

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

22-128

STATISTICAL REVIEW(S)

Statistical Review and Evaluation

NDA#:22128 Sn #: 130 Supp Doc: Date Received: 10/19/2008

Drug/Biologics Name: Selzentry® (Maraviroc, MVC) 150 or 300 mg tablets

Sponsor: Pfizer Global Research & Development

Indication: Treatment of CCR5-tropic HIV-1 infection

Purpose/Background: Per Pfizer requests, this reviewer provided a document to revise the statistical review document sent to DFS dated 6/20/2007 for the priority review of Maraviroc.

Statistical Reviewer: Susan Zhou, Ph.D. Medical Reviewer: Scott Proestel, M.D.

Review Completion Date: October 26, 2008

Statistical Key Words: Clinical Studies, NDA review, Type I error, erratum

1. Introduction

This document includes corrections to statistical review document sent to DFS dated 6/20/2007 for the priority review of Maraviroc. DAVP review team commented that none of these had any significant impact on the conclusions of the primary stats review or approval of Maraviroc.

2. Erratum to NDA 22-128 Statistical Review by Susan Zhou, Ph.D.

Page 12 1st paragraph, "Subjects who did not fail continued treatment with maraviroc 300 mg QD but in an unblinded manner."

Replace with

'Subjects who did not fail were given the opportunity to opt for open label maraviroc 300 mg BID therapy...'

Page 15 - last paragraph, Page 55 - 2nd paragraph and Page 70 - last paragraph, "Hodges-Lehmen"

Replace with

'Hodges-Lehman'.

Page 15 -3rd paragraph, "As a result, reporting 97.5% CI of the treatment difference in primary efficacy endpoint remained unchanged in the maraviroc label."

Replace with

'As a result, reporting 97.5% CI of the treatment difference in primary efficacy endpoint remained unchanged in the study report but removed from maraviroc label.'

Page 17 Section 1.3.7 "We experienced difficulties in requesting the submissions of SAS programs on time"

Replace with

‘See history of data and SAS program submissions (Appendix 1).’

Page 18 Under #4 “datasets submitted during March 23-April 4, 2007”

Replace with

‘datasets received during March 23-March 26, 2007’.

Page 22 - last paragraph, “Swaziland”

Replace with ‘Switzerland’.

Statistical Review Page 37 - 1st bullet, “...the subjects in the maraviroc 150 BID regimen had slightly better results than the maraviroc 300 mg QD regimen ($p>0.20$)”

Replace with

‘...the subjects in the maraviroc 150 BID regimen had slightly better results than the maraviroc 150 mg QD regimen ($p>0.20$)’.

Page 48 - #3, “This feather is different....”

Replace with

‘This feature is different....’.

Page 69 – last sentence, “The sponsor’s presentations in the Advisoryand maraviroc labeling, the sponsor kept the 97.5% CI...”

Replace with

‘The sponsor’s presentations in the Advisoryand initially proposed maraviroc labeling, the sponsor kept the 97.5% CI... However, it was agreed with the review team to remove all confidence intervals from the label’.

Page 70 2nd paragraph, “As a result, reporting 97.25% CI of the treatment difference in primary efficacy endpoint remained unchanged in the maraviroc label.”

Replace with

‘As a result, confidence intervals of the treatment difference in primary efficacy endpoint were all removed in the maraviroc label.’

Page 70 - last paragraph, “sceening”

Replace with ‘screening’.

Page 70 - 3rd paragraph, “In the future, the adjustment of multiplicity in a phase 2b/3 design should be discussed in the protocol and the SAP reviews to avoid the potential confusions.”

Replace with

‘Although the multiplicity adjustment for two doses was fully addressed in the protocol and the SAP, the adjustment for the two studies combined was not discussed.’

Page 71 Section 5.1.7 “The sponsor failed to submit SAS programs on time....:”

Replaced with

‘We requested the sponsor submit all datasets and SAS programs as early as the pre-NDA meeting, the whole process lasted more than four months. Please see history of data and SAS program submissions (Appendix 1).’

Regarding sponsor Comment #12

Page 82 of electronic file, p.2 of Stats Evaluation, “You failed to submit *.csv files ...”

Replace with

‘SAS input data files including two *.csv files to debug the SAS program for creation anlpop.xpt were not submitted. However, in the E-mail submissions 02 February 2007, the sponsor explained the reasons. The review team accepted the explanations and agreed to use the sponsor’s indicators for defining study populations in statistical evaluation of Maraviroc efficacy.’

3. Appendix 1. DAVP’s Requests of Data and SAS programs and Sponsor’s Responses*

#	Date of Request	Communication	Requests
A	28 November 2006	Pre-NDA Meeting	SAS xpt file and SAS programs to EDR
B	1 February 2007	Facsimile	1. Reiterating the requests in A; 2. Additional requests <ul style="list-style-type: none">• SAS programs for anlpop.xpt;• SAS programs for vir27.xpt, vir28.xpt and vir29.xpt.
C	7 February	Statistical Review	Requesting

	2007		<ul style="list-style-type: none"> • Two input files 'pidlist.csv' and 'pv.csv' for anlpop.xpt; • Revised SAS programs to create anlpop.xpt; • SAS programs for hiv.xpt and vir*.xpt; • Updated *.xpt for all studies; • SAS programs for A4001026.
D	26 February 2007	Internal E-mails	<p>Waiting Dr. Proestel's telecom with the sponsor regarding</p> <ul style="list-style-type: none"> • Two input files 'pidlist.csv' and 'pv.csv' for anlpop.xpt; • TLOVR and SAS programs
E	5 March 2007	Facsimile	<p>Reiterating revised SAS programs to create anlpop.xpt and SAS programs for hiv.xpt; Requesting</p> <ul style="list-style-type: none"> • SAS program to obtain covar.xpt; • TLOVR datasets and SAS programs; • Death information in the ITT population; • Explanation of one individual in the FAS population who had a screening HIV-1 viral load 2600 copies/mL.
F	20 March 2007	E-mail/EDR	9 SAS log files sent by E-mail and 11 SAS log files sent to EDR.
G	21 March 2007	E-mail	TLOVR
H	29 March 2007	Facsimile	<p>Requesting</p> <ul style="list-style-type: none"> • Complete datasets for HIV-1 viral load and CD4+ cell count including the follow-up data; • Conducting sensitivity analyses on change from baseline to Week 24 in HIV-1 viral load (primary efficacy endpoint) and CD4 (secondary efficacy endpoint) for A4001027, A4001028 respectively and those using combined data; • SAS programs for the above analyses.
I	1 st Week, April 2007	Telecom and E-mail*	Problems identified in the submitted SAS program per request in H.

*. Some of the E-mail information in Microsoft Outlook could not be recovered after consulting with the CDER Helpdesk.

Susan Zhou, Ph. D.
Mathematical Statistician
Date: November 24, 2008

Concur: Guoxing Soon, Ph.D., Statistics Team Leader

cc:

HFD-530/ Kenny Shade (Project Manager)
HFD-530/ Scott Proestel, M.D. (Medical Reviewer)
HFD-530/ Linda Lewis, M.D. (Medical Team Leader)
HFD-530/ Debbie Birnkrant, M.D. (Medical Director)
HFD-530/ Jeffrey Murray, M.D. (Medical Deputy Director)
HFD-725/ Susan Zhou, Ph.D. (Statistical Reviewer)
HFD-725/ Guoxing Soon, Ph.D. (Statistics Team Leader)
HFD-725/ Mohhamad Huque, Ph.D. (DB4, Director)
HFD-725/ Daphne Lin, Ph.D., (DB4, Deputy Director)
HFD-725/ Lillian Patricia, MS, MBA (OB)

Appears This Way
On Original

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

/s/

Susan Zhou
11/24/2008 03:59:56 PM
BIOMETRICS

Greg Soon
12/3/2008 05:43:45 PM
BIOMETRICS



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number: NDA 022,128/N000
Drug Name: Selzentry[®] (Maraviroc, MVC) 150 or 300 mg tablets
Indication(s): CCR5-tropic HIV-1 infection
Applicant: Pfizer Global Research & Development
Date(s): December 20, 2006
Review Priority: Priority
Biometrics Division: Division IV
Statistical Reviewer: Susan Zhou, Ph.D. (HFD-725)
Concurring Reviewers: Greg Soon, Ph.D. (HFD-725)
Medical Division: DAVP
Clinical Reviewer: Scott Proestel, MD (HFD-530)
Clinical Team Leader: Katherine Laessig, MD (HFD-530)
Project Manager: Kenny Shade, JD, BSN (HFD-530)

Keywords: Phase 2b/3 trials, Week 24 efficacy data, Discontinuation, Sensitivity analysis, Multiplicity, Competing risk model

Table of Contents

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES	1
FOOD AND DRUG ADMINISTRATION	1
STATISTICAL REVIEW AND EVALUATION	1
LIST OF TABLES.....	4
LIST OF FIGURES.....	5
1. EXECUTIVE SUMMARY	6
1.1 CONCLUSIONS AND RECOMMENDATIONS	6
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	10
1.3 STATISTICAL ISSUES AND FINDINGS	11
1.3.1 Significant Heterogeneity in Discontinuation Status	11
1.3.2 Sensitivity Analyses of the Primary Efficacy Endpoint	12
1.3.3 Baseline HIV-1 Viral Load: Average Value or Day 1 Value.....	14
1.3.4 Multiplicity.....	15
1.3.5 Different Matrices in Measuring Efficacy	15
1.3.6 Time to Discontinuation due to Competing Risks	16
1.3.7 DAVP's Suggestions in Requesting SAS Programs	16
2. INTRODUCTION	17
2.1 OVERVIEW.....	17
2.2 DATA SOURCES	17
3. STATISTICAL EVALUATION	19
3.1 EVALUATION OF EFFICACY	19
3.1.1 Study Design	19
3.1.2 Analysis Datasets and Definitions	20
3.1.3 Primary Efficacy Endpoint, Hypotheses and Decision Rules	20
3.1.4 Secondary Efficacy Endpoints	21
3.1.5 Results.....	22
3.1.5.1 PATIENT DISPOSITION, DEMOGRAPHIC AND BASELINE CHARACTERISTICS.....	22
3.1.5.2 ESTIMATED TREATMENT DIFFERENCES IN PRIMARY EFFICACY ENDPOINT USING DIFFERENT BASELINE VL	28
3.1.5.3 SENSITIVITY ANALYSES OF PRIMARY EFFICACY ENDPOINT	29
3.1.5.4 ALTERNATIVE ANALYSIS OF SELECTED SECONDARY EFFICACY ENDPOINT.....	37
TIME TO DISCONTINUATION AT WEEK 24	37
TIME-AVERAGE DIFFERENCE (TAD) IN HIV-1 VL FROM BASELINE TO WEEK 24	44
CHANGES FROM BASELINE TO WEEK 24 IN CD4+ CELL COUNT.....	47
6/11/2007 Last revised	2

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

3.2	EVALUATION OF SAFETY	48
4.	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	53
4.1	GENDER, RACE AND AGE	53
4.1.1	<i>Applicant's Subgroup Analyses</i>	53
4.1.2	<i>Reviewer's Analyses</i>	54
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS	59
4.3	MULTIVARIATE ANALYSES TO EVALUATE AGE AND RACE INTERACTIONS WITH TREATMENT EFFECTS.....	60
5.	CONCLUSIONS AND RECOMMENDATIONS.....	66
5.1	STATISTICAL ISSUES AND COLLECTIVE EVIDENCE	66
5.1.1	<i>Significant Heterogeneity in Discontinuation Status</i>	66
5.1.2	<i>Sensitivity Analyses of the Primary Efficacy Endpoint</i>	66
5.1.3	<i>Baseline HIV-1 Viral Load: Average Value or Day 1 Value</i>	68
5.1.4	<i>Multiplicity</i>	69
5.1.5	<i>Different Matrices in Measuring Efficacy</i>	70
5.1.6	<i>Time to Discontinuation Due to Competing Risks</i>	70
5.1.7	<i>DAVP's Suggestions in Requesting SAS Programs</i>	71
5.2	CONCLUSIONS AND RECOMMENDATIONS	72
6.	APPENDICES.....	76
6.1	REFERENCES.....	76
6.2	TWO ANCOVA MODELS	77
	SIGNATURES/DISTRIBUTION LIST (OPTIONAL).....	79

LIST OF TABLES

	Page
Table 1. Studies A4001027 and A4001028: Demographics and Baseline Characteristics*	24
Table 2. Studies A4001027 and A4001028: Patient Disposition*	25
Table 3. Change from Baseline to Week 24 in Viral Load by Study	33
Table 4. Studies A4001027 and A4001028: Change from Baseline to Week 24 in VL	33
Table 5. Sensitivity Analysis: Mean Change from Baseline to Week 24 in Viral Load	35
Table 6. Sensitivity Analysis: Median Change from Baseline to Week 24 in Viral Load	36
Table 7. Time-Average Difference (TAD) from Baseline to Week 24 in Viral Load.....	46
Table 8. Mean Change from Baseline to Week 24 in CD4+ Cell Count (cells/ μ L)	49
Table 9. Median Change from Baseline to Week 24 in CD4+ Cell Count (cells/ μ L) ¹	50
Table 10. Race, Gender, Age and Region: Median Change in Viral Load from Baseline to Week 24	57
Table 11. Race, Gender, Age and Region: Mean Change in Viral Load from Baseline to Week 24	58
Table 12. Other Subgroup: Median Change in Viral Load from Baseline to Week 24.....	62
Table 13. Other Subgroup: Mean Change in Viral Load from Baseline to Week 24.....	64

LIST OF FIGURES

	Page
Figure 1. Baseline VL in Studies A4001027 & A4001028: Day 1 VL- Average VL.....	31
Figure 2. Study A4001027: Baseline HIV VL: Average V. Day 1	32
Figure 3. Study A4001028: Baseline HIV VL: Average V. Day 1	32
Figure 4. Histograms - Change from Baseline to Week 24 in VL.....	34
Figure 5. Study A4001027: Time to Discontinuation.....	39
Figure 6. Study A4001028: Time to Discontinuation.....	39
Figure 7. Study A4001027: Time to Discontinuation (Enfuvirtide Use in ART)	40
Figure 8. Study A4001028: Time to Discontinuation (Enfuvirtide Use in ART)	40
Figure 9. Study A4001027: Time to Discontinuation (Non-Enfuvirtide Use in ART)	41
Figure 10. Study A4001028: Time to Discontinuation (Non-Enfuvirtide Use in ART)	41
Figure 11. Study A4001027: Time to Discontinuation (VL<100,000 copies/mL)	42
Figure 12. Study A4001028: Time to Discontinuation (VL<100,000 copies/mL)	42
Figure 13. Study A4001027: Time to Discontinuation (VL≥100,000 copies/mL).....	43
Figure 14. Study A4001028: Time to Discontinuation (VL≥100,000 copies/mL).....	43
Figure 15. Study A4001027: Median CD4+ cells/μL.....	51
Figure 16. Study A4001028: Median CD4+ cells/μL.....	51
Figure 17. Study A4001027: Baseline CD4+ cells/μL (Average v. Day 1).....	52
Figure 18. Study A4001028: Baseline CD4+ cells/μL (Average v. Day 1).....	52
Figure 19. Histograms: Change from Baseline to Week 24 in VL among Black Subjects	56

Subgroup analyses of the primary efficacy endpoint on selected baseline characteristics were conducted. This included the number of overall susceptibility score (OSS: 0-2, ≥ 3), the number of overall phenotypic score (PSS: 0-2, ≥ 3), the number of genotypic susceptibility score (GSS: 0-2, ≥ 3), previous enfuvirtide use in the ART (user and non-user), subjects' screening VL (< or $\geq 100,000$ copies/mL) and baseline CD4+ cell count (<100, 100-200, >200 cells/mm³).

- The criteria for grouping the screening OSS, PSS and GSS were based on homogeneity of the outcome within a subgroup and sample sizes after grouping. For example, grouping screening OSS into 0-2 and ≥ 3 were based on initial subgroup analyses of OSS=0,1,2 and ≥ 3 on primary efficacy endpoint. The results showed that the two subgroups of OSS 0-2 and ≥ 3 provided two levels of virologic responses without significant differences within a subgroup after grouping.

Subgroups of baseline OSS, GSS and PSS showed significant treatment differences (MVC versus placebo) regarding the primary efficacy endpoint, while subgroups of previous enfuvirtide use in the ART, subjects' screening VL and baseline CD4+ cell count, respectively, did not appear to be any clinically important treatment differences. Subjects with baseline OSS ≥ 3 had significantly increased mean reductions of Week 24 VL from baseline, 2.3 log₁₀ copies/mL in MVC+OBT groups and 2.0 log₁₀ copies/mL in placebo+OBT group. This resulted almost no meaningful treatment benefit (-0.3 log₁₀ copies/mL) maraviroc over placebo. Conversely, subjects receiving maraviroc and with baseline OSS 0-2 had at least -1.1 log₁₀ copies/mL treatment benefit in mean change from baseline in viral load at Week 24. The results of subgroup analyses based on screening PSS (0-2, ≥ 3) and GSS (0-2, ≥ 3) appeared to be similar to those of the OSS subgroup analyses.

Subgroup analyses of the primary efficacy endpoint on selected demographic characteristics were conducted, including race (white versus black), age (≤ 42 , 43-48, ≥ 49), gender and region (USA versus non-USA). These three age subgroups (≤ 42 , 43-48, ≥ 49) were obtained using the 33% (age=42) and 67% (age=49) percentiles as cut points for the baseline age category, so that the sample sizes among age strata were similar.

There did not appear to be any clinically important gender or region differences in the mean change from baseline to Week 24 in VL.

As to race (black and white), the following results were observed.

- Using an ANCOVA model with treatment groups, baseline CD4+ cell count in square root transformation, baseline VL level, previous enfuvirtide use in ART, baseline OSS (0-2, ≥ 3), black or white and the interactions to treatment as explanatory variables (See Section 4.2), race (black v. white) was significantly associated with mean reduction from baseline to Week 24 in VL among subjects receiving MVC BID

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

($p < 0.01$), MVC QD ($p = 0.06$), but not in placebo ($p = 0.43$). The race by (MVC BID-placebo) treatment interaction was estimated as -0.80 ($se = 0.24$) \log_{10} copies/mL, resulting $p = 0.018$ by the Wald-t test.

- The ANCOVA model also indicated that black subjects receiving maraviroc had only $0.3 \sim 0.4 \log_{10}$ copies/mL more reductions in Week 24 VL compared to placebo. However, white subjects in the MVC regimens showed at least $-0.9 \log_{10}$ copies/mL treatment benefit over placebo.
- Subgroup analyses for white and black respectively showed similar results. However, the analyses were based on univariate analyses where other baseline factors such as baseline CD4+, screening VL, previous use of enfuvirtide, and baseline OSS were not controlled. Please note that the sample size ($n = 26$) in the black placebo subgroup is small. The small number will produce relatively coarse estimates of the efficacy. Hence the observed treatment differences between white and black subjects remained unclear.

As to age effects, the following results were observed.

- The ANCOVA model with treatment groups, baseline CD4+ cell count in square root transformation, baseline VL level, enfuvirtide use in ART, baseline OSS, age and age by treatment interactions as explanatory variables (Section 4.3. ANCOVA Model 2) indicated that overall, age was significantly associated with mean reduction in Week 24 VL from baseline among subjects receiving MVC BID ($p = 0.027$). The older the age, the more reduction in Week 24 VL. The estimated age by (MVC BID – placebo) interaction was -0.020 ($se = 0.014$) \log_{10} copies/mL, resulting $p = 0.16$, by the Wald-t test. Thus, controlling for the baseline characteristics in the model, age by MVC QD treatment effect versus placebo was no longer significant at $\alpha = 0.05$ level.
- Subgroup analyses using three age subgroups (≤ 42 , $43-48$, ≥ 49) showed those receiving MVC BID with age 49 or older had an extra mean reduction ($0.5 \log_{10}$ copies/mL) in Week 24 VL than those with age 42 or younger, compared with placebo subjects ($p = 0.08$). No such findings were observed among subjects receiving MVC QD or placebo groups.

