clozaril 200 to 900 mg/day as the treatment under study, conducted in US, Canada, U.K., France, Italy, Hungary, Croatia, South Africa, Czech Republic, Argentina, and Chile. A total of 980 males or females aged 18-65 and diagnosed with schizophrenia or schizoaffective disorder using DSM-IV criteria and who were deemed at high risk for suicide were included and randomized equally to two treatment groups.

2. Study ABA 451

2.1. Objective

The primary efficacy objective was to demonstrate decreased risk for suicide among schizophrenic patients treated with Clozaril as compared to that treated with Zyprexa using multiple events analysis techniques applied to time (in days, after randomization) to either of a Type 1 Event which is defined as occurrence of a significant suicide attempt or hospitalization due to imminent suicide risk (including increased level of surveillance) as confirmed by a Suicide Monitoring Board (SMB); or a Type 2 Event which is defined as occurrence of a worsening of severity of suicidality as demonstrated by a score of 6 or 7 (much worse or very much worse) on the change score of the Clinical Global Impression for Severity of Suicidality (CGI-SS-BP), rated by the Blinded Psychiatrist (BP), or an implicit worsening of severity of suicidality as indicated by the occurrence of a Type 1 event.

2.2. Study Design

This was a prospective, randomized, open-label, 24-month study with an active control treatment, designed to evaluate the effects of Clozaril and Zyprexa on suicidality in patients with schizophrenia and schizoaffective disorder, who are known to be at high risk for suicide.

A total of 980 males or females patients aged 18-65 years (inclusive) and diagnosed with schizophrenia or schizoaffective disorder using DSM-IV criteria and who were deemed at high risk for suicide were included in the trial. Patients were randomized to treatment with either Clozaril or Zyprexa in an approximate 1:1 ratio within each center. The suggested dosage for patients assigned to Clozaril was 200 to 900 mg/day and for patients assigned to Zyprexa was 5 to 20 mg/day. Throughout the study, prescribed doses were based on a clinical assessment of tolerability and efficacy.

At each medication dispensing visit to the clinic (weekly for the first 26 weeks and thereafter, every 2 weeks for the remainder of the study), the patient's overall psychiatric condition was assessed by a health care professional. On the basis of this assessment, the need for a referral to the Principal Investigator (PI)/sub-investigator for psychiatric evaluation and treatment was considered. At designated visits, the patients' suicidality were evaluated by the PI and the Blinded Psychiatrist (BP) on the CGI-SS (CGI-SS-PI, and CGI-SS-BP) and on the InterSePT Scale for Suicidal Thinking (ISST-PI, and ISST-BP).

2.3. Efficacy Measures

The primary efficacy variable was time (in days, after randomization) to first occurrence of any of the following events: a Type 1 event, which included a significant suicide attempt or hospitalization due to imminent suicide risk including an increased level of surveillance as confirmed by the Suicide Monitoring Board (SMB); or a Type 2 event, defined as worsening of severity of suicidality as demonstrated by a 7-point CGI-SS-BP change scale score of 6 or 7 or an implicit worsening of severity of suicidality as indicated by the occurrence of a Type 1 event.

Completed Suicides: all deaths that occur during the study will require full documentation. Hospital summaries, autopsy reports, and coroner's reports will be collected and utilized by the PI to prepare a narrative which includes all relevant details but deletes information relating to the patient's treatment prior to sending it to the Worldwide Clinical Trials (WCT) medical monitor. If any required material is unavailable the PI must notify the WCT medical monitor in writing. A Suicide Attempt Form (SAF) must also be completed and sent to the WCT medical monitor. All information must be reviewed by the WCT medical monitor before it is sent to both the BP at the site and the SMB that is constituted of three experts in the study. Both will make independent determinations as to whether the death was a suicide and will complete a Suicide Event Form-Blinded Psychiatrist (SEF-BP)/SEF-SMB. Any requests for further information by the SMB or BP must be made through the WCT medical monitor.

Suicide Attempts: during the study, if a suicide attempt is reported, the PI will contact the relevant hospital/physician and obtain all necessary documents to objectively categorize the attempt as life-threatening or non-life-threatening, recording such information as type of event (e.g., violent/non-violent), treatment received (e.g., gastric lavage, surgery, stitches), name of hospital/physician providing care. Using this information, the PI will complete an SAF with a brief narrative and forward the information to the WCT medical monitor within 24 hours of learning of the event. The WCT medical monitor will ensure that the information will not compromise the blinding of the BP at the site or the SMB. The BP and the SMB will make independent decisions as to whether the attempt was serious, will categorize its lethality and will record this information on the SEF-BP and SEF-SMB, respectively.

Hospitalization for Imminent Risk of Suicide: the PI will notify the WCT medical monitor within 24

hours of learning that a patient has been hospitalized for imminent risk of suicide. However, hospitalizations driven more by logistical concerns (e.g., patient's residence being very distant from the site, mild increases of depression, or lack of caregiver) will not be considered as imminent risk and as such the WCT medical monitor will not need to be notified. Whenever a patient has been admitted to a hospital for imminent risk of suicide reason the PI will contact the WCT medical monitor within 24 hours of learning of the hospitalization and will complete the Imminent Risk Requiring Hospitalization Form (IRH) which will include a brief narrative. Information about patients who are hospitalized for imminent risk of suicide will be sent by the WCT medical monitor to the BP and to the SMB. They will independently evaluate the hospitalization and complete an SEF-BP and SEF-SMB regarding whether it met a primary efficacy outcome criterion.