This reviewed conducted alternative analyses on selected key secondary efficacy endpoints including as time-average difference (TAD) in VL from baseline to Week 24 and change from baseline to Week 24 in CD4+ cell count using data pooling the two studies A4001027 and A4001028. In addition, time to discontinuation using the Kaplan-Meier approach was conducted respectively for A4001027 and A4001028 in the early phase of review.

- The sensitivity analyses on TAD in VL from baseline to Week 24 showed robustness

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

in estimated mean treatment differences, regardless of different baseline VL (Day 1 or average VL) used in calculation of TAD, different cut points (Day 154, Day 168, Day 196) used to define the discontinuation and imputation. The mean TAD were 0.76~0.77 and 0.80~0.84 \log_{10} copies/mL, respectively in subjects receiving MVC QD and MVC BID, compared with placebo. All the 99.9% CIs of the mean treatment differences in VL exclude zeros, indicating the supportive evidence of the superiority of MVC in VL reductions, compared to Placebo. MVC BID regimen appears to have slightly better results ($<0.07 \log_{10}$ copies/mL) than the MVC QD regimen in VL reductions.

The results of the alternative analyses of treatment difference in CD4+ cell count increase from baseline to Week 24 are as follows.

- When Day 1 CD4+ cell count was used as baseline, the two analyses using last observation carry forward (LOCF) with or without imputing zero to missing baseline (or on study) showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 57~58, and 50~52 cells/ μ L, respectively for MVC QD and MVC BID groups, compare with placebo. The estimated mean treatment differences results using average CD4+ prior to treatment with study drugs as baseline were 2~6 cells/ μ L lower. All the 99.9% CIs excluded zero, supporting the superiority of maraviroc, compare with placebo in CD4+ cell count increase.
- The Week 24 completers (n=701) were defined as subjects who had on study CD4+ at Week 24 time window. Analyses of the Week 24 completers using Day 1 CD4 as baseline showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 38 and 25 cells/ μ L, respectively for MVC QD and MVC BID, compare with placebo. The results using average CD4+ prior to treatment with study drugs as baseline were 5~8 cells/ μ L lower. All the 99.9% CIs included zero, indicating a reduced significant level.
- Different from change from baseline to Week 24 in VL, 1) It appeared that subjects receiving MVC QD had slightly more increase (<13 cells/ μ L) in Week 24 CD4+ cell count than those who received MVC BID, compare with placebo; and 2) the treatment difference (MVC versus placebo) using median (Hodges-Lehman) may or may not be smaller than the mean for CD4+ cell count.

Using the Kaplan-Meier method (K-M), time to discontinuation was significantly longer in the maraviroc QD and BID regimens compared with the placebo regimen, $p<0.0001$ by the log-rank test for A4001027 and A4001028, respectively. Subgroup analyses suggest that time to discontinuation may be associated with the previous use of enfuvirtide in ART but not associated with screening HIV-1 VL level at a significant level of $p=0.20$. Different temporal patterns in different studies were observed. In A4001027 among those who used

enfuvirtide in ART, the subjects receiving MVC BID were doing somewhat better than those receiving MVC QD. Conversely, in A4001028 among the non-enfuvirtide users in ART, the subjects in the MVC QD group were doing somewhat better than the MVC BID group. However, these qualitative interactions were based on the univariate analyses (K-M) and at the significance level of 0.2.

1.2 Brief Overview of Clinical Studies

Under NDA 21,228/N000, the sponsor submit interim analyses at Week 24 of the efficacy and safety data in two independent Phase 3, registrational Studies A4001027 and A1004028 for accelerated approval.

The A1004027 and A1004028 are two identically designed ongoing, multicenter, randomized, double-blind, placebo-controlled trials of maraviroc, in combination with optimized background therapy (OBT) versus optimized background therapy alone for the treatment of antiretroviral (ART) experienced HIV-1 CCR5-tropic HIV-1 infected adults. Study A4001027 was conducted in America and Canada. Study 4001028 was conducted in America, Australia, and Europe. 1049 patients were assigned and treated to maraviroc 300 mg QD, maraviroc 300 mg BID and placebo treatment groups when given in combination with OBT with a ratio of 2:2:1. The primary efficacy endpoint is change in HIV-1 RNA at Week 48 with an interim analysis at Week 24. The duration of treatment is 48 weeks.

Two additional Phase 2b or 2b/3 studies to assess maraviroc were conducted.

A Phase 2b Study A4001029 was designed as a safety study to assess maraviroc 300 mg QD and BID dose equivalents in antiretroviral-experienced patients infected with dual/mixed-tropic, CXCR4-tropic or non-phenotypable (non-CCR5 tropic) HIV-1. The A4001029 was conducted in North America, Europe and Australia. 190 patients were assigned to maraviroc 300 mg QD, maraviroc 300 mg BID and placebo treatment groups when given in combination with OBT with a ratio of 1:1:1. The primary efficacy endpoint was change in HIV-1 RNA at Week 24. The duration of treatment was 48 weeks.

A Phase 2b/3 Study A4001026 in treatment-naïve patients infected with CCR5 tropic HIV-1 was conducted in North America, Europe, Australia, Latin America and South Africa. 205 patients were assigned to maraviroc 300 mg QD, maraviroc 300 mg BID and efavirenz 600 mg QD treatment groups when given in combination with Zidovudine/Lamivudine with a ratio of 1:1:1. The primary efficacy endpoint was percentage of subjects with HIV-1 RNA < 400 /50 copies/mL through Weeks 48/96. The duration of treatment was 96 weeks. The DSMB reviewed the interim analysis where 205 patients had been treated with blinded therapy for 16 weeks, the Phase 2b run-in part of the study. The DSMB recommended

discontinuation of the maraviroc 300 mg QD treatment group. This group failed to demonstrate non-inferiority criteria in both the time average difference (TAD) and percentage of subjects with HIV-1 < 400 copies/mL. Subjects who did not fail continued treatment with maraviroc 300 mg QD but in an unblinded manner. Subjects in the other two treatment groups remain double-blinded.

This statistical review focused on the 24 weeks efficacy data on two registrational Studies A4001027 and A1004028.

1.3 Statistical Issues and Findings

During the review, several statistical issues in estimating the treatment effects maraviroc versus placebo group have been addressed.

1.3.1 Significant Heterogeneity in Discontinuation Status

In clinical trials, it is common to observe the heterogeneities in proportions of subjects discontinuing from the study, censorship violating the non-informative assumptions, and in different durations of treatment, between treatment groups. It is known that significant distributional differences in discontinuation between treatment groups may be associated with biased estimated treatment effects both in efficacy and safety.

In these two studies, discontinuation rates were much higher in the placebo treatment groups than in the MVC treatment groups. As of the date of the Week 24 cut-off, for data combining Studies A4001027 and A4001028, 133 (63.6%) of the subjects in the placebo group discontinued from study. This was a significantly higher proportion than the MVC QD 143 (34.5%) and MVC BID 138 (32.4%) groups. The main reason was due to lack of clinical response or efficacy, found in 106 (50.7%) of the subjects in the placebo group - significantly higher than MVC QD 81 (19.6%) and MVC BID 91 (21.4%) groups. This is based on sample sizes 209, 414 and 426, respectively, for the placebo, MVC QD and MVC BID groups.

Hence, time to discontinuation was much shorter in the placebo treatment group compared to the two MVC groups. Figures 5 and 6 show the Kaplan-Meier (K-M) curves of time to discontinuation by A4001027 and A4001028, respectively. It appears that the K-M curves in the MVC groups are significantly separate from that in the placebo group, as early as Week 12.

As a result, there was a great concern that an extreme discontinuation rate imbalance between the MVC and placebo groups may result biased estimation in maraviroc efficacy.

1.3.2 Sensitivity Analyses of the Primary Efficacy Endpoint

To verify whether the estimated efficacy sizes by the sponsor are representative of the true effects of maraviroc, and to examining potential effects of the discontinuation discrepancies between placebo and maraviroc groups, this reviewer conducted three types of sensitivity analyses on primary efficacy endpoint:

- Analysis using all available HIV-1 VL data regardless whether a subject was on study drug or not.
- Analysis on Week 24 completer.
- Imputation of missing by -0.4 to 0.3 log₁₀ copies/mL with an increment of 0.1 log₁₀ copies/mL.

The methodology used in the sensitivity analyses was as follows;

- The treatment difference on the primary endpoint was evaluated using the mean and median treatment difference. The later is known as Hodges-Lehman approach, where the median of all possible pairs ($n_1 \cdot n_2$) in treatment difference should be obtained, n_i ($i=1,2$) is the sample size for the i th group.
- Type I error was adjusted when data were combined for A4001027 and A4001028 ($\alpha=0.001$), and also the two comparisons MVC QD or MVC BID versus Placebo ($\alpha=0.0005$). Hence, 99.95% confidence intervals of the mean and median treatment differences were estimated for the primary efficacy endpoint.
- The mean or median treatment differences were sample summary statistics without adjusting for their randomization strata.
- All the sensitivity analyses used Day 1 HIV-1 VL as baseline. If there are more than one value in the Week 24 time window (from Week 22 to Week 28), the one closest to Day 168 was selected.

The results of these sensitivity analyses are summarized in Tables 5 and 6.

The Sensitivity Analysis 1 used all available HIV-1 VL data regardless whether a subject was on study drug or not. This meant that those VL data after the subjects in the placebo group had switched to MVC+ OBT were included in the analysis as if they had been treated with Placebo+OBT. We obtained an extra mean reduction of 0.5 log₁₀ copies/mL and median reductions of 0.45~0.46 log₁₀ copies/mL in the two MVC regimens compared to the Placebo. The 99.95% CIs were all excluding zeros, indicating the superiority of MVC compared to Placebo. The treatment difference between the two MVC regimens was within 0.01 log₁₀ copies/mL.

The Sensitivity Analysis 2 was conducted among Week 24 completers, defined as the *earliest date of two events* exceeded Day 155 since Day 1: (1) the date of discontinuation from study

and (2) the last date of treatment with study regimen. HIV-1 VL data were extended to one week from the date of discontinuation or the date of stopping treatment with study regimen.

- Subjects in the MVC regimens showed extra mean reductions 0.48 and 0.54 log₁₀ copies/mL in the MVC-QD and MVC-BID regimen, compared to Placebo. The extra median reductions were 0.38 log₁₀ and 0.44 log₁₀ copies/mL respectively in the MVC-QD and MVC-BID regimen, compared to Placebo. However, the superiority of MVC compared to Placebo was confirmed using mean comparisons. Results by the Hodges-Lehman approach support the superiority of MVC-BID (p<0.0005), not MVC-QD (p>0.0005), at the type I error of 0.0005 level.

The third type of sensitivity analysis contained eight imputations (Analysis 3-10) to impute the missing of Week 24 HIV-1 VL values from -0.4 to 0.3 log₁₀ copies/mL with an increment of 0.1 log₁₀ copies/mL.

- As the imputed value increasing from -0.4 to 0.3 log₁₀ copies/mL, the mean treatment difference in change from baseline to Week 24 ranging -0.79 to -0.98 log₁₀ copies/mL for (MVC QD-Placebo), and -0.88 to -1.07 log₁₀ copies/mL for (MVC BID-Placebo).
- The estimated median treatment differences were slightly less than the mean treatment differences. As the imputed value increasing from -0.4 to 0.3 log₁₀ copies/mL, the median treatment difference in change from baseline to Week 24 ranging -0.69 to -0.88 log₁₀ copies/mL for MVC QD-Placebo, -0.80 to -0.99 log₁₀ copies/mL for MVC BID-Placebo.
- More than 50% of the subjects in the placebo regimen had discontinued from study by Week 24. Hence, any imputation approach that involved imputing of a single value to the missing in the placebo regimen should make this imputation value a median of the sample after imputation. For example, after imputing missing with -0.3 log₁₀ copies/mL, the median change from baseline to Week 24 should be -0.3 log₁₀ copies/mL in placebo regimen.

We conclude that the superiority of MVC versus placebo in estimating primary efficacy endpoint is essentially well maintained by these sensitivity analyses,. This was concluded even though some of the sensitivity analyses were rather conservative and were designed *not favoring the maraviroc treatment*.

1.3.3 Baseline HIV-1 Viral Load: Average Value or Day 1 Value

In evaluating the primary efficacy endpoint, the sponsor used average of HIV-1 viral load (Average-VL) at screening, at randomization and Day 1 (Day 1-VL) prior to treatment with study drugs as baseline VL. It was noticed that there was a mean time window of 5.9 weeks (mean=41 days, range -86 to -5 days prior to Day 1) between screening and Day 1. In addition, the study population in the A4001027 and A4001028 had a mean of 14 years of HIV-1 infections, and most of the subjects were on stable ART for at least 4 weeks. There was a concern whether the baseline VL may be influenced by the ART prior to Day 1.

- Figures 2 and 3 show the observed pairs and regression lines Day 1-VL and Average-VL by study. Similar estimated slopes in the regression lines for the two studies were obtained.
- Figure 4 shows the histogram of difference between the Day 1-VL and Average-VL. Overall, the Day 1-VL appears below the Average-VL, $p=0.053$ by the signed rank test. The mean (median) difference (Day 1-VL - Average-VL) was -0.003 (-0.013) with a range of $(-1.418, 0.773)$ \log_{10} copies/mL. Percentages of subjects with such difference < -0.3 , -0.2 , and -0.1 \log_{10} copies/mL were 5.3%, 9.7% and 23.9% respectively.

This reviewer conducted sensitivity analysis in primary efficacy endpoint using the Day 1-VL as baseline. Change from baseline at Week 24 data in VL were fitted to analysis of covariance (ANCOVA) models which include treatment regimens, screening HIV-1 VL strata ($<$ or $\geq 100,000$ copies/mL), and enfuvirtide use in the ART.

Tables 3 and 4 summarize the results using different baseline VL. For data pooling the two studies, the estimated mean reduction from baseline to Week 24 was about $0.14 \log_{10}$ copies/mL more using the Day 1-VL as baseline, compared to Average-VL as baseline. Only slight differences were found among three treatment groups. Hence, the estimated treatment differences MVC-Placebo using different baseline VL were seen as similar. However, the lengths of the estimated 97.5% CI using Day 1-VL were $0.15 \log_{10}$ copies/mL wider than those using the Average-VL as baseline (Table 4).

- Please note that the 99.95% CIs for the two studies combined should be used for multiplicity concern. This reviewer used 97.5% CIs so that the results could be comparable between the two methods.

Using Day 1-VL as baseline should be more appropriate for endpoints such as change from baseline in VL and time average difference (TAD) for this ART experienced study populations. However, the estimated treatment differences regarding change from baseline in VL appear to be robust due to the double-blind and randomized study design.

1.3.4 Multiplicity

Multiplicity is unique problem for trials such as the two maraviroc trials A4001027 and A4001028 with phase 2b/3 design. In the maraviroc label, the sponsor decided to use the twice-daily regimen (MVC BID) for maraviroc as recommended regimen. As a result, data in the MVC QD groups were excluded; only data in the MVC BID and placebo groups were used for summarizing the clinical studies and maraviroc safety and efficacy. All results were corresponding to data pooling the two maraviroc trials A4001027 and A4001028. In the Advisory Committee Meeting (April 20, 2007) and maraviroc labeling, the sponsor kept the 97.5% CI or a type I error of 0.025 unchanged for the estimating of primary efficacy endpoint.

We suggest using 99.95% CI or a type I error of 0.0005 for such entity. It appears to be correct to use 97.5% CI for the treatment differences in primary efficacy endpoint obtained by individual study A4001027 or A1004028 respectively. When pooling the two studies together, it is implied that one study has been conducted, then the type I error should be 0.0005 (two treatment arms versus one placebo arm) to account for multiple comparisons, or 0.001 (one treatment arm versus one placebo arm) with no need to adjust for multiple comparison. One suggestion was to change the confidence interval to p-value.

The sponsor claimed that the change of the protocol-defined level of significance in type I error from 0.025 to 0.0005 would cause confusions and hence not be acceptable for the medical communities. As a result, reporting 97.5% CI of the treatment difference in primary efficacy endpoint remained unchanged in maraviroc label.

In the future, the adjustment of multiplicity in a phase 2b/3 design should be discussed in the protocol and the SAP reviews to avoid the potential confusions.

1.3.5 Different Matrices in Measuring Efficacy

The assessment of treatment effect size maraviroc versus placebo by the mean approach could be greatly influenced by outliers in the data since the distributions of change from baseline to Week 24 in VL were all skewed thus not Normal distributed. The problem should be worse when the sample size is small. Different matrices in measuring treatment difference should be considered for the evaluations. **Here we used Hodges-Lehmen's median approach as an alternative.** Similar findings and conclusions were obtained with a few exceptions although the estimated median treatment differences MVC versus placebo were slightly reduced than those using the mean approach.

1.3.6 Time to Discontinuation due to Competing Risks

By the study design, subjects who had AE, pregnancy, etc, and treatment failure, will be discontinued from study. As seen in the study reports and data, the main reason of the subjects discontinued from studies was related to the lack of efficacy. In the two studies combined, there were approximately 20% of the subjects in the two MVC groups and 50% in the placebo groups discontinued due to virologic failure. Time to discontinuation due to different reasons should be analyzed aptly, especially when substantial number of subjects would discontinued from the study due to a study design. This reviewer suggests using the competing risk models⁴⁻⁸ to conduct analyses for the maraviroc discontinuation data in the presence of multiple reasons such as 1) adverse events; 2) virologic failure by the study design; and 3) other reasons such as patient withdrawn consent, lost to follow-up, etc. The competing risk analyses could investigate the treatment differences for the important competing risk reasons for discontinuation simultaneously.

Statistical methodology should be discussed with the sponsor in protocol review as well as the review of statistical analysis plan (SAP).

1.3.7 DAVP's Suggestions in Requesting SAS Programs

We experienced the difficulties in requesting the submissions of SAS programs on time.

As suggested by the DAVP, the statistical review team should request the SAS programs before a NDA filing meeting. The statistical reviewer should refuse to recommend filing the NDA if the sponsor fails to submit SAS programs along with the SAS *.xpt files.

**Appears This Way
On Original**

2. INTRODUCTION

2.1 Overview

Maraviroc is first in a class of new antiretrovirals that inhibit binding of HIV to the CCR5 receptor, thereby blocking an essential initial step in viral replication. It is intended to use for subjects infected with CCR5-tropic HIV-1

The sponsor, Pfizer Global Research & Development, has been conducting four studies, including two independent Phase 2b/3 Studies A4001027 and A1004028 in antiretroviral treatment (ART) experienced HIV-1 infected adults, a Phase 2b/3 study A4001026 in ART-naïve, CCR5-tropic HIV-1 infected subjects, and a Phase 2b study A4001029 in non-CCR5-tropic HIV-1 infected subjects.

Under NDA022,128/N000, the sponsor is seeking an accelerated approval for maraviroc upon review of this submission, where a “priority review” has been granted. The interim analyses at 24 weeks of the two independent Phase 3, registrational Studies A4001027 and A1004028 evaluating maraviroc 300 mg QD and BID dose equivalents, provide the efficacy and safety data for treatment-experienced, CCR5-tropic HIV-1 infected subjects.

This statistical review focused on the 24 weeks efficacy data on Studies A4001027 and A1004028 provided in this submission. The A1004027 and A1004028 are two identically-designed ongoing, multicenter, randomized, double-blind, placebo-controlled trials of maraviroc, in combination with optimized background therapy versus optimized background therapy alone for the treatment of antiretroviral (ART) experienced HIV-1 CCR5-tropic HIV-1 infected adults.

2.2 Data Sources

The submissions under NDA022,128/N000 contain the efficacy results of the Studies A4001027 and A4001028.

The statistical review included the following:

1. Reviewing the protocols, statistical analysis plans, and efficacy results and conclusions in ‘Interim Full Clinical Study Report’ for the ‘24 Week Clinical Study Report’ (‘study-a4001027.pdf’ and ‘study-a4001027.pdf’).
2. Converting SAS transportable *.xpt files and conducting efficacy analyses following the instructions in their document files ‘define.pdf’ for the definitions of variables in the datasets of subdirectories ‘a4001027’, ‘a4001028’, ‘a4001029’, ‘virology’, etc., under

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

CDER Electronic Document Room (EDR) directory
\\Cdsesub1\evsprod\NDA022128\n0000\m5\datasets\.

3. Assisting other reviewers to compile datasets, and conduct statistical analyses.
4. Verifying sponsor's SAS log files, SAS program files and SAS xpt files such as the Time to Loss of Virologic Response (TLOVR) datasets submitted during March 23-April 4, 2007.

The sponsor has well documented variables in the datasets for each subdirectory in the 'define.pdf'. The quality of the efficacy data is acceptable. This reviewer can replicate the applicant's results of the estimated primary and key secondary efficacy endpoints using summary statistics such as efficacy endpoints at Week 24 in the datasets. After verifying some of the key efficacy variables, this reviewer did not find any differences with the applicant's data.

- The formats and variable names are consistent across studies with a few exceptions. For example, the indicator for subject ID (PID) has different lengths in different datasets. Datasets should be compiled better to delete unnecessary replicates and missing.

Originally, only SAS xpt files were submitted to the EDR. Per review team's persistent requests, the sponsor subsequently submitted SAS programs (received on March 27, 2007) and SAS log files (received on March 23, 2007) for review.

- When verifying SAS log files to compile the outcome datasets, it appears that the sponsor did not submit all the input datasets such as CSV files needed. However, the SAS programs and log files were logically correct with minor problems such as the utilities of some SAS Functions.

**Appears This Way
On Original**

3. STATISTICAL EVALUATION

Since the two Phase 3 Studies A1004027 (MOTIVATE-1) and A1004028 (MOTIVATE-2) are identically designed, except for the geographical difference, the statistical evaluations of maraviroc efficacy has been based on the datasets combining the two studies with a few exceptions.

3.1 Evaluation of Efficacy

3.1.1 Study Design

The Studies A1004027 and A1004028 are ongoing, multicenter, randomized, double-blind, placebo-controlled studies in patients infected with CCR5-tropic HIV-1. Eligible patients are those with more than 5,000 copies/mL of HIV-1 RNA, patients who had at least 6 months of prior therapy with at least one agent from three of the four antiretroviral drug classes listed below, or documented resistance or intolerance to at least one member of each class:

- ≥ 1 nucleoside reverse transcriptase inhibitors (NRTI);
- ≥ 1 non-nucleoside reverse transcriptase inhibitors (NNRTI);
- ≥ 2 protease inhibitors (PI); and
- and/or enfuvirtide (Fusion inhibitor).

Subjects were randomized in a ratio of 2:2:1 to one of the three arms:

- Optimized Background Therapy + maraviroc QD (300 mg dose equivalent once daily)
- Optimized Background Therapy + maraviroc BID (300 mg dose equivalent twice daily)
- Optimized Background Therapy

The randomization was stratified by screening HIV-1 RNA level ($<$ or $\geq 100,000$ copies/mL) and previous use of enfuvirtide in the background regimen.

Both A4001027 and A4001028 were superiority studies comparing the safety and antiviral activity of 2 maraviroc treatment regimens versus placebo, each in combination with Optimized Background Therapy (OBT). Identically-designed except for the geographical differences: the A1004027 enrolled subjects in North American, and the A1004028 enrolled subjects in the US, Australia and the European countries. Subjects were planned to be treated for 48 weeks.

The sample sizes would provide a greater than 95% power to detect a treatment difference of 0.5 \log_{10} copies/mL (maraviroc versus placebo), standard deviation = 0.8 \log_{10} copies/mL, a two-sided type I error of 0.025, by the Bonferroni adjustment for multiple comparisons.

There was no stopping rule for tropism shift in the absence of documented virologic failure.

3.1.2 Analysis Datasets and Definitions

In Studies A4001027 and A4001028, three analysis populations were used to report of safety and efficacy information: Full Analysis Sets (FAS), Per-Protocol (PP) and the Safety analysis set.

The definitions of the analysis populations are as follows.

- The FAS consists of all randomized subjects who receive at least 1 dose of study medication.
- The PP consists of all randomized subjects who receive at least 1 dose of study medication; who was treated for at least 14 days or discontinued before this time due to treatment failure; who has more than 80% compliant with randomized treatment; and who has no violation of any inclusion or exclusion criteria.
- Safety Analysis Set - consisting of all randomized subjects who receive at least 1 dose of study medication. Subjects will be reported in the dose group they actually receive.
- Both the FAS and PP sets were further defined according to the assignment of randomization or actual treatment received. For example, 'FAS – As Randomized' versus 'FAS – As Treated', 'PP – As Randomized' and 'PP – As Treated'.

For the primary efficacy endpoint, the applicant used both FAS and PP – As Randomized' and '– As Treated' datasets. For the secondary efficacy endpoints, only 'FAS – As Treated' and 'PP – As Treated' datasets were used. Safety Analysis Set was used to analyze the safety endpoints.

This reviewer conducted sensitivity analyses of primary and secondary efficacy endpoints using the 'FAS-As Randomized' except for a few occasions.

3.1.3 Primary Efficacy Endpoint, Hypotheses and Decision Rules

Primary efficacy endpoint is change from baseline in \log_{10} HIV-1 RNA at Weeks 24 and 48. This is a superiority trial to test H_0 versus H_A .

H_0 : difference in change from baseline \log_{10} HIV-1 RNA between maraviroc and placebo = 0

Versus

H_A : difference in change from baseline \log_{10} HIV-1 RNA between maraviroc and placebo >0.

Plasma HIV-1 RNA was determined at screening, randomization, Day 1, Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and at the time of early termination, using the RT-PCR Roche Amplicor v1.5 standard and ultra-sensitive assays. The standard assay has lower limit of quantification

LLOQ=400 copies/mL, and the ultra-sensitive has LLOQ=50 copies/mL. If a measurement was < 400 copies/mL, the Ultrasensitive assay was automatically performed and the second measurement was used for analysis. Otherwise, the results by the Standard assay were used. If a measurement was beyond the upper limit of quantification (ULOQ), different coding methods were observed (See a SAS program submitted in March 26, 2007).

- If a measurement > ULOQ=75,000 copies/mL in a standard assay, the measurement was assigned to 75,001 copies/mL, a dilution assay of the same specimen was performed and coded.
- If a measurement > ULOQ=10,000 copies/mL by the ultrasensitive assay, then this measurement was deleted, and the second measurement of the same specimen was used for analysis, regardless whether the second measurement is above or below 10,000 copies/mL.

Multiplicity was adjusted for the analysis of the primary efficacy endpoint, not secondary efficacy endpoints. When analyzing data for each study, A4001027 or A4001028 respectively, the two-sided 97.5% confidence interval (CI) to estimate the treatment difference of the primary efficacy endpoint, and 95% CIs were obtained for secondary efficacy endpoints. When analyzing data combining Studies A4001027 and A4001028, the two-sided 99.95% CI for the treatment difference (MVC-Placebo) of the primary efficacy endpoint, and the two-sided 99.9% CIs for the treatment differences of the secondary efficacy endpoints were estimated. If the confidence interval is completely to the left side and completely excluded 0, the superiority of maraviroc over placebo will be concluded.

3.1.4 Secondary Efficacy Endpoints

Secondary efficacy endpoints are defined as follows.

- 1) Percentage of subjects with HIV-1 RNA < 400 copies/mL at Weeks 24 and 48;
- 2) Percentage of subjects with HIV-1 RNA < 50 copies/mL at Weeks 24 and 48;
- 3) Percentage of subjects with HIV-1 RNA < 400 copies/mL or having at least 0.5 log₁₀ decrease in HIV-1 RNA from baseline at Weeks 24 and 48;
- 4) Percentage of subjects with HIV-1 RNA < 400 copies/mL or having at least 1.0 log₁₀ decrease in HIV-1 RNA from baseline at Weeks 24 and 48;
- 5) Time-average Difference (TAD) in log₁₀ HIV-1 RNA at Weeks 24 and 48;
- 6) Time to virologic failure at Week 48 using DAVP's loss-of-virologic-response TLOVR algorithm;
- 7) Time to treatment failure at Weeks 24 and 48;
- 8) Change from baseline in CD4+ (CD8+) cell count through Weeks 24 and 48;
- 9) HIV-1 tropism at baseline and at the time of failure;
- 10) Genotype and phenotype at screen, Weeks 24 and 48, at the time of failure or early termination visit;
- 11) Association between baseline resistance and virologic response;

Except for time to virologic failure and time to treatment failure which will be analyzed at Week 48, all other secondary endpoints will be analyzed at Weeks 24 and 48.

3.1.5 Results

3.1.5.1 Patient Disposition, Demographic and Baseline Characteristics

Randomization and FAS

In Study 4001027, 1816 subjects were recruited at screen, among them 601 (33.1%) were randomized and 585 (97.3%) were treated with maraviroc QD (n=232, MVC QD), maraviroc BID (n=235, MVC BID) and Placebo (n=118). Hence, the FAS population in Study 4001027 consists of 585 subjects.

In Study A4001028, 1428 subjects were recruited at screen, among which 475 (33.3%) were randomized and 464 (97.7%) were treated. Hence, the FAS population in Study 4001028 consists of 464 subjects.

In both studies, the main reason for patient's screen failure was presence of a dual/mixed or CXCR4 tropism.

Demographic and Baseline Characteristics

The demographic and baseline characteristics by treatment groups in the FAS are summarized in Table 1.

In A4001027, the FAS population consists of 88.9% subjects from America including Puerto Rico and 11.1% from Canada. 39-40.4% of the treated subjects had HIV-1 VL \geq 100,000 copies/mL, and 42.4-45.5% had previous treatment with enfuvirtide. Mean age was 46, 84-89% was male, and 81-84% was Caucasian, mean baseline viral load was 4.84-4.86 log₁₀ copies/mL and median CD4+ cell count was approximately 150-168 cells/mm³. At baseline, subjects had mean duration 13.9-14.3 years in HIV-1 diagnoses, 83.9-87.1% with protease inhibitor (PI) and/or Delavirdine in previous ARV therapy, 50.9-59.1% of the subjects had Genotypic Sensitivity Score (GSS) of 0 or 1 for the drug, 29.7-41.0% had Phenotypic Sensitivity Score (PSS) of 0 or 1, and 33.9-48.1% had an Overall Sensitivity Score (OSS) of 0 or 1.

In A4001028, the FAS population consists of 148 (31.9%) subjects from America, 42 (9.1%) from Australia and 274 (59.1%) from ten European countries: Belgium, France, Germany, Great Britain, Italy, Spain, Netherlands, Poland, Sweden and Swaziland. 42-42.4% of the treated subjects had HIV-1 VL \geq 100,000 copies/mL, and 37.4-44.0% had previous treatment with enfuvirtide. Mean age was 45-57, 90-91% was male, and 82-87% was Caucasian, mean baseline

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

viral load was 4.84-4.89 log₁₀ copies/mL and median CD4+ cell count was 174-182 cells/mm³. Subjects had mean duration 13.8-14.4 years in HIV-1 diagnoses, 64.8-79.1% with PI and/or Delavirdine in previous ARV therapy, 48.4-56.6% of the subjects had GSS of 0 or 1, 35.2-36.3% had PSS of 0 or 1, and 41.9-42.9% had OSS of 0 or 1.

The demographics and baseline characteristics appear to be balanced between treatment regimens in each study.

Patients' Disposition

The patient's discontinuation status and reasons for discontinuation through the date of Week 24 cut-off (not at Week 24 from Day 1) are summarized in Table 2.

In A4001027, the 35.8%, 34.0% and 62.7% of the subjects respectively in the MVC QD, MVC BID and Placebo regimen discontinued at Week 24. The majority of the dropouts was due to insufficient efficacy: 49 (59.0%), 56 (70.0%) and 59 (79.7%), adverse event: 12 (14.5%), 10 (12.5%) and 6 (8.0%), respectively in the MVC QD, MVC BID and Placebo regimen.

In A4001028, 33.0, 30.4 and 64.8% of the subjects in the MVC QD, MVC BID and Placebo regimen respectively, discontinued at Week 24. The dropouts were mainly due to insufficient efficacy: 17.6, 18.3 and 51.7% respectively, in the MVC QD, MVC BID and Placebo regimen.

Overall, 414 (39.5%) subjects discontinued from studies at Week 24. This includes 133 (63.6%) in the Placebo regimen and 286 (34.0%) in the two MVC regimens. The main reason for the dropouts was attributable to lack of efficacy. The relative risks of lack of efficacy MVC (QD, BID) versus Placebo were (2.4, 2.1) in Study 4001027 and (2.9, 2.8) in Study 4001028, respectively. No significant differences were found between the two MVC regimens.

The overall discontinuation rates were significantly lower in the two MVC regimens than the Placebo regimens. This feature was not uncommon in HIV-1 drug trials. However, it may create a marked effect on estimation of treatment differences MVC versus Placebo, both on efficacy and safety. Statistical methods should be applied to test sensitivity of the conclusions.

**Appears This Way
On Original**

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

Table 1. Studies A4001027 and A4001028: Demographics and Baseline Characteristics*

Characteristic	Study A4001027			Study A4001028		
	Maraviroc QD	Maraviroc BID	Placebo	Maraviroc QD	Maraviroc BID	Placebo
N	232	235	118	182	191	91
Male Sex, n (%)	210 (91)	212 (90)	106 (90)	153 (84)	170 (89)	79 (87)
White Race, n (%)	187 (81)	197 (84)	99 (84)	149 (82)	166 (87)	79 (87)
Mean Age (range), yrs	46 (19-75)	46 (25-69)	46 (31-71)	45.2 (17-75)	47.0 (21-73)	45.3 (29-72)
Mean HIV-1 RNA (SD), log ₁₀ copies/mL	4.85 (0.641)	4.86 (0.614)	4.84 (0.556)	4.87 (0.664)	4.84 (0.621)	4.89 (0.696)
Screening Stratum of HIV-1 RNA level, n (%)						
<100,000 copies/mL	135 (58.2)	139 (59.1)	70 (59.3)	103 (56.6)	104 (54.5)	53 (58)
≥100,000 copies/mL	93 (40.1)	95 (40.4)	46 (39.0)	77 (42.3)	81 (42.4)	38 (42)
Screening Stratum of Envirvutide Use						
Yes	100 (43.1)	107 (45.5)	50 (42.4)	68 (37.4)	76 (39.8)	40 (44.0)
No	132 (56.9)	128 (54.5)	68 (57.6)	114 (62.6)	115 (60.2)	51 (56.0)
Median CD4+ Cell Count (range), cells/μL	167.5 (1.0 – 811.5)	150.0 (2.0 – 677.5)	163.3 (1.0 – 675.0)	174.3 (0.5 – 965.5)	182.0 (3.0 – 820.0)	174.3 (2.0 – 544.5)
Mean Duration of Diagnosis (years)	14.0 (1.0-27.8)	13.9 (2.3-24.3)	14.3 (3.4-25.1)	14.3 (5.1-23.1)	13.8 (4.1-26.1)	14.4 (4.1-24.0)
PI ^a and/or Delavirdine in OBT, n (%)	202 (87.1)	191 (81.3)	99 (83.9)	118 (64.8)	144 (75.4)	72 (79.1)
Genotypic Susceptibility Score (GSS) – n (%)						
0	52 (22.4)	59 (25.1)	31 (26.3)	39 (21.4)	43 (22.5)	20 (22.0)
1	82 (35.3)	80 (34.0)	29 (24.6)	64 (35.2)	58 (30.4)	24 (26.4)
2	38 (16.4)	48 (20.4)	21 (17.8)	25 (13.7)	32 (16.8)	20 (22.0)
≥3	57 (24.6)	47 (20.0)	34 (28.8)	52 (28.6)	57 (29.8)	25 (27.5)
Phenotypic Susceptibility Score (PSS) – n (%)						
0	25 (10.8)	24 (10.2)	17 (14.4)	20 (11.0)	26 (13.6)	12 (13.2)
1	70 (30.2)	73 (31.1)	18 (15.3)	46 (25.3)	42 (22.0)	20 (22.0)
2	51 (22.0)	69 (29.4)	35 (29.7)	42 (23.1)	38 (19.9)	23 (25.3)
≥3	83 (35.8)	66 (28.1)	45 (38.1)	71 (39.0)	84 (44.0)	34 (37.4)
Overall Susceptibility Score (OSS) – n (%)						
0	30 (12.9)	27 (11.5)	19 (16.1)	22 (12.1)	30 (15.7)	16 (17.6)
1	78 (33.6)	86 (36.6)	21 (17.8)	55 (30.2)	50 (26.2)	23 (25.3)
2	51 (22.0)	65 (27.7)	38 (32.2)	37 (20.3)	39 (20.4)	21 (23.1)
≥3	69 (29.7)	54 (23.0)	37 (31.4)	65 (35.7)	71 (37.2)	29 (31.9)
Δ32 Genotype (W/W, W/D)	200/17	207/13	101/11	157/15	166/15	75/5

^a Except for tipranavir/ritonavir.

ARVs = Antiretroviral agents; QD = Once daily dosing; BID = Twice daily dosing; W/W = Wild-type, wild-type; W/D = Wild-type, deletion.

* Modified Table 5, Antiviral Drugs Advisory Committee Briefing Document, April 24, 2007.

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 2. Studies A4001027 and A4001028: Patient Disposition*

	Maraviroc QD	Maraviroc BID	Placebo
Study A4001027			
FAS as Randomized	232	235	118
Completer through Week 24	149 (64.2)	155(66.0)	44 (37.3)
Discontinuation	83 (35.8)	80 (34.0)	74 (62.7)
Death	1	1	1
Insufficient Response	49 (21.1)	56 (23.8)	59 (50.0)
Adverse Event	12	10	6
Lost to follow-up	8	5	2
Consent withdrawal	12	10	5
Protocol violation	4	1	3
Other	6	1	0
Study A4001028			
FAS as Randomized	182	191	91
Completer through Week 24	122 (67.0)	133(69.6)	32 (35.2)
Discontinuation	60 (33.0%)	58 (30.4%)	59 (64.8%)
Death	2	4	0
Insufficient Response#	32 (17.6%)	35 (18.3%)	47 (51.7%)
Adverse Event	7	6	2
Lost to follow-up	3	3	1
Consent withdrawal	11	7	6
Protocol violation	3	1	2
Other	2	2	1

* Through Date of Week 24 Data Cut-off. #. Insufficient clinical response/

Appears This Way
 On Original

Statistical Methodologies

In the primary evaluation of the primary efficacy endpoint by the sponsor,

- Time window for Week 24 (Day 169) was between Week 22 (Day 155) and Week 28 (Day 196).
- The mean HIV-1 RNA viral load (VL) at screening, at randomization and at Day 1 prior to treatment using the study regimen was used as the baseline HIV-1 VL.
- A missing VL in change from baseline in \log_{10} HIV-1 RNA at Week 24 has been assigned to 1) 'no change' for those with no baseline or on-study HIV-1 VL (n=7 in Study A4001027 n=8 in Study A4001028), and those who discontinued from study prior to Week 24, and 2) the last observation prior to Week 24 for those who did not discontinue at Week 24 but no HIV-1 VL in the Week 24 time window, known as the last observation carry forward (LOCF) approach.
- The mean treatment differences (MVC-Placebo) and 97.5% CIs were obtained using Analysis of Covariance Models (ANCOVA) with treatment, and two randomization strata as explanatory variables.