CGI-SS is a 2-part assessment that evaluated patient suicidality. In the first part, which is referred to throughout the report as the 5-point CGI-SS severity score, the rater chose the most severe level of suicidality experienced by the patient over the previous 7 days on a 5-point scale ranging from 1 (not at all suicidal) to 5 (attempted suicide). In the second part, which is referred to as the 7-point CGI-SS change score, the rater assessed how much the patient's suicidality had changed since baseline on a 7-point scale ranging from 1 (very much improved) to 7 (very much worse). The patient's suicidality will be formally evaluated at baseline, Week 2, Week 4 and every 4 weeks up to Week 52, and then at Weeks 60, 68, 80, 92, and 104 by the PI; and at baseline and at Weeks 8, 16, 24, 32, 40, 48, 52, 60, 68, 80, 92, and 104 by the BP.

The endpoints defined in the original protocol (dated 1/16/98) for Study ABA 451 were either time from baseline until first significant suicide attempt or hospitalization due to imminent risk of suicide confirmed by the blinded SMB; or change from baseline in the CGI-SS-BP severity score (5-point scale).

2.4. Statistical Analysis Plan

The primary analysis is WLW (Wei, Lin, and Weissfeld (1989)) method, submitted in Amendment 6 dated 12/19/2000. The WLW method is a semiparametric method used to analyze multivariate failure time data, and models the marginal distribution with a Cox's proportional hazards model without imposing any particular structure of dependence on each event time. Following the approach of WLW for the primary analysis, time to the first occurrence of a Type 1 event and time to the first occurrence of a Type 2 event were modeled using a proportional hazards model, with pooled country as strata and treatment group as the only covariate. The WLW method provided estimates of treatment effects for the 2 types of events, the combined estimator of treatment effect is defined as:

$$\hat{\beta}_c = c\hat{\beta}_1 + c\hat{\beta}_2.$$

where $\hat{\beta}_1$ and $\hat{\beta}_2$ represent the maximum partial likelihood estimators of β_1 and β_2 , respectively. Based on negotiations with the FDA, $c=0.5$ will be used in the analysis. Using this estimate of the combined treatment effect and the corresponding robust standard error, a single two-sided test was performed at a 5% level of significance using the null hypothesis of no treatment effect on either

event type. The test statistic ($T = \text{estimate}/\text{SE}$) has an asymptotic standard normal distribution. Only the ITT population was used for the analysis of the primary efficacy variables. In this report, the analyses of revised primary efficacy variables suggested by the FDA are presented as the main analyses for primary efficacy.

Due to the revision of the primary efficacy variables in this study based on an agreement with the FDA prior to database lock, the analyses defined in the original protocol (dated 1/16/98) were considered as supportive analyses for primary efficacy. The main analysis defined in the original protocol (dated 1/16/98) of the time to suicide attempt or hospitalization to prevent suicide (later called Type 1 event) was to be performed based on Cox's proportional hazards regression model. Explanatory variables in this model included treatment and the following baseline measurements: number of lifetime suicide attempts, active substance/alcohol abuse, pooled country, sex, and age group (at 3 levels: 18-32 years, 33-44 years, and ≥ 45 years). The main analysis of the change from baseline in the CGI-SS-BP severity score (5-point scale) was to be performed using an analysis of covariance (ANCOVA) model with the following explanatory variables: treatment, number of suicide attempts, active substance/alcohol abuse, country, sex, age group (at 3 levels: 18-32 years, 33-44 years, and ≥ 45 years), and baseline CGI-SS-BP severity score.

2.5. Study Population

2.5.1. Patient disposition

A total of 980 patients were randomized in Study ABA 451. Of these, 956 (97.6%) actually received study medication, 379 (38.7%) discontinued early, and 671 (68.5%) completed the study. A summary of patient entry to the study and completion by treatment group, including reasons for withdrawals and status of retrieved dropout patients, is given in Table 2.5.1. (adapted from Study report, Table 7-1).

Table 2.5.1. Patient disposition for each treatment group

	Clozaril	Zyprexa
Patient disposition	No. (%) of patients	
Screened: 1065 total		
Randomized	490 (100%)	490 (100%)
Did not receive study drug	10 (2.0%)	11 (2.2%)
Dispensed study drug	480 (98.0%)	479 (97.8%)
Took study drug	479 (97.8%)	477 (97.3%)
Completed ¹	332 (67.8%)	339 (69.2%)
Total discontinuations from the study	192 (39.2%)	187 (38.2%)
Reason for discontinuation:		
Adverse event	41 (8.4%)	33 (6.7%)
Abnormal laboratory value	2 (0.4%)	0 (0.0%)
Abnormal test procedure result	1 (0.2%)	0 (0.0%)
Unsatisfactory therapeutic effect on psychosis	5 (1.0%)	9 (1.8%)
Unsatisfactory therapeutic effect on suicide risk	0 (0.0%)	6 (1.2%)
Protocol violation	29 (5.9%)	20 (4.1%)
Patient withdrew consent	50 (10.2%)	49 (10.0%)
Lost to follow-up	33 (6.7%)	39 (8.0%)
Administrative problems	23 (4.7%)	26 (5.3%)
Death ²	8 (1.6%)	5 (1.0%)
Retrieved Dropouts	61 (12.4%)	60 (12.2%)
Completed ³	34 (55.7%)	37 (61.7%)
Discontinued	27 (44.3%)	23 (38.3%)

¹Includes subjects whose Study Completion form indicated that the subject had completed and Retrieved Dropout (RDO) subjects whose last assessment occurred after Week 102 in relation to randomization date.