Three sensitivity analyses of the primary efficacy endpoint were conducted as described in SAP of Protocol A4001027 and Protocol A4001028 (Source: P 3282-3283, a4001027.pdf, a401028.pdf):

- *Sensitivity 1 (LOCF)*: LOCF will be used. This analysis will include only those subjects with an assessment of HIV-1 RNA at baseline and at least one assessment while on study treatment, else will be missing. For subjects who discontinue the study prematurely, the last value on blinded study therapy will be used.
- *Sensitivity 2 (LOCF / No Change)*: Subjects who discontinued before or at Week 24 (or 48) due to any reason, apart from protocol defined treatment failure (see Section 6.1.2 *Time to Treatment Failure* for definition) will impute a change from baseline of 0. Subjects who have missing baseline assessments will impute a change from baseline of 0. The LOCF approach will be used for missing data for protocol defined treatment failures and subjects who are not discontinued.
- *Sensitivity 3 (Treatment failure classification)*: Those patients who have met at least 1 of the treatment failure criteria but have not been discontinued by the investigator from the study will be considered as treatment failures.

In evaluation of the secondary efficacy endpoints,

- The continuous variables are TAD through Week 24, change from baseline to Week 24 in CD4+ and CD8+ cell count. Handling missing in Week 24 TAD was the same as for the primary efficacy endpoint. For change from baseline to Week 24 in CD4+ and CD8+ cell counts, LOCF was used for missing data. ANCOVA models were used for the above three secondary endpoints.

- The dichotomous variables are proportion of subjects with HIV-1 VL <400, <50, <400 or at least 1 log₁₀ reduction from baseline to Week 24, <400 or at least 0.5 log₁₀ reduction from baseline to Week 24. For missing data, 'non-responder' is generally assigned. However, if a subject has a missing value at Week 24 but their value is <400/50 copies/mL at both Weeks 20 and 32, then they will be considered a responder. If either their Week 20 or Week 32 value is missing, or they have discontinued prior to Week 32 then they will be considered a non-responder. Logistic regression models with treatment and two randomization strata as explanatory variables were used to estimate the treatment differences and 95% CIs for these secondary dichotomous efficacy variables.
- Time windows are the same as in the primary efficacy endpoint.

All the above analyses were conducted for FAS and PP both as randomized and as treated population.

This reviewer verified the protocol-defined primary efficacy endpoint and selected secondary efficacy endpoints. Figure 4 shows the histograms of the primary efficacy endpoint by treatment regimen. The distributions of change from baseline to Week 24 in HIV-1 VL appear to be skewed due to heavy discontinuations. As a result, this reviewer estimated the median treatment differences (MVC-Placebo) in HIV-1 VL using method by Hodges and Lehman (H-L method). Sensitivity analyses different from the applicant's were conducted. For example, datasets in sub-study populations, data including off study information, etc. were also considered to evaluate the robustness of the protocol-defined statistical approaches in evaluation of efficacy endpoints, especially the primary efficacy endpoint.

Multiplicity was adjusted for the analysis of the primary efficacy endpoint, not secondary efficacy endpoints. When analyzing data for each study, A4001027 or A4001028 respectively, the two-sided 97.5% confidence interval (CI) was obtained for the treatment difference of the primary efficacy endpoint, and 95% CIs were obtained for the treatment difference of secondary efficacy endpoints. When analyzing data combining Studies A4001027 and A4001028, the two-sided 99.95% CI for the treatment difference (MVC-Placebo) of the primary efficacy endpoint, and the two-sided 99.9% CIs for the treatment differences of the secondary efficacy endpoints were estimated.

**Appears This Way
On Original**

3.1.5.2 Estimated Treatment Differences in Primary Efficacy Endpoint Using Different Baseline VL

There was a concern whether the baseline VL may be influenced by the ART prior to Day 1, based on data pooling Studies A4001027 and A4001028. A mean time window of 5.9 weeks (mean=41 days, range -86 to -5 days prior to Day 1) between screening and Day 1 was obtained. At screening, this study population had a mean of 14 years of HIV-1 infections, and most of the subjects were on stabled ART for at least 4 weeks.

To investigate the effect of baseline VL in estimating the primary efficacy endpoint, this reviewer conducted analyses to compare VL at Day 1 prior to treatment with study drugs (Day 1-VL) and the sponsor's baseline VL, defined as average of HIV-1 VL at screening, at randomization and Day 1 prior to treatment with study drugs (Average-VL). In the text below, we refer to the two types of baseline VL as Day 1-VL and Average-VL.

In the evaluations of primary efficacy endpoint using Day 1-VL and Average-VL, other criteria such as handling missing in Week 24 time window and modeling remain the same as the sponsor's.

- A missing VL in change from baseline in \log_{10} HIV-1 RNA at Week 24 has been assigned to 1) 'no change' for those with no baseline or on-study HIV-1 VL, and those who discontinued from study prior to Week 24, and 2) the last observation prior to Week 24 for those who did not discontinued at Week 24 but no HIV-1 VL in the Week 24 time window, known as the last observation carry forward (LOCF) approach.
- An ANCOVA model which includes treatment regimens, screening HIV-1 VL strata ($<$ or \geq 100,000 copies/mL), and enfuvirtide use in the ART as explanatory variables was used to estimate the mean treatment differences in change from baseline to Week 24 in VL.

Figure 1 shows histogram of difference between the Day 1-VL and Average-VL. Overall, the Day 1-VL appears to be lesser than the Average-VL, $p=0.053$, by the signed rank test. The mean (median) difference (Day 1-VL – Average-VL) is -0.003 (-0.013) \log_{10} copies/mL, with a range of $(-1.418, 0.773)$ copies/mL. Percentages of subjects with such differences < -0.3 , -0.2 , and -0.1 \log_{10} copies/mL are 5.3%, 9.7% and 23.9% respectively.

Tables 3 and 4 summarize the results of fitting the same ANCOVA model but using different baseline VL, for Studies A4001027, A4001028, respectively (Table 3) and for the two

studies combined (Table 4). The results using the ANOVA model with treatment, randomized strata indicators in the model are summarized below.

- For each study, the differences in estimated adjusted mean reduction from baseline to Week 24 using different baseline VL were less than 0.2 log₁₀ copies/mL across three treatment groups. Hence, the estimated mean treatment differences MVC-Placebo using different baseline VL methods were similar. However, the estimated 97.5% CIs using the Day 1-VL appear to be slightly wider because the standard errors were greater using the Day 1-VL as baseline than those using the Average-VL as baseline.
- Similar results were obtained for the two studies combined: the differences in estimated adjusted mean reduction from baseline to Week 24 using different baseline VL were about 0.14 log₁₀ copies/mL across three treatment groups and the estimated mean treatment differences MVC-Placebo using different baseline VL methods were similar. However, the estimated 97.5% CIs using the Day 1-VL appear to be slightly wider because the standard errors were greater using the Day 1-VL as baseline than those using the Average-VL as baseline.
- Please note that the 99.95% CIs for data pooling two studies should be used for multiplicity concern. This reviewer used 97.5% CIs so that the results should be comparable.

Day 1-VL should be more appropriate for the continuous efficacy endpoints such as change from baseline in VL and time average difference (TAD) for this ART experienced study populations. However, the estimated treatment differences regarding change from baseline to Week 24 in VL appear to be robust for different baseline VL.

3.1.5.3 Sensitivity Analyses of Primary Efficacy Endpoint

To verify whether the estimated efficacy sizes by the sponsor are representative of the true effects of maraviroc, and to examine the potential effects of substantial discrepancies in the discontinuation status between placebo and maraviroc groups, this reviewer conducted three types of sensitivity analyses on primary efficacy endpoint as follows.

- Analysis using all available HIV-1 VL data regardless of whether a subject was on study drug.
- Analysis on completers at Week 24.
- Imputation of missing by -0.4 to 0.3 log₁₀ copies/mL with an increment of 0.1 log₁₀ copies/mL.

The detailed methodology used in the sensitivity analyses was as follows.

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

- The treatment difference on the primary endpoint was evaluated using both mean and median difference. Applied Hodges-Lehman approach, the median treatment difference is the median of n_1 by n_2 pairs in treatment difference between the two treatment groups¹ where n_1 and n_2 are sample sizes for the two groups. A distribution-free confidence interval (*Moses*) was then be constructed based on a Wilcoxon Rank Sum Test¹.
- Type I error was adjusted when data were combined for A4001027 and A4001028 ($p=0.001$), and also the two comparisons MVC QD or MVC BID versus Placebo ($p=0.0005$). Hence, 99.95% confidence intervals of the mean and median treatment differences were estimated.
- The mean or median treatment differences are sample summary statistics without adjusting for their randomization strata.
- All the sensitivity analyses used Day 1-VL as baseline. If there was more than one value in the Week 24 time window, the one closest to Day 168 was selected.

The results of these sensitivity analyses are summarized in Tables 5 and 6.

The Sensitivity Analysis 1 used all available HIV-1 VL data regardless whether a subject was on study drug, meaning that those VL data after the subjects in the placebo group had switched to MVC+ OBT were included as if they had been treated with Placebo+OBT. We obtained an additional mean reduction of 0.5 \log_{10} copies/mL and median reductions of 0.45~0.46 \log_{10} copies/mL in the two MVC regimens compared to the Placebo. The 99.95% CIs were all excluding zeros, indicating the superiority of MVC compared to Placebo. The treatment difference between the two MVC regimens is within 0.01 \log_{10} copies/mL.

The Sensitivity Analysis 2 was conducted among completers, defined as the *earliest date of two events* exceeded Day 155 since Day 1: (1) the date of discontinuation from study and (2) the last date of treatment with study regimen. HIV-1 VL data were extended to 1 week from the date of discontinuation or the date of stopping treatment with study regimen. Subjects in the MVC regimens showed an additional mean reductions 0.48 \log_{10} and 0.54 \log_{10} copies/mL respectively in the MVC-QD and MVC-BID regimen, compared to placebo. The extra median reductions were 0.38 \log_{10} and 0.44 \log_{10} respectively in the MVC-QD and MVC-BID regimen, compared to placebo. The superiority of MVC compared to Placebo was confirmed using mean comparisons. Results by the Hodges-Lehman approach support the superiority of MVC-BID ($p<0.0005$), not MVC-QD ($p>0.0005$), at the type I error of 0.0005 level.

The third sensitivity analysis (Analyses 3-10) contained eight imputations to impute the missing of Week 24 HIV-1 VL values from -0.4 to 0.3 log₁₀ copies/mL with an increment of 0.1 log₁₀ copies/mL. As the imputed value increasing from -0.4 to 0.3 log₁₀ copies/mL, the mean treatment difference in change from baseline to Week 24 ranging -0.79 to -0.98 log₁₀ copies/mL for (MVC QD-Placebo), and -0.88 to -1.07 log₁₀ copies/mL for (MVC BID-Placebo). The estimated median treatment differences were slightly less than the mean differences. As the imputed value increasing from -0.4 to 0.3 log₁₀ copies/mL, the median treatment difference in change from baseline to Week 24 ranging -0.69 to -0.88 log₁₀ copies/mL for MVC QD-Placebo, and -0.80 to -0.99 log₁₀ copies/mL for MVC BID-Placebo.

More than 50% of the subjects in the placebo regimen had discontinued from study by Week 24. Hence, if one imputes a single value δ to the missing, then δ will be the median of the sample in the placebo group after imputation. For example, after imputing missing with -0.3 log₁₀ copies/mL, the median change from baseline to Week 24 should be -0.3 log₁₀ in placebo group.

We conclude that the superiority of MVC versus placebo in estimating primary efficacy endpoint is essentially well maintained by the sensitivity analyses, even though some of the sensitivity analyses were rather conservative and were designed *not in favor of* maraviroc treatment.

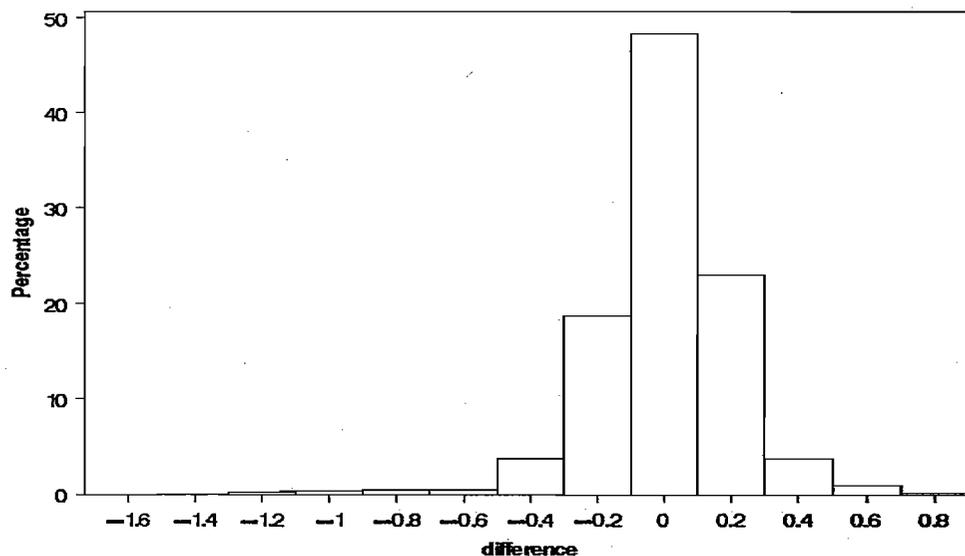


Figure 1. Baseline VL in Studies A4001027 & A4001028: Day 1 VL- Average VL

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

Study 1027: Baseline HIV VL: Average v. Last ($Y=0.56+0.88x$)

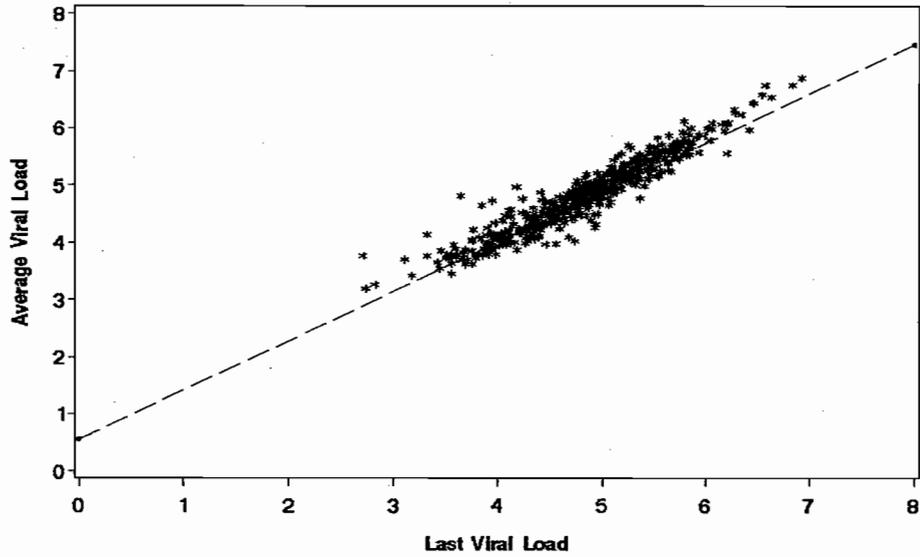


Figure 2. Study A4001027: Baseline HIV VL: Average V. Day 1

Study 1028: Baseline HIV VL: Average v. Last ($Y=0.71+0.86x$)

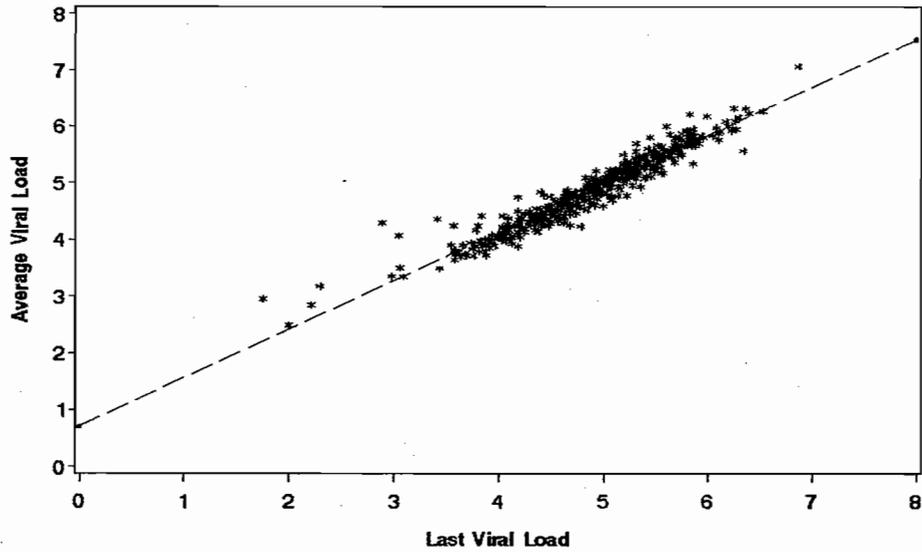


Figure 3. Study A4001028: Baseline HIV VL: Average V. Day 1

Table 3. Change from Baseline to Week 24 in Viral Load by Study

	Mean	Se	Mean Treatment Difference and 97.5% CI
Study A4001027			
ANCOVA Results by the Sponsor²			
Maraviroc QD (n=232)	-1.812	0.092	-0.788 (-1.141,-0.435)
Maraviroc BID (n=235)	-1.952	0.091	-0.922 (-1.275,-0.570)
Placebo (n=118)	-1.030	0.129	
ANCOVA Results Using Day 1-VL as Baseline²			
Maraviroc QD (n=232)	-1.636	0.130	-0.798 (-1.259,-0.336)
Maraviroc BID (n=235)	-1.762	0.132	-0.923 (-1.378,-0.459)
Placebo (n=118)	-0.839	0.160	
Study A4001028			
ANCOVA Results by the Sponsor²			
Maraviroc QD (n=182)	-1.950	0.105	-1.021 (-1.426,-0.616)
Maraviroc BID (n=191)	-1.971	0.103	-1.042 (-1.444,-0.640)
Placebo (n=91)	-0.929	0.147	
ANCOVA Results Using Day 1-VL as Baseline²			
Maraviroc QD (n=182)	-1.863	0.142	-1.019 (-1.533,-0.505)
Maraviroc BID (n=191)	-1.889	0.141	-1.045 (-1.558,-0.531)
Placebo (n=91)	-0.844	0.180	

1. Basic statistics: mean, standard error (se) and confidence interval (CI).

2. ANCOVA model includes treatment, screening HIV-1 VL, and enfuvirtide use.

Table 4. Studies A4001027 and A4001028: Change from Baseline to Week 24 in VL

	Mean	Standard Error	Treatment difference and 97.5% CI
ANCOVA Results Using Day 1 HIV-1 VL as Baseline¹			
Maraviroc QD (n=414)	-1.740	0.096	-0.893 (-1.236,-0.550)
Maraviroc BID (n=426)	-1.822	0.096	-0.975 (-1.319,-0.632)
Placebo (n=209)	-0.847	0.119	
Sponsor's ANCOVA Results (Average-VL as Baseline)^{2,3}			
Maraviroc QD (n=414)	-1.876	0.069	-0.888 (-1.153,-0.623)
Maraviroc BID (n=426)	-1.960	0.068	-0.973 (-1.237,-0.709)
Placebo (n=209)	-0.987	0.097	

1. Model includes treatment, screening HIV-1 VL (< or ≥ 100,000 copies/mL) and enfuvirtide use.

2. Treatment Failure and Missing as No Change.

3. Source: Table 9, Antiviral drugs advisory committee briefing document.

Studies 1027 & 1028: Mean Change in HIV-1 VL from Baseline at Week 24

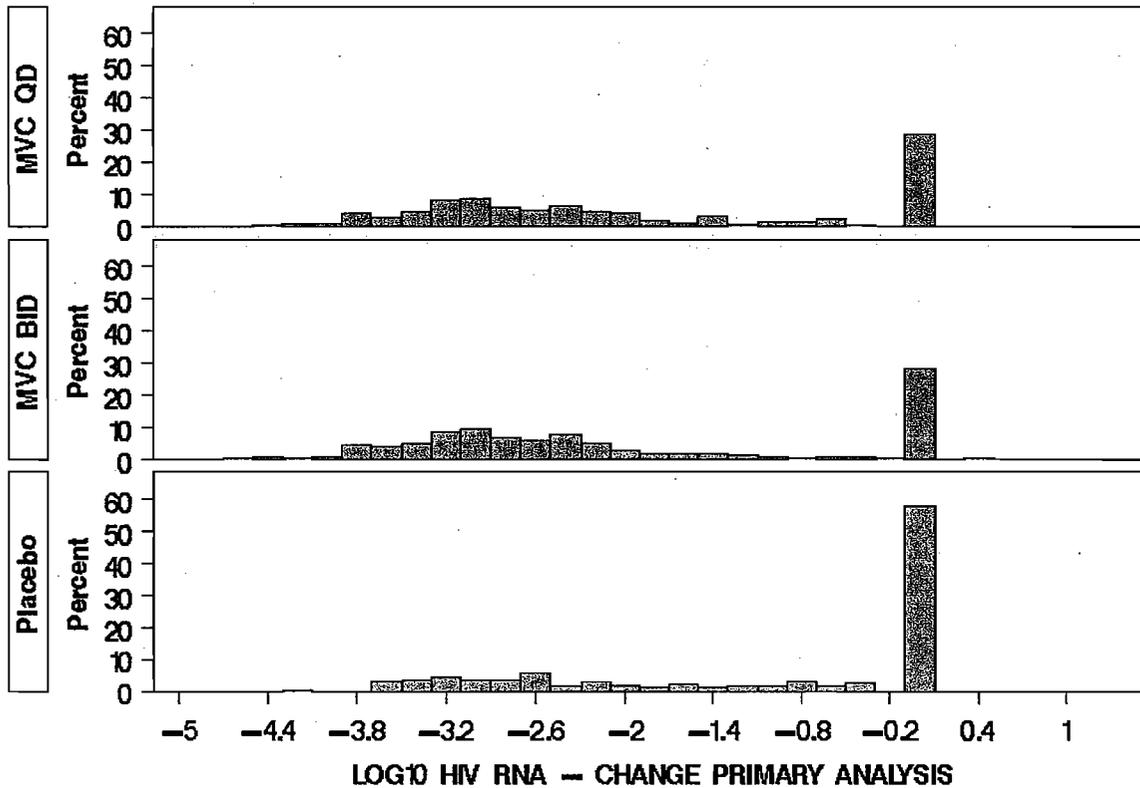


Figure 4. Histograms - Change from Baseline to Week 24 in VL

Appears This Way
On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 5. Sensitivity Analysis: Mean Change from Baseline to Week 24 in Viral Load

	n	mean	se	Treatment Difference			
				mean	se	99.95% CI	
Analysis 1: Regardless of Switching Treatment (n=871)							
MVC-QD	334	-2.33	0.07	-0.49	0.12	-0.89	-0.08
MVC-BID	366	-2.34	0.07	-0.50	0.12	-0.90	-0.09
Placebo	171	-1.84	0.10				
Analysis 2: Completers (n=707)							
MVC-QD	294	-2.58	0.06	-0.47	0.13	-0.93	-0.02
MVC-BID	315	-2.65	0.05	-0.55	0.13	-1.00	-0.10
Placebo	98	-2.10	0.12				
Analysis 3: Imputing Missing with -0.4 log₁₀ (n=1049)							
MVC-QD	414	-1.95	0.06	-0.75	0.10	-1.11	-0.40
MVC-BID	426	-2.07	0.06	-0.87	0.10	-1.22	-0.51
Placebo	209	-1.20	0.08				
Analysis 4: Imputing Missing with -0.3 log₁₀ (n=1049)							
MVC-QD	414	-1.93	0.06	-0.78	0.11	-1.15	-0.41
MVC-BID	426	-2.04	0.06	-0.89	0.11	-1.26	-0.53
Placebo	209	-1.15	0.08				
Analysis 5: Imputing Missing with -0.2 log₁₀ (n=1049)							
MVC-QD	414	-1.90	0.07	-0.81	0.11	-1.18	-0.43
MVC-BID	426	-2.02	0.07	-0.92	0.11	-1.30	-0.54
Placebo	209	-1.10	0.09				
Analysis 6: Imputing Missing with -0.1 log₁₀ (n=1049)							
MVC-QD	414	-1.88	0.07	-0.83	0.11	-1.22	-0.44
MVC-BID	426	-1.99	0.07	-0.95	0.11	-1.33	-0.56
Placebo	209	-1.05	0.09				
Analysis 7: Imputing Missing with 0 log₁₀ (n=1049)							
MVC-QD	414	-1.85	0.07	-0.86	0.11	-1.26	-0.46
MVC-BID	426	-1.97	0.07	-0.97	0.11	-1.37	-0.58
Placebo	209	-0.99	0.09				
Analysis 8: Imputing Missing with 0.1 log₁₀ (n=1049)							
MVC-QD	414	-1.82	0.07	-0.88	0.12	-1.29	-0.47
MVC-BID	426	-1.94	0.07	-1.00	0.12	-1.41	-0.59
Placebo	209	-0.94	0.09				
Analysis 9: Imputing Missing with 0.2 log₁₀ (n=1049)							
MVC-QD	414	-1.80	0.07	-0.91	0.12	-1.33	-0.49
MVC-BID	426	-1.92	0.07	-1.03	0.12	-1.45	-0.61
Placebo	209	-0.89	0.10				
Analysis 10: Imputing Missing with 0.3 log₁₀ (n=1049)							
MVC-QD	414	-1.77	0.07	-0.94	0.12	-1.37	-0.50
MVC-BID	426	-1.89	0.07	-1.05	0.12	-1.48	-0.62
Placebo	209	-0.84	0.10				

Table 6. Sensitivity Analysis: Median Change from Baseline to Week 24 in Viral Load

	N	median	Med ₁ - Med ₃	minimum	maximum	Treatment Difference			
						median	99.95% CI		
Analysis 1: Regardless of Switching Treatment (n=871)									
MVC-QD	334	-2.614	-0.473	-5.579	5.753	-0.442	-0.853	-0.057	
MVC- BID	366	-2.649	-0.508	-5.696	5.874	-0.454	-0.867	-0.082	
Placebo	171	-2.141							
Analysis 2: Completers (n=707)									
MVC-QD	294	-2.737	-0.401	-5.546	4.348	-0.381	-0.874	0.043	
MVC- BID	315	-2.809	-0.473	-5.663	4.367	-0.445	-0.940	-0.029	
Placebo	98	-2.336							
Analysis 3: Imputing Missing with -0.4 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-1.851	-5.138	4.140	-0.605	-1.281	-0.008	
MVC- BID	426	-2.368	-1.968	-5.255	4.140	-0.786	-1.490	-0.217	
Analysis 4: Imputing Missing with -0.3 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-1.951	-5.138	4.140	-0.605	-1.331	-0.022	
MVC- BID	426	-2.368	-2.068	-5.255	4.140	-0.831	-1.562	-0.231	
Analysis 5: Imputing Missing with -0.2 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.051	-5.138	4.140	-0.659	-1.372	-0.029	
MVC- BID	426	-2.368	-2.168	-5.255	4.140	-0.831	-1.624	-0.206	
Analysis 6: Imputing Missing with -0.1 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.151	-5.138	4.140	-0.669	-1.413	-0.059	
MVC- BID	426	-2.368	-2.268	-5.255	4.140	-0.877	-1.655	-0.204	
Analysis 7: Imputing Missing with 0.0 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.251	-5.138	4.140	-0.726	-1.485	-0.047	
MVC- BID	426	-2.368	-2.368	-5.255	4.140	-0.919	-1.718	-0.231	
Analysis 8: Imputing Missing with 0.1 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.351	-5.138	4.240	-0.746	-1.520	-0.100	
MVC- BID	426	-2.368	-2.468	-5.255	4.240	-0.941	-1.738	-0.231	
Analysis 9: Imputing Missing with 0.2 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.451	-5.138	4.340	-0.785	-1.567	-0.153	
MVC- BID	426	-2.368	-2.568	-5.255	4.340	-0.987	-1.788	-0.257	
Analysis 10: Imputing Missing with 0.3 log₁₀ (n=1049)									
MVC-QD	414	-2.251	-2.551	-5.138	4.440	-0.808	-1.597	-0.153	
MVC- BID	426	-2.368	-2.668	-5.255	4.440	-1.020	-1.840	-0.300	

1. i=1, Maraviroc QD, i=2, Maraviroc BID, 3-Placebo. 2. Based on n_i × n_j pairs, i=1,2.

Appears This Way
 On Original

3.1.5.4 Alternative Analysis of Selected Secondary Efficacy Endpoint

Time to Discontinuation at Week 24

Table 2 summarizes subjects' disposition and reasons for continuation for Studies A4001027 and A4001028, respectively. As of the date of Week 24 cut-off, percentages of subjects who discontinued were 62.7% and 64.8%, respectively for Studies A4001027 and A4001028. The main reason was insufficient clinical response: more than 50% of the subjects in the placebo groups and approximately 20% of subjects in the MVC groups discontinued from the study due to insufficient clinical responses. It appears that the significant heterogeneity of the discontinuation status, mainly by insufficient clinical response, should have great impact on the evaluation of maraviroc efficacy and safety.

Time to discontinuation was investigated using the Kaplan-Meier approach.

Figures 5 and 6 show the overall Kaplan-Meier (K-M) curves in time to discontinuation by treatment group (week). It is observed that subjects in the placebo regimen had significant fast pace in discontinuation than the two maraviroc regimens.

The K-M estimates in time to discontinuation by treatment group (week) and randomized stratum were provided in Figures 7-14. In addition to the significant longer time to discontinuation in the maraviroc regimens compared to the placebo, other observations are summarized as follows.

Among subjects with Enfuvirtide use in ART prior to trial,

- In Study A4001027, the subjects in the maraviroc 300 mg QD regimen had slightly better results than the maraviroc 150 mg BID regimen ($p=0.19$, log-rank test). In study 1028, however, the subjects in the maraviroc 150 mg BID regimen had slightly better results than the maraviroc 300 mg QD regimen ($p>0.20$).

Among subjects without Enfuvirtide use in ART prior to trial,

- In Study A4001027, the subjects in the maraviroc 150 mg BID regimen had slightly longer time to discontinuation than the maraviroc 300 mg QD regimen ($p=0.18$, log-rank test). On contrary, in study 1028, the subjects in the maraviroc 300 mg QD regimen had slightly longer time to discontinuation than the maraviroc 150 mg BID regimen ($p>0.20$), and each regimen had shorter follow-up between 28-32 weeks.

Among subjects with HIV viral load $< 100,000$ copies/mL, no significant differences were found at $p=0.2$ level.

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation.

- In study 1027, the subjects in the maraviroc 300 mg QD regimen had slightly better results than the maraviroc 150 mg BID regimen, before Week 20, and vice versa after Week 20. Similar observations were seen in study 1028, except for the shorter follow-up between 28-32 weeks for subjects in the maraviroc 150 BID and placebo regimens.

Among subjects with HIV viral load \geq 100,000 copies/mL, no significant differences were found at $p=0.2$ level.

- In studies A4001027 and A4001028, the subjects in the maraviroc 150 mg BID regimen had slightly better results than the maraviroc 300 mg QD regimen. Subjects in placebo regimen had a shorter follow-up of 32-36 weeks than other the maraviroc regimens (48 weeks).

We have the following conclusions. Using the Kaplan-Meier method (K-M), time to discontinuation was significantly longer in the maraviroc QD and BID regimens compared with the placebo regimen, $p<0.0001$ by the log-rank test for A4001027 and A4001028, respectively. Subgroup analyses suggested that time to discontinuation may be associated with the previous use of enfuvirtide in ART but not associated with screening HIV-1 VL level at a significant level of $p=0.20$. Different temporal patterns in different studies were observed. In A4001027 among those who used enfuvirtide in ART, the subjects receiving MVC BID were doing somewhat better than those receiving MVC QD. Conversely, in A4001028 among the non-enfuvirtide users in ART, the subjects in the MVC QD group were doing somewhat better than the MVC BID group. However, these qualitative interactions were based on the univariate analyses (K-M) and at the significance level of 0.2.

Appears This Way
On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

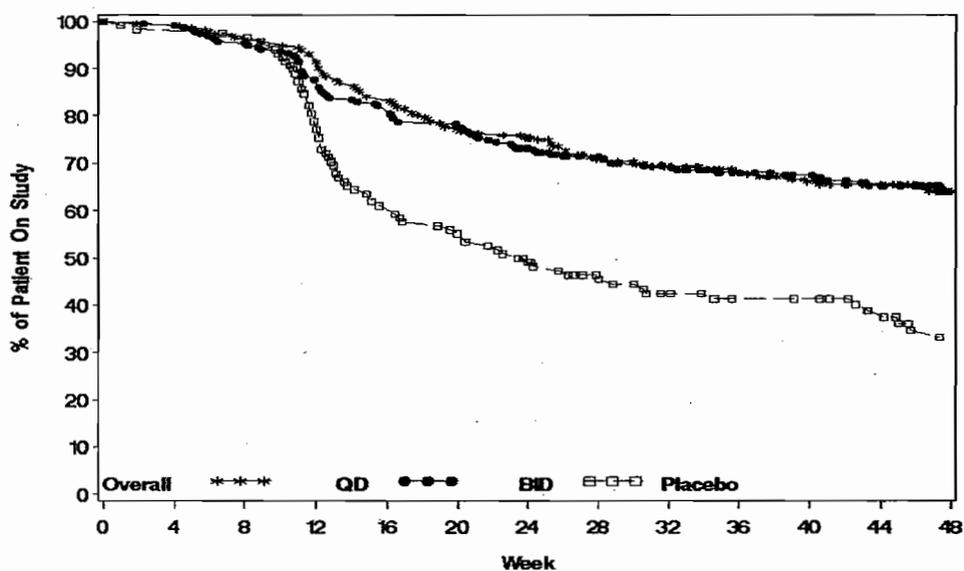


Figure 5. Study A4001027: Time to Discontinuation

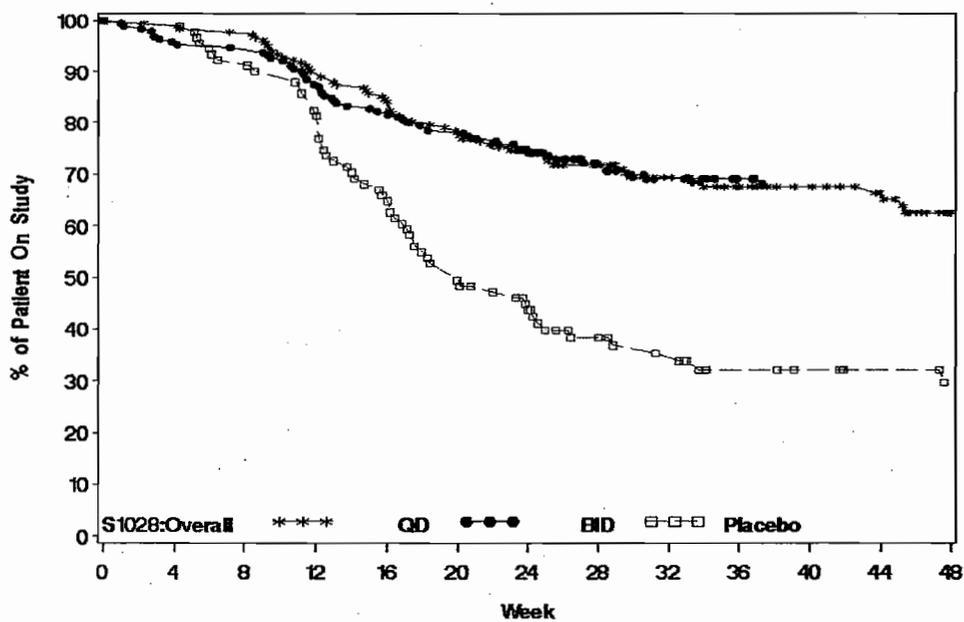


Figure 6. Study A4001028: Time to Discontinuation

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

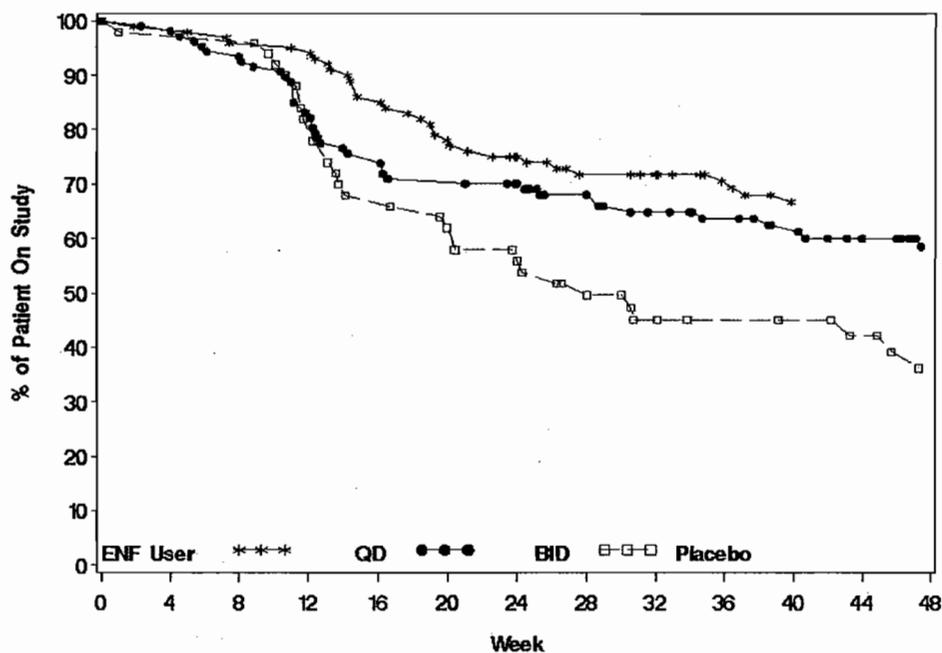


Figure 7. Study A4001027: Time to Discontinuation (Enfuvirtide Use in ART)

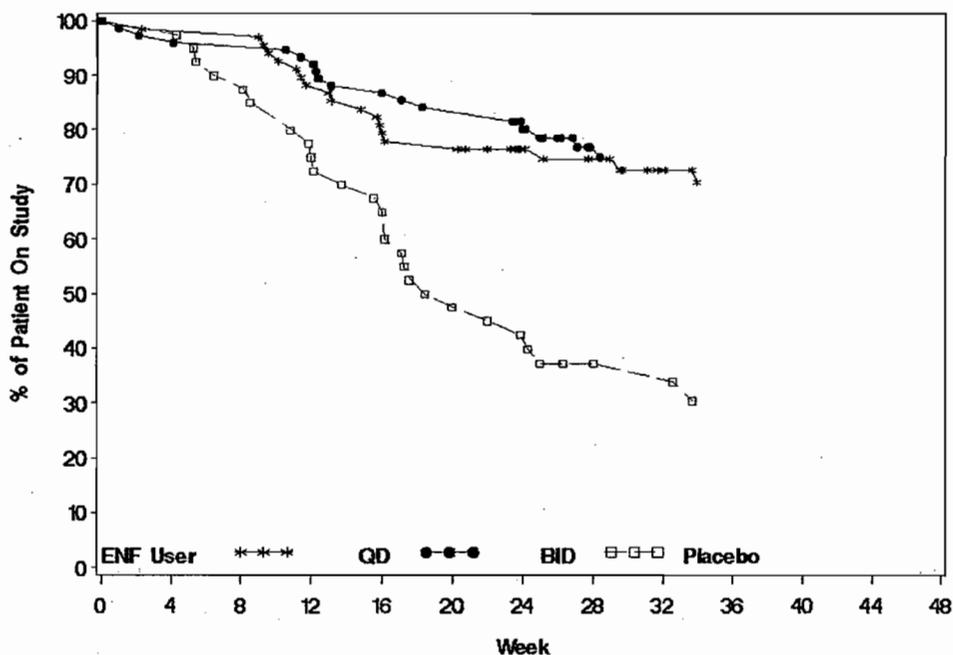


Figure 8. Study A4001028: Time to Discontinuation (Enfuvirtide Use in ART)