²Includes only discontinuations for death, a more complete summary of deaths can be found in Section 10.2.1.1.

³Includes RDO subjects whose last assessment occurred after Week 102 in relation to randomization date. Completed and discontinued percentages use total number of RDO subjects in denominator.

Source: Post-Text Table 7.1-1

Withdrawal of consent was the most common reason for discontinuation in both groups, occurring in 10.2% and 10.0% of Clozaril and Zyprexa patients, respectively. Reasons for discontinuation did not differ significantly between the two treatment groups.

2.5.2. Baseline demographic and background characteristics

Demographic and key baseline characteristics are summarized for the ITT population in Tables 2.5.2.1 and 2.5.2.2 (adapted from Study report, Tables 7-3 and 7-4).

Table 2.5.2.1. Summary of demographic characteristics at baseline by treatment group

	No. (%) of patients unless otherwise noted	
	Clozaril (n=490)	Zyprexa (n=490)
Age (year)		
Mean (SD)	37.1 (10.3)	37.0 (10.3)
Median	37	36
Range	18-69	18-65
18-32	168 (34.3%)	178 (36.3%)
33-44	216 (44.1%)	204 (41.6%)
≥45	106 (21.6%)	108 (22.0%)
Sex		
Male	301 (61.4%)	301 (61.4%)
Female	189 (38.6%)	189 (38.6%)
Race		
Caucasian	356 (72.7%)	337 (68.8%)
Black	65 (13.3%)	86 (17.6%)
Oriental	6 (1.2%)	7 (1.4%)
Other	63 (12.9%)	60 (12.2%)
Weight (kg)– Females		
	n=181	n=180
Mean (SD)	74.0 (20.1)	73.2 (18.4)
Median	70	70
Range	40-152	36-133
Weight (kg)– Males		
	n=283	n=289
Mean (SD)	82.8 (18.3)	84.3 (20.9)
Median	80.9	80
Range	45-156	44-166

Source: Post-Text Table 7.4-1

The treatment groups were comparable for sex, age, weight, and race. There were more males than females in the study (61.4% of all patients were male).

Table 2.5.2.2. Disease characteristics and lifetime suicide history at baseline by treatment group

	Clozaril (n=490)	Zyprexa (n=490)
Diagnosis	No. (%) of patients	
Schizophrenia	300 (61.2%)	309 (63.1%)
Schizoaffective	190 (38.8%)	181 (36.9%)
Treatment resistant ¹	135 (27.6%)	128 (26.1%)
Suicide History		
No. of lifetime suicide attempts	(n=490)	(n=489)
0	77 (15.7%)	86 (17.6%)
1	124 (25.3%)	99 (20.2%)
2-3	154 (31.4%)	157 (32.0%)
4-5	60 (12.2%)	75 (15.3%)
>5	75 (15.3%)	72 (14.7%)
Missing	0 (0.0%)	1 (0.2%)
Mean (SD)	3.6 (7.5)	3.2 (4.5)
Median	2	2
Range	0-120	0-50
No. of lifetime hospitalizations to prevent a suicide attempt	(n=487)	(n=483)
0	79 (16.1%)	75 (15.3%)
1	126 (25.7%)	145 (29.6%)
2-3	145 (29.6%)	135 (27.6%)
4-5	63 (12.9%)	52 (10.6%)
>5	74 (15.1%)	76 (15.5%)
Missing	3 (0.6%)	7 (1.4%)
Mean (SD)	3.7 (7.7)	3.2 (4.8)
Median	2	2
Range	0-100	0-50

¹ Treatment resistance determined by clinical assessment

Source: Post-Text Tables 7.1-1 (diagnosis) and 7.4-2 (other characteristics)

2.6. Sponsor's Efficacy Results

2.6.1. Primary efficacy results

The main analyses for all efficacy variables were performed on the ITT population. The primary analysis was WLW analysis of time to the first occurrence of a Type 1 event or a Type 2 (adapted from Study report Table 9-1).

Table 2.6.1. Primary analysis: Multiple event analysis of time to first occurrence of Type 1 and Type 2 events

Event Type ¹	Coefficient of Treatment Effect (Beta ^{2,3}) (SE)	p-value ²	Hazard Ratio ^{2,3}	95% C.I. for Hazard Ratio ²
Type 1	-0.280 (0.130)	0.0316	0.76	0.58, 0.98
Type 2	-0.250 (0.121)	0.0388	0.78	0.61, 0.99
Combined	-0.265 (0.123)	0.0309	--	--

¹Type 1 event = a significant suicide attempt or hospitalization due to imminent suicide risk (including increased level of surveillance), confirmed by SMB.
²Type 2 event = worsening of suicidality severity as demonstrated by 7-point CGI-SS-BP change scale score of 6 or 7, or by implicit worsening of suicidality severity as demonstrated by occurrence of a Type 1 event.
²Refer to detailed statistical analysis plan in Appendix 5.1, §1.1.5.3, for information on calculation of these parameters.
³Hazard ratio < 1 and beta < 0 indicate that Clozaril is better than Zyprexa.