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

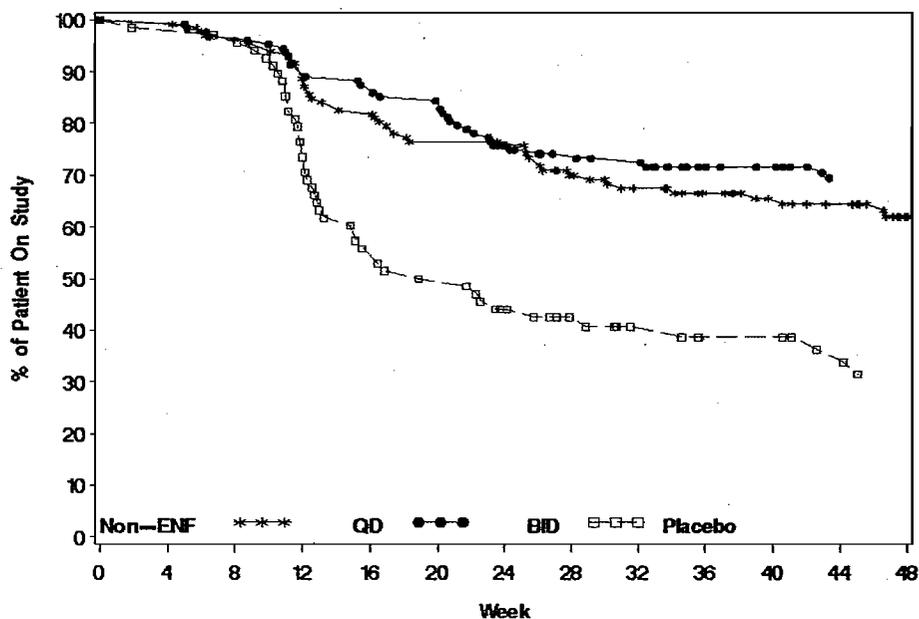


Figure 9. Study A4001027: Time to Discontinuation (Non-Enfuvirtide Use in ART)

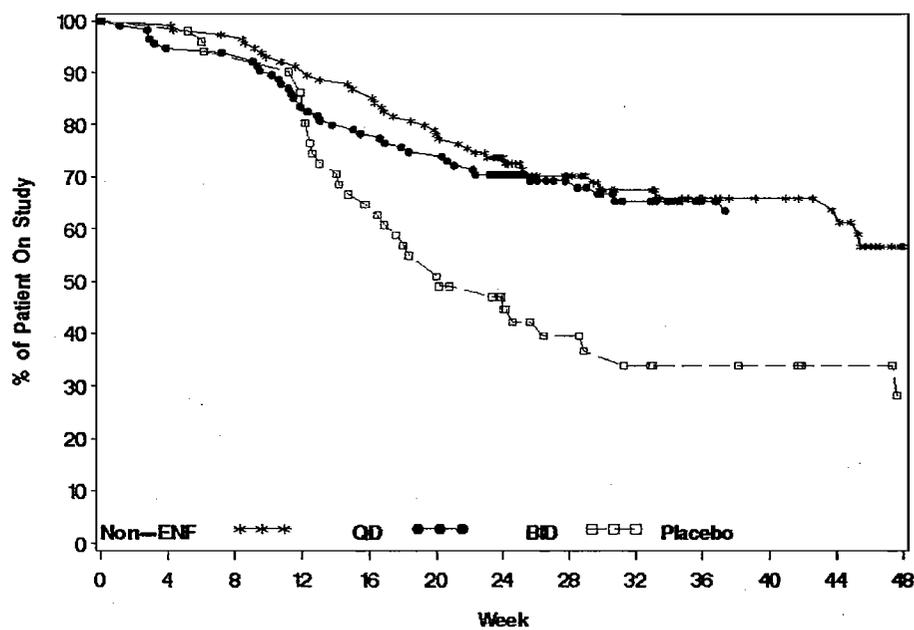


Figure 10. Study A4001028: Time to Discontinuation (Non-Enfuvirtide Use in ART)

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

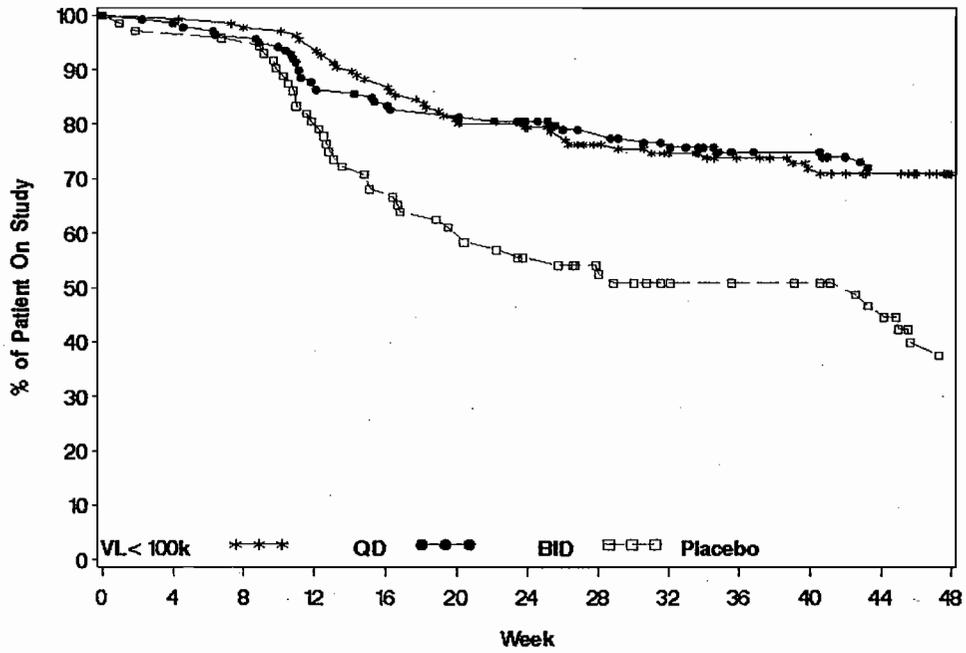


Figure 11. Study A4001027: Time to Discontinuation (VL < 100,000 copies/mL)

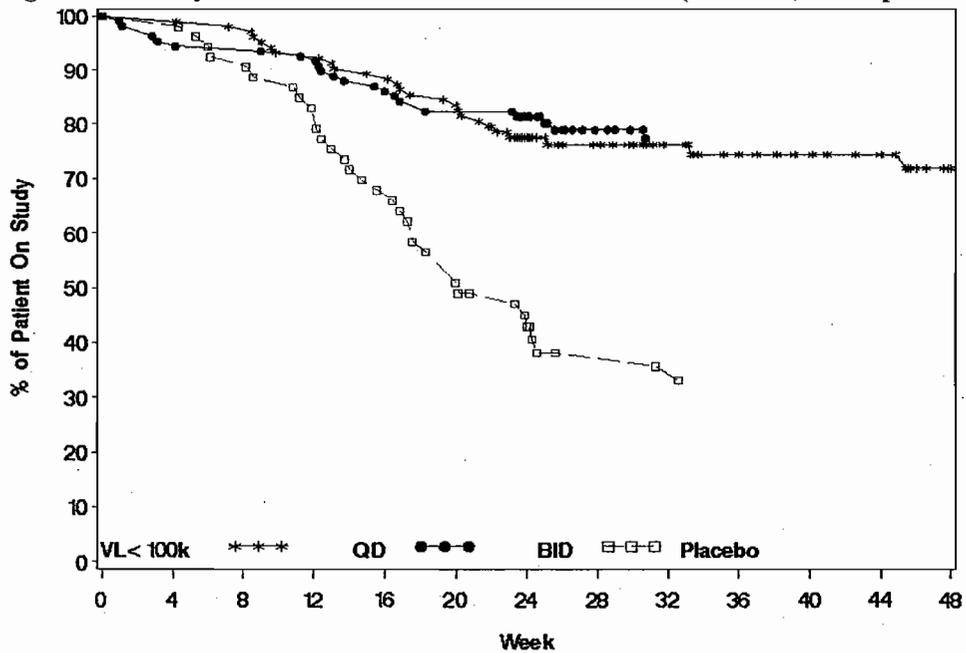


Figure 12. Study A4001028: Time to Discontinuation (VL < 100,000 copies/mL)

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

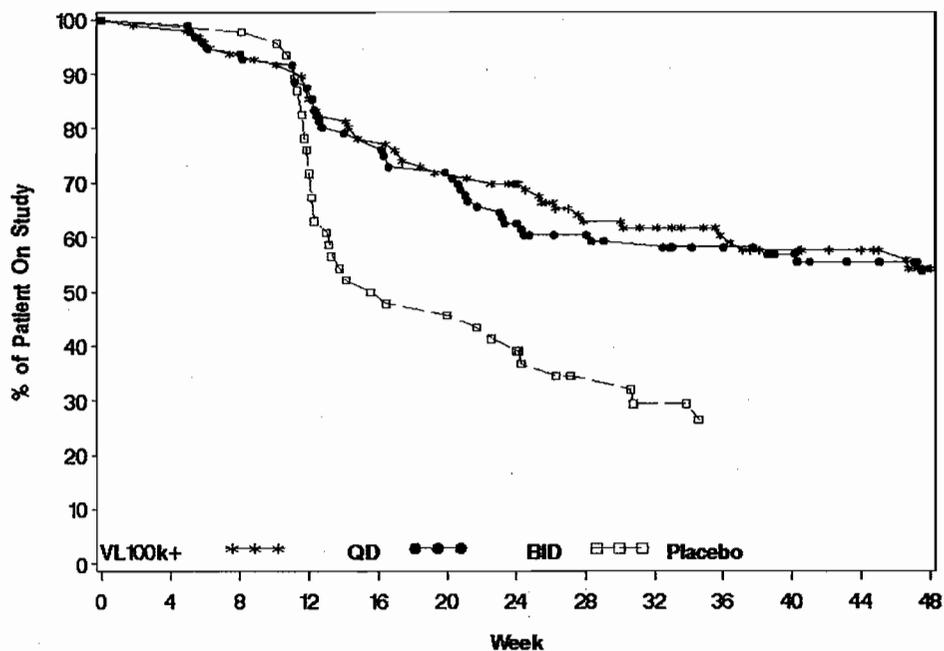


Figure 13. Study A4001027: Time to Discontinuation (VL \geq 100,000 copies/mL)

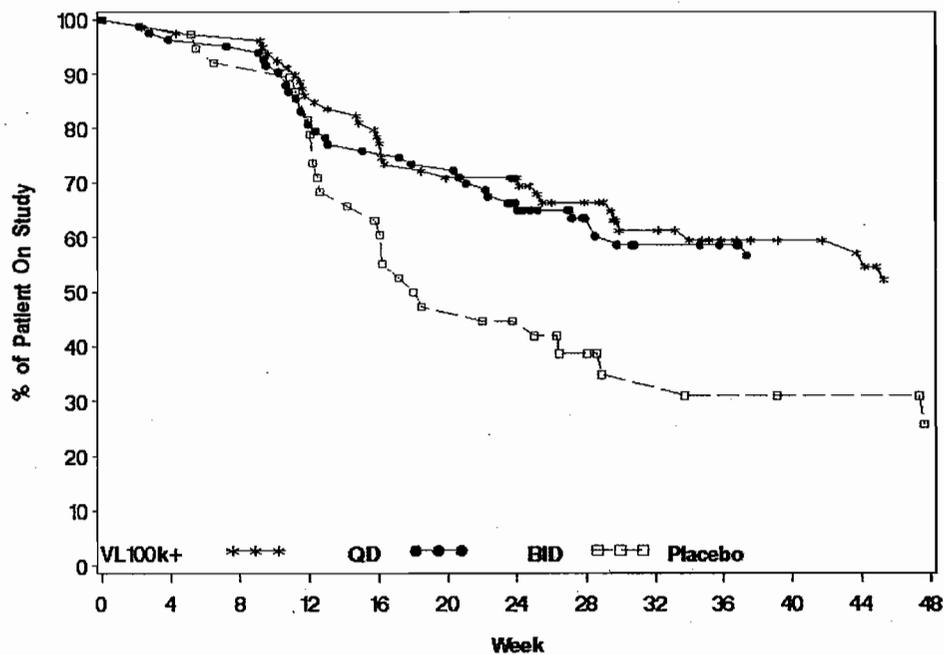


Figure 14. Study A4001028: Time to Discontinuation (VL \geq 100,000 copies/mL)

Time-Average Difference (TAD) in HIV-1 VL from Baseline to Week 24

In calculation of time-average difference (TAD) in VL from baseline to Week 24, the applicant used Day 196 (≤ 196) as a cut point to quantify a subject's discontinuation. If a subject discontinued prior to Day 197, the upper bound of Week 24, then this subject's TAD at Week 24 was imputed as zero. The applicant's results in Table 35 of 'Interim Full Clinical Study Report' (Page 100, Table 35 for A4001027, Page 101, Table 35 for A4001028) were based on FAS-As Treated population. The average value of VL at screening, randomization and Day 1 prior to treatment with study regimen was used as baseline VL.

This reviewer performed six analyses to verify the applicant's results for Studies A4001027 and A4001028, separately and combined. In addition, three types of sensitivity analyses were conducted using data pooling Studies A4001027 and A4001028 (Analyses 4-6. Different from the applicant's, this reviewer used type I error 0.001 to obtain the confidence intervals of the treatment difference (MVC-Placebo) in TAD at Week 24. Further more, all analyses were based on FAS - As Randomized population. The six TAD analyses are as follows;

- Analysis 1-for Study A4001027;
- Analysis 2-for Study A4001028;
- Analysis 3-for Studies A4001027 and A4001028 combined;
- Analysis 4-for Studies A4001027 and A4001028 combined and Day 1-VL as baseline;
- Analysis 5-for Studies A4001027 and A4001028 combined and using Day 168 as a cut point, and to impute zero to a subject's Week 24 TAD if this subject discontinued prior to Day 169; and
- Analysis 6-for Studies A4001027 and A4001028 combined and using Day 154 as a cut point, and to impute zero to a subject's Week 24 TAD if this subject discontinued prior to Day 155.

Please note that Analyses 1-4 used Day 196 as a cut point, and to impute zero to a subject's Week 24 TAD if this subject discontinued prior to Day 196, similar to the sponsor's.

Results are summarized in Table 7. We observed the following.

Results by Analyses 1-2 showed that the mean (se) for each treatment group, mean treatment differences (MVC QD-Placebo) in Week 24 TADs in the Studies A4001027 and A4001028 were similar to the applicant's 'Adjusted mean (se)', treatment difference (Maraviroc-Placebo).

- It is unclear whether the word 'Adjusted' in the applicant's report should be avoided. In summarizing the primary efficacy endpoint, the applicant used 'Adjusted mean' to quantify those by the ANCOVA models adjusting for the treatment and baseline randomization strata.

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

The Analysis 3 showed for data pooling A4001027 and A4001028, the mean treatment differences in TAD were -0.77 and -0.84 log₁₀ copies/mL, respectively for MVC QD – Placebo and MVC BID-Placebo.

The Analysis 4 showed that the estimated treatment differences in TADs using Day 1-VL as baseline were within 0.02 log₁₀ copies/mL from the corresponding results in Analysis 3 where the Average-VL was used as baseline.

The Analyses 5 and 6 showed robustness of treatment differences in TADs regardless of different cut points (Day 154, Day 168) used to define the discontinuation and imputation. The estimated treatment differences were within 0.03 log₁₀ copies/mL from those in Analysis 3 where a cut point of Day 196 was used to define discontinuation. This is because most of the subjects were discontinued earlier (<Week 24) due to insufficient virologic responses after Week 8.

The sensitivity analyses on time-average difference (TAD) in VL from baseline to Week 24 using data pooling the two studies showed robustness in estimated mean treatment differences, regardless of different baseline VL (Day 1 or average VL) used in calculation of TAD, different cut points (Day 154, Day 168, Day 196) used to define the discontinuation and imputation. The mean TAD were 0.76~0.77 and 0.80~0.84 log₁₀ copies/mL, respectively in subjects receiving MVC QD and MVC BID, compared with placebo. All the 99.9% CIs of the mean treatment differences in VL exclude zeros, indicating the supportive evidence of the superiority of MVC in VL reductions, compared to Placebo. MVC BID regimen appears to have slightly better results (<0.07 log₁₀ copies/mL) than the MVC QD regimen in VL reductions.

**Appears This Way
On Original**

Table 7. Time-Average Difference (TAD) from Baseline to Week 24 in Viral Load

Treatment Group	N	Mean (se)	Median	Range	Treatment Difference	
					Mean (se)	CI ¹
1. Study A4001027²						
MVC-QD	232	-1.64 (0.08)	-2.06	-3.86,0.00	-0.69 (0.13)	-0.95, -0.43
MVC-BID	235	-1.75 (0.08)	-2.22	-4.09,0.00	-0.80 (0.13)	-1.06, -0.54
Placebo	118	-0.95 (0.11)	0.00	-3.14,0.10		
2. Study A4001028²						
MVC-QD	182	-1.73 (0.09)	-2.16	-3.94,0.00	-0.87 (0.15)	-1.17,-0.58
MVC-BID	191	-1.75 (0.09)	-2.18	-3.52,0.00	-0.89 (0.15)	-1.18,-0.60
Placebo	91	-0.86 (0.12)	0.00	-3.60,0.00		
3. Studies A4001027 & A4001028²						
MVC-QD	414	-1.68 (0.06)	-2.10	-3.94,0.00	-0.77 (0.10)	-1.10,-0.44
MVC-BID	426	-1.75 (0.06)	-2.20	-4.09,0.00	-0.84 (0.10)	-1.16,-0.51
Placebo	209	-0.91 (0.08)	0.00	-3.60,0.10		
4. Studies A4001027 & A4001028 (Day 1-VL as Baseline)²						
MVC-QD	414	-1.69 (0.06)	-2.13	-3.98,0.00	-0.77 (0.10)	-1.10,-0.44
MVC-BID	426	-1.75 (0.06)	-2.17	-4.13,0.00	-0.82 (0.10)	-1.15,-0.49
Placebo	209	-0.92 (0.08)	0.00	-3.60,0.00		
5. Studies A4001027 & A4001028 (Imputing 0 if Discontinued prior to Day 169)						
MVC-QD	414	-1.75 (0.06)	-2.15	-3.94,0.00	-0.76 (0.10)	-1.08,-0.44
MVC-BID	426	-1.79 (0.06)	-2.21	-4.09,0.00	-0.80 (0.10)	-1.12,-0.48
Placebo	209	-0.99 (0.08)	-0.15	-3.60,0.10		
6. Studies A4001027 & A4001028 (Imputing 0 if Discontinued prior to Day 155)						
MVC-QD	414	-1.69 (0.06)	-2.13	-3.98,0.00	-0.77 (0.10)	-1.10,-0.44
MVC-BID	426	-1.75 (0.06)	-2.17	-4.13,0.00	-0.82 (0.10)	-1.15,-0.49
Placebo	209	-0.92 (0.08)	0.00	-3.60,0.10		

1. 95% confidence interval (CI) for an individual study and 99.9% CI for A4001027 and A4001028 combined.