Source: Post-Text Table 9.1-1

As shown above, an overall treatment effect favoring Clozaril was statistically significant with p-value 0.0309.

2.6.2. Original Protocol Analyses

The analyses defined in the original protocol are presented in Table 2.6.2. (adapted from Study report Table 9-2).

Table 2.6.2. Per protocol primary analysis: Analysis of time to first occurrence of Type 1 event and analysis of change from baseline in 5-point CGI-SS-BP severity score

Event Type	Regression Coefficient for Treatment (SE) ²	p-value ³	95% C.I. of Regression Coefficient ²	HR ²	95% C.I. for HR ²
Type 1 event ¹	-0.304 (0.132)	0.0211	--	0.74	0.57, 0.96
5-point CGI-SS-BP severity score	0.007 (0.048)	0.8884	-0.09, 0.10	--	--

¹Type 1 event = a significant suicide attempt or hospitalization due to imminent suicide risk (including increased level of surveillance), confirmed by SMB.
²Refer to detailed statistical analysis plan in Appendix 5.1, §1.1.5.3, for information on calculation of these parameters.
³P-value for Type 1 event was generated using a full Cox's proportional hazards regression model; p-value for 5-point CGI-SS-BP severity score was generated using an ANCOVA model

Source: Post-Text Tables 9.1-3 and 9.1-4

P-values are .0211 for Type 1, and .8884 for 5-point CGI-SS-BP, respectively.

2.6.3. Survival analyses

Kaplan-Meier estimates of cumulative probabilities for Type 1 and Type 2 events were estimated for the two treatment groups and are summarized for the ITT population in Table 2.6.3. (adapted from Study report Table 9-4).

Table 2.6.3.1. Kaplan-Meier estimates for the cumulative probability of a Type 1 or Type 2 event by visit (ITT population)

	CLOZARIL (n=490)				ZYPREXA (n=490)				95% C.I. of the difference
	n1 ²	N2 ³	Cum Prob	95% C.I.	n1 ²	n2 ³	Cum Prob	95% C.I.	
Type 1 Event									
Week 0 (Day 0)	490	0	0.00	(0.00, 0.00)	490	0	0.00	(0.00, 0.00)	(0.00, 0.00)
Week 8 (Day 70)	393	43	0.09	(0.09, 0.10)	411	50	0.11	(0.10, 0.11)	(-0.03, 0.05)
Week 24 (Day 182)	346	69	0.16	(0.15, 0.17)	365	81	0.17	(0.16, 0.18)	(-0.03, 0.07)
Week 52 (Day 378)	308	91	0.21	(0.20, 0.22)	312	112	0.25	(0.23, 0.26)	(-0.02, 0.09)
Week 80 (Day 574)	277	100	0.23	(0.22, 0.25)	269	128	0.29	(0.27, 0.30)	(-0.01, 0.11)
Week 104 (Day 742)	36	102	0.24	(0.23, 0.25)	40	141	0.32	(0.31, 0.34)	(0.02, 0.14)
Type 2 Event									
Week 0 (Day 0)	490	0	0.00	(0.00, 0.00)	490	0	0.00	(0.00, 0.00)	(0.00, 0.00)
Week 8 (Day 70)	389	47	0.10	(0.10, 0.11)	410	51	0.11	(0.10, 0.11)	(-0.03, 0.04)
Week 24 (Day 182)	334	81	0.18	(0.17, 0.20)	356	90	0.19	(0.18, 0.21)	(-0.04, 0.06)
Week 52 (Day 378)	290	109	0.25	(0.24, 0.27)	294	132	0.29	(0.28, 0.31)	(-0.02, 0.10)
Week 80 (Day 574)	261	119	0.28	(0.27, 0.30)	251	150	0.34	(0.32, 0.35)	(0.00, 0.12)
Week 104 (Day 742)	34	120	0.28	(0.27, 0.30)	36	161	0.37	(0.35, 0.38)	(0.02, 0.15)

¹ Kaplan-Meier estimates compute the probability of an event (cumulative). Two weeks were added to the visit week when calculating the actual day, e.g., Visit Week 8 = Day (8+2)x7 = Day 70.
² n1 represents number of patients at risk.
³ n2 represents the number of cumulative events.

Source: Post-Text Table 9.1-2

The log-rank tests give p-values .0195 for Type 1, and .027 for Type 2, respectively.

3. Reviewer's Analysis

The reviewer verified the sponsor's efficacy analyses according to the protocol. In this section, the issues of open-label and censoring, retrieved dropouts, and subgroup will be performed on the ITT population.

3.1. The Issue of Open-Label

Both 7-point CGI-SS and 5-point CGI-SS for the PI and BP will be discussed. The relation between

the number of referred and the number of Type 1 event will also be discussed.

3.1.1. CGI-SS

Table 3.1.1 gives summary information for 5-point CGI-SS (Study report, Appendix 5.3.1 Table 9.3-21). The p-value is calculated from ANOVA with terms for treatment, pooled country, and CGI-SS at baseline. The percentage is calculated using actual number of patients.

Table 3.1.1. Summary for 5-point CGI-SS (LOCF)

At end of study	Blinded Psychiatrist (BP)		Principal Investigator (PI)	
	Clozaril (n=490)	Zyprexa (n=490)	Clozaril (n=490)	Zyprexa (n=490)
n	484	481	482	480
mean change from baseline \pm sd	-0.812 \pm 1.1	-0.807 \pm 1.1	-1.102 \pm 1.1	-0.983 \pm 1.1
p-value	.8708		.2767	
Distribution				
-4	3 (0.62%)	2 (0.42%)	2 (0.41%)	3 (0.63%)
-3	33 (6.24%)	30 (6.82%)	41 (8.51%)	29 (6.04%)
-2	87 (17.98%)	102 (21.21%)	143 (29.67%)	129 (26.88%)
-1	131 (27.07%)	121 (25.16%)	130 (26.97%)	140 (29.17%)
0	214 (44.21%)	201 (41.79%)	154 (31.95%)	158 (32.92%)
1	13 (2.69%)	17 (3.53%)	10 (2.07%)	17 (3.54%)
2	1 (0.21%)	6 (1.25%)	1 (0.21%)	4 (0.83%)
3	0	2 (0.42%)	0	0
4	2 (0.41%)	0	1 (0.21%)	0

Table 3.1.2 gives summary information for 7-point CGI-SS (Study report, Appendix 5.3.1 Table 9.3-15). The p-value is calculated from ANOVA with terms for treatment and pooled country.

Table 3.1.2. Summary for 7-point CGI-SS (LOCF)

At end of study	Blinded Psychiatrist (BP)		Principal Investigator (PI)	
	Clozaril (n=490)	Zyprexa (n=490)	Clozaril (n=490)	Zyprexa (n=490)
n	422	440	464	468
mean \pm sd	2.64 \pm 1.4	2.73 \pm 1.4	2.40 \pm 1.3	2.54 \pm 1.4
p-value	.2803		.0946	
Distribution				
1 = very much improved	125 (29.62%)	123 (27.95%)	170 (36.64%)	155 (33.12%)
2 = much improved	86 (20.38%)	90 (20.45%)	100 (21.55%)	105 (22.44%)
3 = minimally improved	49 (11.61%)	49 (11.14%)	52 (11.21%)	49 (10.47%)
4 = no change	145 (34.36%)	150 (34.09%)	123 (26.51%)	131 (27.99%)
5 = minimally worse	13 (3.08%)	16 (3.64%)	18 (3.88%)	18 (3.85%)
6 = much worse	3 (0.71%)	11 (2.5%)	1 (0.22%)	7 (1.5%)
7 = very much worse	1 (0.24%)	1 (0.23%)	0	3 (0.64%)

Since p-values for the PI are always smaller than that for the BP, this indicates that the PI might have bias favoring Clozaril.

3.1.2. Number of Referred (PI) and Event (SMB)

If there is a completed suicide, suicide attempt, or hospitalization for imminent risk of suicide, it is the PI who will prepare a brief narrative, complete SAF or IRH, and report to the SMB through WCT.

The Type 1 event is determined by the SMB but it is the PI who determines who should be referred to the SMB. There are 122 patients in Clozaril and 157 patients in Zyprexa, respectively, who were referred to the SMB. The difference between the number of referred patients is 35. However, there are 102 patients in Clozaril and 141 patients in Zyprexa, respectively, who were judged by the SMB as the Type 1 event. The difference of the number of Type 1 event is 39. It is also noted that the correlation between the number of referred and the number of event is high. Based on the above, it is seen that the less referred will have the less Type 1 event, or the number of event is determined by the number of referred.

A patient might be referred multiple times. 122 patients in Clozaril were referred 261 times, in which 217 times were judged by the SMB as Type 1 event, and 44 times were not. 157 patients in Zyprexa were referred 316 times, in which 266 times were judged by the SMB as Type 1 event, and 50 times were not.

The above analyses indicate that if the PI might have bias favoring Clozaril and the PI's referral determines the number of event, then the open-label does introduce the bias.

3.2. The Issue of Censoring

The time to treatment failure (TTF) and time to lost to follow-up will be discussed.

3.2.1. Time to Treatment Failure

The time to treatment failure (TTF) is defined from date of randomization until the date that a patient withdrew (failed or censored) from the study. A patient who completed the study will be censored at the end of study.

For Type 1, there are 454 failed and 36 completed in Clozaril, and 450 failed and 40 completed in Zyprexa, respectively. The log-rank test for TTF gives p-value .6845. The survival probabilities from the KM estimates at 104 weeks are .0735 for Clozaril and .0816 for Zyprexa, respectively.

For Type 2, there are 456 failed and 34 completed in Clozaril, and 454 failed and 36 completed in

Zyprexa, respectively. The log-rank test for TTF gives p-value .7877. The survival probabilities from the KM estimates at 104 weeks are .0694 for Clozaril and .0735 for Zyprexa, respectively.

3.2.2. Lost to follow-up

For Type 1, lost to follow-up are 33 (8 failed and 25 censored) in Clozaril, and 39 (13 failed and 26 censored) in Zyprexa, respectively. Table 3.2.2.1 gives number of failed, censored, and mean and standard deviation of the primary efficacy variable TIMETO (time to event).

Table 3.2.2.1 Lost to Follow-up

Type=1	Clozaril (n=33)	TIMETO Mean (sd)	Zyprexa (n=39)	TIMETO Mean (sd)
failed	8	97 (101)	13	163 (198)
censored	25	169 (233)	26	241 (226)

After counting the censored in the lost to follow-up as failed, there are 127 failed, 327 censored, and 36 completed in Clozaril, and 167 failed, 283 censored, and 40 completed in Zyprexa, respectively. The log-rank test gives p-value .031. The survival probabilities from the KM estimates at 104 weeks are .6985 for Clozaril and .6296 for Zyprexa, respectively.