2. Imputing 0 for those who discontinued from study prior to Day 196.

Appears This Way
 On Original

Changes from Baseline to Week 24 in CD4+ Cell Count

In CD4+ cell count data combining A4001027 and A4001028, 1031 (98.3%) subjects had both baseline and on study CD4 measurements. 16 subjects did not have on study CD4+ measurements but had baseline CD4+ measurements (n=16) and 2 subjects did not have baseline CD4+ but had on study CD4+ measurements. In addition, the longitudinal sample median lines by treatment regimen are displayed in Figures 15 and 16. It appears that the two MVC groups had CD4+ cell count increase as early as Week 2 and continued to Week 24. However, these curves are not adjusted for drop outs.

Similar to the method of baseline calculation in VL, there was a concern about the baseline CD4+ cell count. Figures 17 and 18 display the linear regression lines for the paired baseline CD4+ data - Average of CD4+ prior and at Day 1 (Average-CD4+) versus Day 1-CD4+ (Day 1-CD4+), for the Studies A4001027 and A4001028, respectively. In the text below, we refer to Average-CD4+ and Day 1-CD4+ as two types of baseline CD4+ cell count (cells/ μ L). At baseline the mean (median) CD4+ cell counts in A4001027 and A4001028 were 179 (152) cells/ μ L and 200 (166) cells/ μ L, respectively, using the Day 1 CD4+ data.

The comparison between Average-CD4+ and Day 1-CD4+ showed the following.

- 17.3% of the subjects had their differences between Average-CD4+ and Day 1-CD4+ greater than 30 cells/ μ L.
- The mean (median) time in CD4+ cell measurements between screening and Day 1 was -41 days with a range of (-106,-3) days.
- Figures 17 and 18 showed regression models and observed pair (Average-CD4+ versus Day 1-CD4+) by study. Similar patterns and the CD4+ cell count variability for the two studies were seen.

This reviewer conducted three sensitivity analyses on change from baseline to Week 24 in CD4+ cell count, using both DAY 1-CD4+ and Average-CD4+ as the baseline. Both mean change from baseline to Week 24, and median change from baseline to Week 24 using non-parametric Hodges-Lehman approach were applied. The estimated treatment differences in change from baseline to Week 24 in CD4+ cell count were sample summary statistics without adjusting for the two randomization strata.

Other methodology for these sensitivity analyses was as follows.

- 1) Week 24 completers (n=701), defined as subjects who had on study CD4+ at Week 24 time window;
- 2) LOCF without imputation for missing (n=1031);
- 3) LOCF for those who had at least one on study CD4+ and baseline CD4+ (n=1031), and 'no change' for those (n=18) who had missing in either baseline CD4+ or on study CD4+.

Tables 8 and 9 summarize the results.

1. Using Day 1-CD4+ cell count as baseline,

- The analyses of completers showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 38 and 25 cells/ μ L, respectively for MVC QD-Placebo and MVC BID-Placebo.
- The two analyses using LOCF showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 57~58, and 50~52 cells/ μ L, respectively for MVC QD -Placebo and MVC BID-Placebo.

2. Comparing the results using same approach but different baseline CD4+ cell count,

- The estimated mean treatment differences in CD4+ mean increases from baseline to Week 24 were slightly greater using Day 1-CD4+ as baseline than those using Average-CD4+ as baseline.

3. Comparing the mean and median treatment differences for each analysis, no particular directions were found, in most of the cases, the median treatment difference is more likely smaller than the mean treatment difference, with some exceptions. This feature is different from the sensitivity analyses for the VL regarding the primary efficacy endpoint, where the absolute value of the median treatment difference was more likely smaller than the absolute value of the mean treatment difference.

4. There were slightly difference between MVC QD and MVC BID regimen with regard to the CD4+ cell count changes. Using the mean method, the MVC QD appears to have slightly more increase in CD4+ cell count than those in the MVC BID group.

5. Except for the two analyses using completers, all the 99.5% CIs exclude zero, supporting the superiority of maraviroc, compare to placebo. However, the completer analyses showed a reduced benefit of treatment with MVC compare to Placebo.

6. Statistical method should be used to reduce the deviations and degree of violation of normality.

3.2 Evaluation of Safety

Safety analyses have been conducted by the medical officer Dr. Scott Proestel. Please refer to his review.

Appears This Way
On Original

Table 8. Mean Change from Baseline to Week 24 in CD4+ Cell Count (cells/ μ L)

	n	mean	se	Treatment Difference			
				mean	se	99.95% CI	
Analysis 1: Completers using Day 1-CD4 as Baseline (n=701)							
MVC-QD	294	137.1	7.6	38.4	12.9	-6.6	83.3
MVC-BID	310	123.2	6.0	24.5	12.1	-17.5	66.5
Placebo	97	98.7	10.4				
Analysis 2: LOCF using Day 1-CD4+ as Baseline (n=1031)							
MVC-QD	407	114.2	6.1	57.8	9.3	25.5	90.1
MVC-BID	418	107.9	5.1	51.5	8.7	21.3	81.6
Placebo	206	56.4	7.0				
Analysis 3: LOCF & Imputing 0 using Day 1-CD4+ as Baseline (n=1049)							
MVC-QD	414	112.3	6.0	56.7	9.2	24.8	88.6
MVC-BID	426	105.8	5.1	50.1	8.6	20.3	79.9
Placebo	209	55.7	6.9				
Analysis 4: Completers using Average-CD4+ as Baseline (n=701)							
MVC-QD	294	130.3	7.5	29.5	12.7	-14.8	73.8
MVC-BID	310	120.7	5.6	19.9	11.8	-21.3	61.1
Placebo	97	100.8	10.3				
Analysis 5: LOCF using Average-CD4+ as Baseline (n=1031)							
MVC-QD	407	106.9	5.9	52.1	9.0	20.8	83.4
MVC-BID	418	103.8	4.8	49.1	8.3	20.2	78.0
Placebo	206	55.8	6.6				
Analysis 6: LOCF & Imputing 0 using Average-CD4+ as Baseline (n=1049)							
MVC-QD	414	106.9	5.9	51.0	8.9	20.1	82.0
MVC-BID	426	103.8	4.8	47.9	8.2	19.3	76.5
Placebo	209	55.8	6.6				

Appears This Way
 On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 9. Median Change from Baseline to Week 24 in CD4+ Cell Count (cells/ μ L)¹

	n	Range ² Med ₁ -Med ₃	Treatment difference and 99.95% CI ³
1. Completers using Day 1 CD4+ as Baseline (n=701)			
Maraviroc QD (n=294)	117	(-789,883)	32 (-8,74)
Maraviroc BID (n=310)	106	(-737,673)	25 (-11,63)
Placebo (n=97)	78		
2. LOCF using Day 1 CD4+ as Baseline (n=1031)			
Maraviroc QD (n=407)	89	(-789,1220)	48 (23,75)
Maraviroc BID (n=418)	90	(-737,1010)	50 (26,76)
Placebo (n=206)	33		
3. LOCF & Imputing 0 using Day 1 CD4+ as Baseline (n=1049)			
Maraviroc QD (n=414)	86	(-789,1220)	46 (22,73)
Maraviroc BID (n=426)	89	(-737,1010)	49 (24,74)
Placebo (n=209)	31		
4. Completers using Average CD4+ as Baseline (n=701)			
Maraviroc QD (n=294)	115	(-716,900)	25 (-15,66)
Maraviroc BID (n=310)	105	(-730,627)	22 (-14,59)
Placebo (n=97)	80		
5. LOCF using Average CD4+ as Baseline (n=1031)			
Maraviroc QD (n=407)	86	(-716,1136)	44 (19,71)
Maraviroc BID (n=418)	88	(-730,862)	51 (25,75)
Placebo (n=206)	31		
6. LOCF & Imputing 0 using Average CD4+ as Baseline (n=1049)			
Maraviroc QD (n=414)	84	(-716,1136)	43 (18,69)
Maraviroc BID (n=426)	87	(-730,862)	49 (24,73)
Placebo (n=209)	30		

1. Hodges-Lehman's method.

2. i=1, Maraviroc QD, i=2, Maraviroc BID, 3-Placebo.

3. CI-confidence Interval.

Appears This Way
 On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Study 1027: CD4+ Cell Count

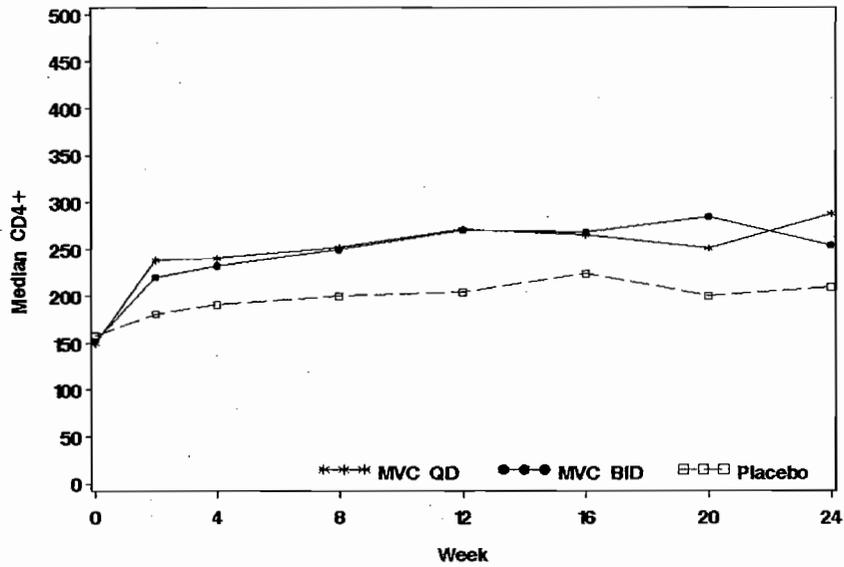


Figure 15. Study A4001027: Median CD4+ cells/ μ L

Study 1028: CD4+ Cell Count

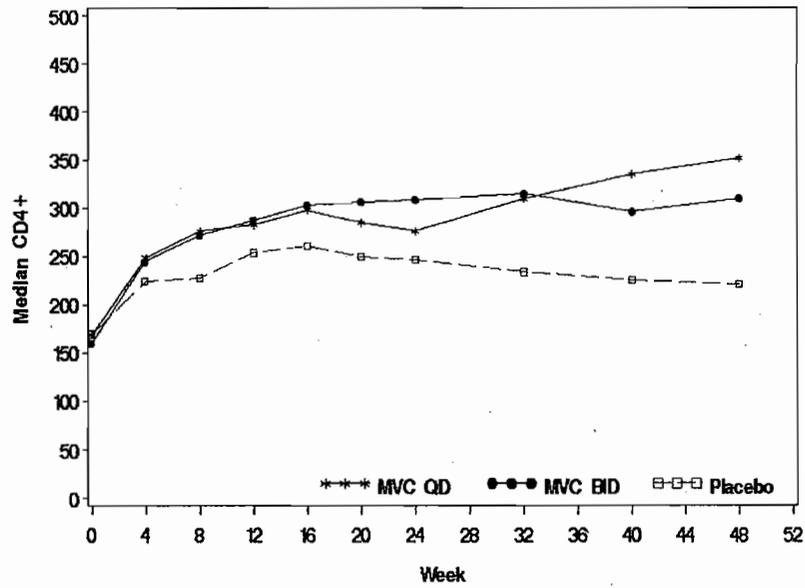


Figure 16. Study A4001028: Median CD4+ cells/ μ L

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

Study 1027: Baseline CD4+: Average v. Day 1 ($Y=9.826+0.958x$)

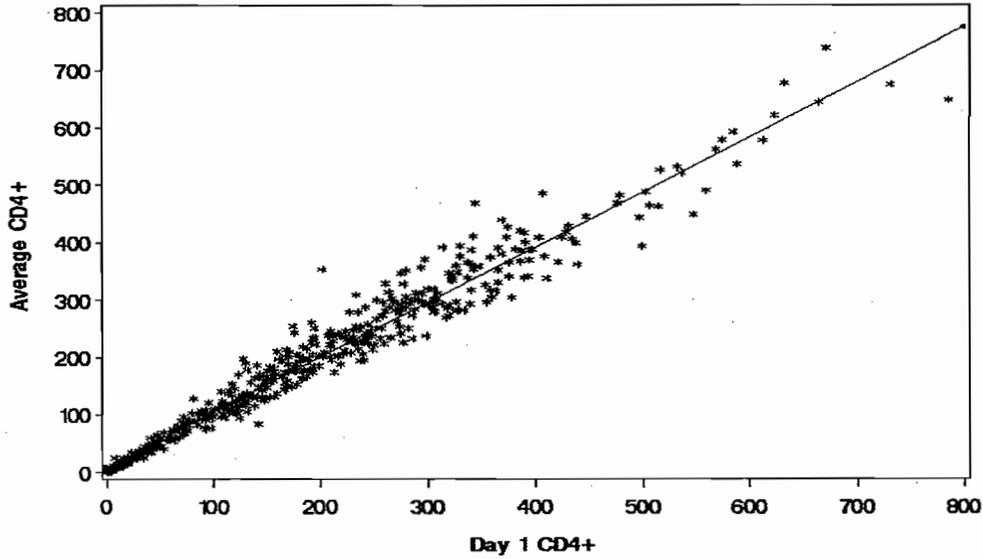


Figure 17. Study A4001027: Baseline CD4+ cells/ μ L (Average v. Day 1)

Study 1028: Baseline CD4+: Average v. Day 1 ($Y=9.438+0.974x$)

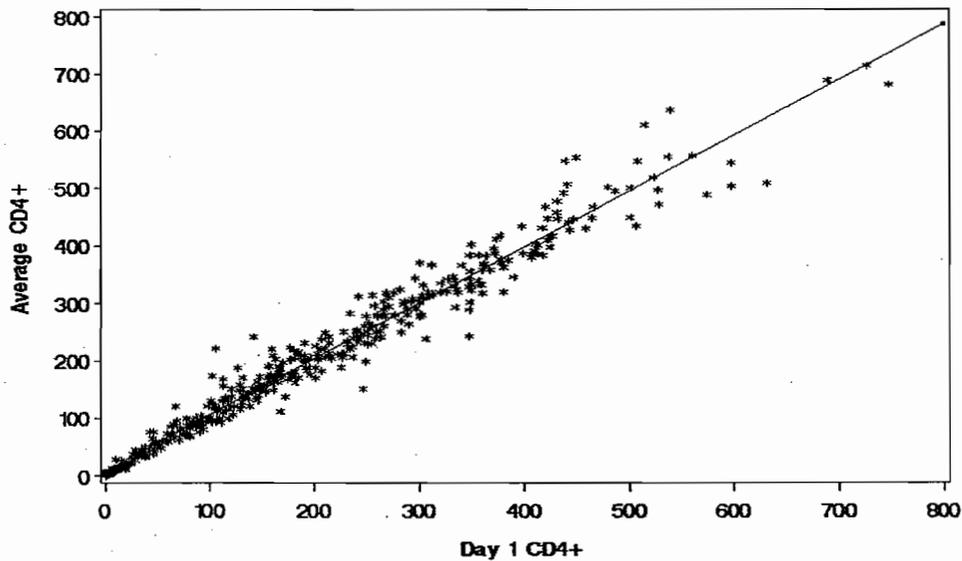


Figure 18. Study A4001028: Baseline CD4+ cells/ μ L (Average v. Day 1)

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

This reviewer conducted subgroup analyses of the primary endpoint using data pooling Studies A4001027 and A4001028. Section 4.1 summarized subgroup analyses on demographical subgroups of age (≤ 42 , 43-48, ≥ 49), gender, race (white, black) and region (USA, Non-USA). Section 4.2 summarized subgroup analyses on other selected baseline characteristics. These included the number of overall susceptibility score (OSS: 0-2, ≥ 3), previous enfuvirtide use in the ART (user and non-user), subjects' screening VL ($<$ or $\geq 100,000$ copies/mL) and baseline CD4+ cell count (< 100 , 100-200, > 200 cells/mm³). In addition, patients' phenotypic susceptibility score (PSS) and genotypic susceptibility score (GSS) have also been evaluated since it has been often to use the GSS/PSS to quantify drug resistance for HIV-1 drug trials. In Section 4.3, this reviewer further explored age and race effects and their interactions with maraviroc treatment using the ANCOVA models, where the effects of other important baseline factors were controlled.

The principles of grouping a variable were two folds: 1. Grouping a variable such as age using the cut points of 33% and 67% percentiles in order to obtain similar sample size for each stratum. 2. Grouping the categorical variable such as baseline OSS into 0-2 and ≥ 3 so that within each stratum the level of virologic responses were homogeneous. We categorized baseline PSS or GSS using the same principle for grouping OSS.

The subgroup analyses on the primary efficacy endpoints were basically post hoc to explore possible factors associated with maraviroc treatment effect over placebo. One should be cautious to make conclusions, especially when the sample sizes for some stratum was small.

4.1 Gender, Race and Age

4.1.1 Applicant's Subgroup Analyses

The applicant conducted subgroup analyses and the results were summarized in *Section 7.4*. Subgroup Analyses of Viral Load and CD4+ Cell Count Change from Baseline to Week 24.

The findings were as follows. In Studies A4001027 and A4001028,

- There did not appear to be any clinically important differences in the mean change from baseline in viral load at Week 24 based on analysis by age, region, HIV exposure category and time from HIV diagnosis.
- There was also no difference by gender; however, the number of female subjects was small: there are 22, 23 and 11 female subjects in study A4001027 and 29, 20 and 12 female subjects in study A4001028, in maraviroc QD, maraviroc BID and placebo, respectively.

- Regarding subgroup analysis by race, the summary of change from baseline in viral load at Week 24 by race indicated that the black placebo subgroup had an unusually high mean change from baseline compared with the white placebo subgroup. This may be explained by the small number of subjects in the black placebo subgroup, (n=14 in A4001027, n=11 in A4001028), and also by the skewed nature of the data in the black placebo treatment group towards higher decreases in viral load. (Reviewer Comments: for FAS as randomized population, there were 15 black subjects in the cover.xpt file in Study A4001027).
- In A4001027, the mean change from baseline for the subgroup of black subjects receiving maraviroc QD and BID was -1.75 log₁₀ copies/mL and -1.56 log₁₀ copies/mL, respectively, compared with -1.61 log₁₀ copies/mL for placebo. The median change from baseline for black subjects receiving maraviroc QD and BID, respectively, was -2.05 copies/mL and -1.70 log₁₀ copies/mL compared with -1.14 log₁₀ copies/mL for placebo. For subjects receiving maraviroc QD the median viral load change from baseline to Week 24 for white subjects was -2.34 log₁₀ copies/mL compared with -2.05 log₁₀ copies/mL for black subjects. For subjects receiving maraviroc BID the median viral load change from baseline for white subjects was -2.66 log₁₀ copies/mL and -1.70 log₁₀ copies/mL for black subjects.
- In A4001028, the mean change from baseline for the subgroup of black subjects receiving maraviroc QD and BID was -1.86 log₁₀ copies/mL and -1.98 log₁₀ copies/mL, respectively, compared with -1.40 log₁₀ copies/mL for placebo. The median change from baseline for black subjects receiving maraviroc QD and BID, respectively, was -2.23 copies/mL and -2.32 log₁₀ copies/mL compared with -0.69 log₁₀ copies/mL for placebo. For subjects receiving maraviroc QD the median viral load change from baseline to Week 24 for white subjects was -2.55 log₁₀ copies/mL compared with -2.23 log₁₀ copies/mL for black subjects. For subjects receiving maraviroc BID the median viral load change from baseline for white subjects was -2.47 log₁₀ copies/mL and -2.32 log₁₀ copies/mL for black subjects.

4.1.2 Reviewer's Analyses

This reviewer conducted subgroup analyses on demographical subgroups of age (≤ 42 , 43-48, ≥ 49), gender, race (white, black) and region (USA, Non-USA). These three age subgroups (≤ 42 , 43-48, ≥ 49) were obtained using the 33% (age=42) and 67% (age=49) percentiles as cut points for the baseline age category, so that the sample sizes for subgroups were similar.