For Type 2, after counting the censored in the lost to follow-up as failed, there are 145 failed, 311 censored, and 34 completed in Clozaril, and 187 failed, 267 censored, and 36 completed in Zyprexa, respectively. The log-rank test gives p-value .0387. The survival probabilities from the KM estimates at 104 weeks are .6571 for Clozaril and .5859 for Zyprexa, respectively.

3.3. Retrieved Dropout

The sponsor used the retrieved dropout (RDO) assessment in defining ITT. The retrieved dropout assessments were made on patients who withdrawn prematurely from the study, and were collected and recorded in the Retrieved Drop Out CRF Booklet. It is possible that a patient might not be still on the treatment between the withdrawn date and the retrieved date.

For Type 1, there are 61 in Clozaril and 60 in Zyprexa, respectively, who are identified as RDO. Table 3.3.1 gives mean and standard deviation of TIMETO and DUREOS (end of study duration days).

Table 3.3.1. Retrieved Dropout

Type=1	Clozaril (n=61)	TIMETO Mean (sd)	DUREOS Mean (sd)	Zyprexa (n=60)	TIMETO Mean (sd)	DUREOS Mean (sd)
failed	20	176 (175)	145 (153)	25	143 (162)	266 (250)
censored	33	553 (204)	134 (143)	24	558 (224)	142 (165)
completed	8	765 (22)	372 (286)	11	764 (18)	214 (199)

If the end of study date is used instead of the retrieved date, then the smaller of TIMETO and DUREOS will be used as the time to event. Furthermore, if a failed patient's TIMETO were greater than DUREOS, then the patient would be classified as censored.

There are 12 failed in Clozaril and 3 failed in Zyprexa, respectively, whose TIMETO are greater than DUREOS. After counting those 12 failed and 8 completed in Clozaril and 3 failed and 11 completed in Zyprexa as censored, respectively, and using the smaller of TIMETO and DUREOS as the time to event, Table 3.3.2 gives summary information.

Table 3.3.2. DUREOS

Type=1	Clozaril (n=61)	TIMETO Mean (sd)	DUREOS Mean (sd)	Zyprexa (n=60)	TIMETO Mean (sd)	DUREOS Mean (sd)
failed	8	62 (50)	225 (186)	22	120 (139)	294 (254)
censored	53	161 (185)	161 (185)	38	156 (172)	156 (172)
completed	0			0		

Using Table 3.3.2, there are 90 failed, 372 censored, and 28 completed in Clozaril, and 138 failed, 323 censored, and 29 completed in Zyprexa, respectively. The log-rank test gives p-value .0035. The survival probabilities from the KM estimates at 104 weeks are .7749 for Clozaril and .6689 for Zyprexa, respectively.

For Type 2, Table 3.3.3 gives the similar information as Table 3.3.1.

Table 3.3.3. Retrieved Dropout

Type=2	Clozaril (n=61)	TIMETO Mean (sd)	DUREOS Mean (sd)	Zyprexa (n=60)	TIMETO Mean (sd)	DUREOS Mean (sd)
failed	20	176 (175)	145 (153)	27	137 (153)	263 (241)
censored	33	553 (204)	134 (143)	22	561 (225)	135 (171)
completed	8	765 (22)	372 (286)	11	764 (18)	214 (199)

Table 3.3.4 gives the similar information as Table 3.3.2.

Table 3.3.4. DUREOS

Type=2	Clozaril (n=61)	TIMETO Mean (sd)	DUREOS Mean (sd)	Zyprexa (n=60)	TIMETO Mean (sd)	DUREOS Mean (sd)
failed	8	62 (50)	225 (186)	24	115 (128)	288 (244)
censored	53	161 (185)	161 (185)	36	153 (176)	153 (176)
completed	0			0		

Using Table 3.3.4, there are 108 failed, 356 censored, and 26 completed in Clozaril, and 158 failed, 306 censored, and 25 completed in Zyprexa, respectively. The log-rank test gives p-value .0069. The survival probabilities from the KM estimates at 104 weeks are .7276 for Clozaril and .6212 for Zyprexa, respectively.

Based on the above analyses, RDO analysis doesn't show bias favoring Clozaril but one needs to consider the labeling if patients might not be still on the treatment between EOS and RDO.

3.4. Subgroup Analyses

Analyses on gender, age, region, and diagnosis group will be performed.

3.4.1. Gender

For Type 1, there are 301 males (66 failed, 217 censored, and 18 completed) in Clozaril, and 301 males (92 failed, 187 censored, and 22 completed) in Zyprexa, respectively. The log-rank test gives p-value .0576. The survival probabilities from the KM estimates at 104 weeks are .7434 for Clozaril and .6521 for Zyprexa, respectively.

For Type 1, there are 189 females (36 failed, 135 censored, and 18 completed) in Clozaril, and 189 females (49 failed, 122 censored, and 18 completed) in Zyprexa, respectively. The log-rank test gives p-value .1736. The survival probabilities from the KM estimates at 104 weeks are .7857 for Clozaril and .7176 for Zyprexa, respectively.

For Type 2, there are 301 males (78 failed, 207 censored, and 16 completed) in Clozaril, and 301 males (105 failed, 175 censored, and 21 completed) in Zyprexa, respectively. The log-rank test gives p-value .0742. The survival probabilities from the KM estimates at 104 weeks are .6951 for Clozaril and .604 for Zyprexa, respectively.

For Type 2, there are 189 females (42 failed, 129 censored, and 18 completed) in Clozaril, and 189 females (56 failed, 118 censored, and 15 completed) in Zyprexa, respectively. The log-rank test gives p-value .1903. The survival probabilities from the KM estimates at 104 weeks are .75 for Clozaril and .6775 for Zyprexa, respectively.

3.4.2. Age

For Type 1 in age group 18-32, there are 168 (35 failed, 120 censored, and 13 completed) in Clozaril, and 178 (54 failed, 109 censored, and 15 completed) in Zyprexa, respectively. The log-rank test gives p-value .111. The survival probabilities from the KM estimates at 104 weeks are .7589 for Clozaril and .6633 for Zyprexa, respectively.

For Type 1 in age group 33-44, there are 216 (46 failed, 153 censored, and 17 completed) in Clozaril, and 204 (59 failed, 128 censored, and 17 completed) in Zyprexa, respectively. The log-rank test gives p-value .2061. The survival probabilities from the KM estimates at 104 weeks are .7493 for Clozaril and .6804 for Zyprexa, respectively.

For Type 1 in age group ≥ 45 , there are 106 (21 failed, 79 censored, and 6 completed) in Clozaril, and 108 (28 failed, 72 censored, and 8 completed) in Zyprexa, respectively. The log-rank test gives p-value .2391. The survival probabilities from the KM estimates at 104 weeks are .7826 for Clozaril and .6975 for Zyprexa, respectively.

For Type 2 in age group 18-32, there are 168 (40 failed, 116 censored, and 12 completed) in Clozaril, and 178 (63 failed, 102 censored, and 13 completed) in Zyprexa, respectively. The log-rank test gives p-value .0631. The survival probabilities from the KM estimates at 104 weeks are .7218 for Clozaril and .6066 for Zyprexa, respectively.

For Type 2 in age group 33-44, there are 216 (54 failed, 145 censored, and 17 completed) in Clozaril, and 204 (66 failed, 122 censored, and 16 completed) in Zyprexa, respectively. The log-rank test gives p-value .3096. The survival probabilities from the KM estimates at 104 weeks are .706 for Clozaril and .6419 for Zyprexa, respectively.

For Type 2 in age group ≥ 45 , there are 106 (26 failed, 75 censored, and 5 completed) in Clozaril, and 108 (32 failed, 69 censored, and 7 completed) in Zyprexa, respectively. The log-rank test gives p-value .369. The survival probabilities from the KM estimates at 104 weeks are .7303 for Clozaril and .6588 for Zyprexa, respectively.

3.4.3. Region

Six pooled regions will be discussed: U.S. & Canada, U.K. & South Africa, France & Italy, Hungary, Croatia & Czech Republic, and Argentina & Chile. Except France & Italy, the survival probabilities from the KM estimates at 104 weeks for both Type 1 and Type 2 for Clozaril are greater than that for Zyprexa.

For Type 1 in U.S. & Canada, there are 208 (62 failed, 135 censored, and 11 completed) in Clozaril, and 207 (83 failed, 111 censored, and 13 completed) in Zyprexa, respectively. The log-rank test gives p-value .1367. The survival probabilities from the KM estimates at 104 weeks are .6273 for

Clozaril and .5247 for Zyprexa, respectively.

For Type 1 in U.K. & South Africa, there are 74 (8 failed, 60 censored, and 6 completed) in Clozaril, and 74 (15 failed, 55 censored, and 4 completed) in Zyprexa, respectively. The log-rank test gives p-value .0892. The survival probabilities from the KM estimates at 104 weeks are .8782 for Clozaril and .7681 for Zyprexa, respectively.

For Type 1 in France & Italy, there are 63 (16 failed, 43 censored, and 4 completed) in Clozaril, and 62 (13 failed, 39 censored, and 10 completed) in Zyprexa, respectively. The log-rank test gives p-value .2867. The survival probabilities from the KM estimates at 104 weeks are .7081 for Clozaril and .7783 for Zyprexa, respectively.

For Type 1 in Hungary, there are 49 (4 failed, 42 censored, and 3 completed) in Clozaril, and 50 (9 failed, 39 censored, and 2 completed) in Zyprexa, respectively. The log-rank test gives p-value .1897. The survival probabilities from the KM estimates at 104 weeks are .9112 for Clozaril and .8064 for Zyprexa, respectively.

For Type 1 in Croatia & Czech Republic, there are 48 (10 failed, 33 censored, and 5 completed) in Clozaril, and 52 (12 failed, 34 censored, and 6 completed) in Zyprexa, respectively. The log-rank test gives p-value .8545. The survival probabilities from the KM estimates at 104 weeks are .77 for Clozaril and .7587 for Zyprexa, respectively.

For Type 1 in Argentina & Chile, there are 48 (2 failed, 39 censored, and 7 completed) in Clozaril, and 45 (9 failed, 31 censored, and 5 completed) in Zyprexa, respectively. The log-rank test gives p-value .0192. The survival probabilities from the KM estimates at 104 weeks are .9583 for Clozaril and .7875 for Zyprexa, respectively.

For Type 2 in U.S. & Canada, there are 208 (66 failed, 131 censored, and 11 completed) in Clozaril, and 207 (91 failed, 104 censored, and 12 completed) in Zyprexa, respectively. The log-rank test gives p-value .0799. The survival probabilities from the KM estimates at 104 weeks are .6029 for Clozaril and .4788 for Zyprexa, respectively.

For Type 2 in U.K. & South Africa, there are 74 (12 failed, 56 censored, and 6 completed) in Clozaril, and 74 (18 failed, 52 censored, and 4 completed) in Zyprexa, respectively. The log-rank test gives p-value .1815. The survival probabilities from the KM estimates at 104 weeks are .8221 for Clozaril and .7272 for Zyprexa, respectively.

For Type 2 in France & Italy, there are 63 (19 failed, 40 censored, and 4 completed) in Clozaril, and 62 (19 failed, 35 censored, and 8 completed) in Zyprexa, respectively. The log-rank test gives p-value .5742. The survival probabilities from the KM estimates at 104 weeks are .6504 for Clozaril and .6779 for Zyprexa, respectively.

For Type 2 in Hungary, there are 49 (7 failed, 40 censored, and 2 completed) in Clozaril, and 50 (10 failed, 38 censored, and 2 completed) in Zyprexa, respectively. The log-rank test gives p-value .5689. The survival probabilities from the KM estimates at 104 weeks are .8422 for Clozaril and .7851 for Zyprexa, respectively.

For Type 2 in Croatia & Czech Republic, there are 48 (11 failed, 33 censored, and 4 completed) in Clozaril, and 52 (13 failed, 34 censored, and 5 completed) in Zyprexa, respectively. The log-rank test gives p-value .9035. The survival probabilities from the KM estimates at 104 weeks are .7472 for Clozaril and .7381 for Zyprexa, respectively.

For Type 2 in Argentina & Chile, there are 48 (5 failed, 36 censored, and 7 completed) in Clozaril, and 45 (10 failed, 30 censored, and 5 completed) in Zyprexa, respectively. The log-rank test gives p-value .1199. The survival probabilities from the KM estimates at 104 weeks are .8923 for Clozaril and .7629 for Zyprexa, respectively.

3.4.4. Schizophrenia and Schizoaffective Group

For Type 1 in Schizophrenia group, there are 300 (51 failed, 227 censored, and 22 completed) in Clozaril, and 309 (82 failed, 204 censored, and 23 completed) in Zyprexa, respectively. The log-rank test gives p-value .0153. The survival probabilities from the KM estimates at 104 weeks are .7888 for Clozaril and .6847 for Zyprexa, respectively.

For Type 1 in Schizoaffective group, there are 190 (51 failed, 125 censored, and 14 completed) in Clozaril, and 181 (59 failed, 105 censored, and 17 completed) in Zyprexa, respectively. The log-rank test gives p-value .4121. The survival probabilities from the KM estimates at 104 weeks are .69 for Clozaril and .6282 for Zyprexa, respectively.

For Type 2 in Schizophrenia group, there are 300 (67 failed, 212 censored, and 21 completed) in Clozaril, and 309 (98 failed, 192 censored, and 19 completed) in Zyprexa, respectively. The log-rank test gives p-value .0365. The survival probabilities from the KM estimates at 104 weeks are .7411 for Clozaril and .6487 for Zyprexa, respectively.

For Type 2 in Schizoaffective group, there are 190 (53 failed, 124 censored, and 13 completed) in Clozaril, and 181 (63 failed, 101 censored, and 17 completed) in Zyprexa, respectively. The log-rank test gives p-value .3311. The survival probabilities from the KM estimates at 104 weeks are .6787 for Clozaril and .6053 for Zyprexa, respectively.

4. Conclusion

The key issue is whether there is any bias caused by open-label as the Principal Investigator (PI) is unblinded. P-values on CGI-SS for the PI are always smaller than that for the BP, .0946 for the PI and .2803 for the BP in 7-point CGI-SS, and .2767 for the PI and .8708 for the BP in 5-point CGI-

SS. The difference between p-values indicates that the PI might have bias favoring Clozaril.

The Type 1 event is determined by the SMB but it is the PI who determines who should be referred to the SMB. There are 122 patients in Clozaril and 157 patients in Zyprexa, respectively, who were referred to the SMB. The difference between the number of referred patients is 35. However, there are 102 patients in Clozaril and 141 patients in Zyprexa, respectively, who were judged by the SMB as the Type 1 event. The difference of the number of Type 1 event is 39. It is also noted that the correlation between the number of referred and the number of event is high. Based on the above, it is seen that the less referred will have the less Type 1 event, or the number of event is determined by the number of referred. Consequently this affects the number of censored patients.

In this study, for Type 1 event, there are 102 failed, 352 censored, and 36 completed in Clozaril, and 141 failed, 309 censored, and 40 completed in Zyprexa, respectively. The significance of analysis based on Type 1 event is due to 39, the difference of the number of event. The significance of analysis based on Type 2 event is mainly due to the contribution of Type 1 event because analysis on CGI-SS is not statistically nominally significant.

The conclusion is that although the primary analysis is statistically nominally significant with p-value .0309, one should interpret the result with caution due to the issues discussed in this review.

**This is a representation of an electronic record that was signed electronically and
this page is the manifestation of the electronic signature.**

/s/

Kun He
8/9/02 12:11:18 PM
BIOMETRICS

Kun Jin
8/9/02 12:17:59 PM
BIOMETRICS

George Chi
8/21/02 11:58:29 AM
BIOMETRICS