The results based on change from baseline to Week 24 in VL are summarized in Tables 10 and 11. There did not appear to be any clinically important gender or region differences in the mean

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

change from baseline to Week 24 in VL. As to race and age subgroups, the following results were observed.

We observed significant treatment differences between white and black subjects.

- The mean change from baseline to Week 24 was -1.9 and -2.0 log₁₀ copies/mL, respectively, for white subjects receiving maraviroc QD and BID, compared with -1.0 log₁₀ copies/mL for white subjects receiving placebo. Apparently, maraviroc showed at least a -0.9 log₁₀ copies/mL treatment benefit over placebo among white subjects.
- Black subjects receiving maraviroc showed almost no treatment benefit over placebo. The mean change from baseline to Week 24 was -1.5 and -1.4 log₁₀ copies/mL, respectively, in MVC QD and BID groups. This compared to -1.2 log₁₀ copies/mL for in placebo group, resulting the mean (median) treatment difference of 0.2-0.3 (0.0) log₁₀ copies/mL MVC over placebo.
- Figure 19 shows histograms of change from baseline to Week 24 in VL by treatment regimen among black subjects. Similarly to the distributions of the primary efficacy endpoint, the distributions were not symmetric, and were skewed to the left. The race differences observed using Hodges-Lehmen's median approach differed slightly with those using the mean approach. However, the conclusions remained the same.
- Please note that the sample size (n=26) in the black placebo subgroup was small. As the small number may produce relatively coarse estimates of the efficacy⁹, the observed treatment differences between white and black subjects remained unclear.

As to age, it appeared that age may be associated with MVC treatment effects.

- The older the age, the more reduction in Week 24 VL. Subgroup analyses showed those receiving MVC BID with age 49 or older had an extra mean reduction (0.5 log₁₀ copies/mL) in Week 24 VL than those with age 42 or younger, compared with placebo subjects, p=0.08 by the Wald test. No such findings were observed among subjects receiving MVC QD or placebo groups.
- Among three age subgroups, the two MVC regimens in the oldest age group (≥49 years old) and the MVC QD in the 43-48 age group showed the best treatment efficacy compared to placebo: mean (median) VL reductions of 1.0 (1.0), 1.5 (1.2) and 1.1 (1.0) log₁₀ copies/mL, respectively, in the MVC QD and BID regimens of the (≥49) age group, and MVC QD regimen in the 43-48 age group, compared to placebo. The youngest age group (≤42 years old) and the MVC BID regimen in the 43-48 age group showed less treatment benefits: mean (median) reductions 0.7~0.9 (0.4~0.5) log₁₀ copies/mL in these MVC regimens, compared with placebo.

Please note that the ratios of black versus white, male versus female, or among age subgroups may not be similar to those in HIV-1 infected population. For example, CDC¹⁰ reported that among all HIV-1 infected adults and adolescents in USA during 2004-2005, 26~27% were female subjects. However, the FAS population in the two Studies A4001027 and A4001028

consisted of 11.3% (n=119) female subjects. The number of female HIV-1 infected subjects should be increased in the future HIV-1 drug trial.

Black Subjects: Mean Change in HIV-1 VL from Baseline at Week 24

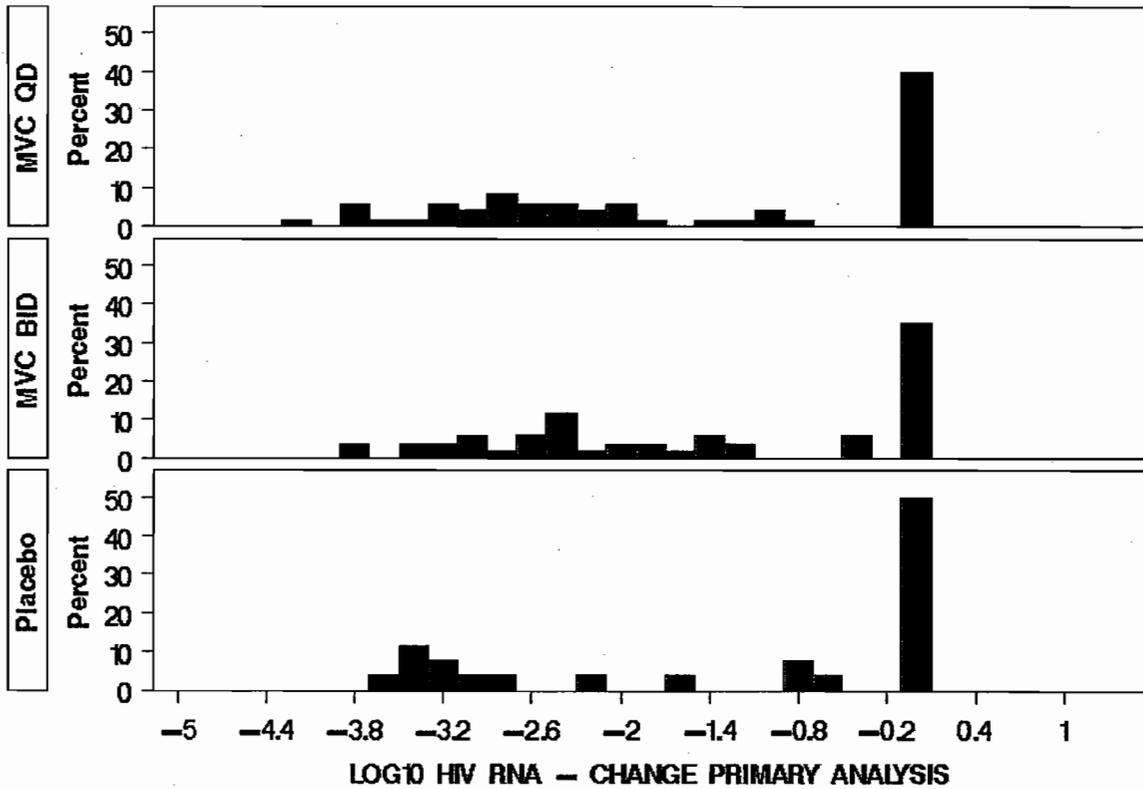


Figure 19. Histograms: Change from Baseline to Week 24 in VL among Black Subjects

Appears This Way
On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 10. Race, Gender, Age and Region: Median Change in Viral Load from Baseline to Week 24

		N	Sample Median	Treatment Difference (H-L) 99.95% CI		
1. White and Black						
White	MVC QD	336	-2.3	-0.8	-1.7	-0.1
	MVC BID	363	-2.5	-1.1	-2.0	-0.2
	Placebo	178	0.0			
Black	MVC QD	70	-1.9	0.0	-2.0	0.8
	MVC BID	51	-1.4	0.0	-1.8	0.8
	Placebo	26	-0.3			
2. Age						
≤42	MVC QD	156	-2.1	-0.4	-1.5	0.0
	MVC BID	147	-2.4	-0.4	-2.0	0.0
	Placebo	75	0.0			
43-48	MVC QD	127	-2.3	-1.0	-2.3	0.0
	MVC BID	120	-2.2	-0.5	-2.2	0.0
	Placebo	63	0.0			
≥49	MVC QD	131	-2.3	-1.0	-2.2	0.0
	MVC BID	159	-2.6	-1.5	-2.4	-0.3
	Placebo	71	0.0			
3. Gender						
Female	MVC QD	51	-2.4	-0.8	-2.9	0.0
	MVC BID	44	-2.5	-0.6	-2.9	0.0
	Placebo	24	0.0			
Male	MVC QD	363	-2.2	-0.7	-1.6	0.0
	MVC BID	382	-2.4	-1.0	-1.8	-0.1
	Placebo	185	0.0			
4. USA versus Other Countries						
USA	MVC QD	257	-2.2	-0.6	-1.7	0.0
	MVC BID	276	-2.3	-0.7	-1.8	0.0
	Placebo	135	0.0			
Non-USA	MVC QD	157	-2.4	-0.8	-2.1	0.0
	MVC BID	150	-2.6	-1.2	-2.4	0.0
	Placebo	74	0.0			

Appears This Way
 On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 11. Race, Gender, Age and Region: Mean Change in Viral Load from Baseline to Week 24

		n	Mean	Se	Treatment Difference			
					mean	Se	99.95% CI	
1. White and Black								
White	MVC QD	336	-1.9	0.1	-1.0	0.1	-1.4	-0.5
	MVC BID	363	-2.0	0.1	-1.1	0.1	-1.5	-0.6
	Placebo	178	-1.0	0.1				
Black	MVC QD	70	-1.5	0.2	-0.3	0.3	-1.5	0.9
	MVC BID	51	-1.4	0.2	-0.2	0.3	-1.4	1.0
	Placebo	26	-1.2	0.3				
2. Age								
≤42	MVC QD	156	-1.8	0.1	-0.7	0.2	-1.4	0.0
	MVC BID	147	-1.9	0.1	-0.7	0.2	-1.4	0.0
	Placebo	75	-1.1	0.2				
43-48	MVC QD	127	-1.9	0.1	-1.1	0.2	-1.8	-0.4
	MVC BID	120	-1.7	0.1	-0.9	0.2	-1.6	-0.2
	Placebo	63	-0.8	0.2				
≥49	MVC QD	131	-1.9	0.1	-1.0	0.2	-1.6	-0.3
	MVC BID	159	-2.2	0.1	-1.2	0.2	-1.9	-0.6
	Placebo	71	-1.0	0.2				
3. Gender								
Female	MVC QD	51	-2.1	0.2	-1.0	0.4	-2.2	0.3
	MVC BID	44	-2.0	0.2	-0.9	0.4	-2.2	0.3
	Placebo	24	-1.1	0.3				
Male	MVC QD	363	-1.8	0.1	-0.9	0.1	-1.3	-0.5
	MVC BID	382	-2.0	0.1	-1.0	0.1	-1.4	-0.6
	Placebo	185	-1.0	0.1				
4. USA versus Other Countries								
USA	MVC QD	257	-1.8	0.1	-0.8	0.1	-1.3	-0.3
	MVC BID	276	-1.8	0.1	-0.9	0.1	-1.4	-0.4
	Placebo	135	-0.9	0.1				
Non-USA	MVC QD	157	-2.0	0.1	-1.0	0.2	-1.7	-0.3
	MVC BID	150	-2.2	0.1	-1.1	0.2	-1.8	-0.5
	Placebo	74	-1.1	0.2				

Appears This Way
 On Original

By definitions, the overall susceptibility score (OSS), phenotypic susceptibility score (PSS) and genotypic susceptibility score (GSS) are **algorithms that assign a value to a subject's virus based on the outcome of the viral resistance testing performed at the screening visit and the antiretroviral agents selected by investigators to construct the OBT.** These scores were calculated by summing the in vitro susceptibilities of the **subject's virus to each antiretroviral agent within OBT.** Hence, the higher the susceptibility score, the more efficacious the OBT was likely to be.

The subgroup analyses on OSS, GSS and PSS showed that the results for GSS or PSS are somewhat similar to those by the subgroup of OSS. Subjects with baseline higher PSS score (≥ 3), higher GSS score (≥ 3) are likely to have reduced treatment efficacy, as seen among subjects with baseline OSS score ≥ 3 .

- The baseline OSS is highly associated with baseline PSS and GSS, $p < 0.0001$ by the Chi-square tests. The percentages of concordance are (OSS=PSS, 904/1033=85.7%) and (OSS=GSS, 643/1033=62.2%).
- A higher PSS score of ≥ 3 was also associated with a median difference of $-0.3 \log_{10}$ copies/mL in both MVC regimens, compare to Placebo. Likewise, a GSS ≥ 3 was also associated with a minimal treatment benefit. For example, the median treatment differences of -0.2 and $-0.1 \log_{10}$ copies/mL in MVC QD and MVC BID were obtained, compared with placebo. Details please see Tables 12 and 13.

4.3 Multivariate Analyses to Evaluate Age and Race Interactions with Treatment Effects

To further evaluate the age or race interactions with maraviroc treatment effects over placebo, we conducted analyses using two ANCOVA models to control for other baseline characteristics. By interactions we mean an interplay among predictors such as age or race in this case that produces an effect on the outcome (maraviroc treatment effect versus placebo in this case) that is different from the sum of the effects of the individual predictors¹².

- The ANCOVA Model 1 (n=1007) included treatment, baseline OSS (0-2, ≥ 3), previous enfuvirtide use in ART, screening HIV VL level ($<$ or $\geq 100,000$ copies/mL), baseline CD4+ cell count in square root transformation and race by treatment interactions as explanatory variables.
- The ANCOVA Model 2 (n=1031) included treatment, baseline OSS (0-2, ≥ 3), previous enfuvirtide use in ART, screening HIV VL level ($<$ or $\geq 100,000$ copies/mL), baseline CD4+ cell count in square root transformation and age by treatment interactions as explanatory variables.

The methodology in coding the variables was as follows.

- The treatment groups were dichotomous variables, defined as: QD (1, if subject was in MVC QD group, 0-otherwise), BID (1, if subject was in MVC BID group, 0-otherwise) and placebo (1, if subject was in placebo group, 0-otherwise).
- The two randomization indicators 'enfuse' and 'bahiv' were 0-1 dichotomous variables in the maraviroc databases to define previous enfuvirtide use in ART (0-not user, 1=user) and screening HIV VL level (1: $\geq 100,000$ copies/mL, 0: $< 100,000$ copies/mL).
- The name of baseline CD4+ cell count in square root transformation was 'sqcd4'.
- A 0-1 dichotomous variable 'osshigh' was defined as osshigh=1, if baseline OSS ≥ 3 and osshigh=0, if baseline OSS=0-2.
- Baseline age was a continuous variable, and the age by treatment interaction terms were age1 (age \times QD), age2 (age \times BID), age3 (age \times placebo).
- Race was defined as 0-1 variable: white (1 for white, and 0 for black). Race and treatment interactions were defined as white1 (white \times QD), white2 (white \times BID) and white3 (white \times placebo).

The SAS outputs for the two ANCOVA models are listed in Appendix 6.2. In the following, we summarize the observations based on the significant level of 0.05 by the Wald-t tests.

The Model 1 showed that the most significant terms were baseline OSS ($p < 0.0001$) and baseline CD4+ in square root transformation ($p < 0.0001$). This was followed by previous enfuvirtide use in ART ($p = 0.0003$), the interaction term of white \times BID (white2, $p = 0.0028$) and QD ($p = 0.012$). The race by (MVC BID- placebo) treatment interaction was estimated as -0.80 ($se = 0.24$) \log_{10} copies/mL, resulting $p = 0.018$ by the Wald-t test. Controlling for baseline CD4+ cell count, baseline OSS and the two randomization strata, white subjects in the two MVC groups had at least $0.9 \log_{10}$ copies/mL more reductions in VL compared to white subjects in the placebo group, but black subjects in the MVC groups had $0.3\text{--}0.4 \log_{10}$ copies/mL more reductions in VL compared to black subjects in the placebo group. These findings were consistent with the subgroup analyses on race (white v. black) via univariate analysis.

The Model 2 showed that the most significant terms were baseline OSS ($p < 0.0001$), baseline CD4+ in square root transformation ($p < 0.0001$) and previous enfuvirtide use in ART ($p < 0.0001$). This was followed by the interaction term of age \times BID (age2, $p = 0.027$). More VL reduction was associated with age, the older the age the subject in the MVC BID group, the more the reduction. No age by treatment effects were found in the MVC QD (age1) and placebo group (age3). The estimated age by MVC BID versus placebo interaction was -0.020 ($se = 0.014$) \log_{10} copies/mL, resulting $p = 0.16$, by the Wald-t test. Thus, controlling for the baseline characteristics in the model, age by MVC QD treatment effect versus placebo was no longer significant at $\alpha = 0.05$ level.

Table 12. Other Subgroup: Median Change in Viral Load from Baseline to Week 24

		N	median	Treatment Difference (H-L)		
				Median	99.95% CI	
1. Baseline OSS						
0-2	MVC QD	273	-2.0	-1.1	-2.0	0.0
	MVC BID	297	-2.3	-1.6	-2.3	-0.5
	Placebo	138	0.0			
≥3	MVC QD	134	-2.8	-0.2	-0.9	0.3
	MVC BID	125	-2.8	-0.2	-1.0	0.3
	Placebo	66	-2.6			
2. Previous Enfuvirtide Use in OBT						
Yes	MVC QD	168	-2.4	-1.0	-2.1	0.0
	MVC BID	183	-2.4	-1.1	-2.2	0.0
	Placebo	90	0.0			
No	MVC QD	246	-2.2	-0.6	-1.7	0.0
	MVC BID	243	-2.4	-0.8	-2.0	0.0
	Placebo	119	0.0			
3. Screening Viral Load (≥100,000 copies/mL)						
Yes	MVC QD	175	-2.0	-0.7	-2.0	0.0
	MVC BID	179	-2.6	-0.5	-2.4	0.0
	Placebo	84	0.0			
No	MVC QD	239	-2.3	-0.6	-1.7	0.0
	MVC BID	247	-2.4	-0.8	-1.9	-0.1
	Placebo	125	0.0			
4. CD4+ at Baseline (cells/mm³)						
<100	MVC QD	139	-0.6	0.0	-1.5	0.0
	MVC BID	142	-0.4	0.0	-1.8	0.0
	Placebo	62	0.0			
100-200	MVC QD	96	-2.3	-0.6	-2.2	0.0
	MVC BID	108	-2.6	-1.1	-2.4	0.0
	Placebo	56	0.0			
>200	MVC QD	178	-2.5	-1.2	-2.1	-0.3
	MVC BID	176	-2.6	-1.2	-2.2	-0.2
	Placebo	90	-0.3			
5. Baseline PSS						
0-2	MVC QD	254	-2.0	-1.2	-2.1	-0.1
	MVC BID	272	-2.3	-1.6	-2.3	-0.4
	Placebo	125	0.0			
≥3	MVC QD	154	-2.8	-0.3	-1.1	0.1
	MVC BID	150	-2.6	-0.3	-1.1	0.1
	Placebo	79	-2.2			
6. Baseline GSS						
0-2	MVC QD	300	-2.0	-1.0	-1.9	0.0

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

	MVC BID	320	-2.3	-1.5	-2.2	-0.4
	Placebo	145	0.0			
≥3	MVC QD	109	-2.8	-0.2	-1.0	0.3
	MVC BID	104	-2.7	-0.1	-1.0	0.4
	Placebo	59	-2.7			

Appears This Way
On Original

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

Table 13. Other Subgroup: Mean Change in Viral Load from Baseline to Week 24

		n	mean	se	Treatment Difference			
					mean	se	99.95% CI	
1. Baseline OSS								
0-2	MVC QD	273	-1.6	0.1	-1.1	0.1	-1.5	-0.7
	MVC BID	297	-1.8	0.1	-1.3	0.1	-1.7	-0.9
	Placebo	138	-0.5	0.1				
≥3	MVC QD	134	-2.3	0.1	-0.3	0.2	-1.0	0.4
	MVC BID	125	-2.3	0.1	-0.3	0.2	-1.0	0.5
	Placebo	66	-2.0	0.2				
2. Previous Enfuvirtide Use in OBT								
Yes	MVC QD	168	-2.0	0.1	-1.0	0.2	-1.6	-0.4
	MVC BID	183	-2.0	0.1	-1.1	0.2	-1.7	-0.5
	Placebo	90	-1.0	0.1				
No	MVC QD	246	-1.8	0.1	-0.8	0.2	-1.3	-0.3
	MVC BID	243	-1.9	0.1	-0.9	0.2	-1.4	-0.4
	Placebo	119	-1.0	0.1				
3. Screening Viral Load (≥100,000 copies/mL)								
Yes	MVC QD	175	-1.9	0.1	-1.0	0.2	-1.7	-0.4
	MVC BID	179	-1.9	0.1	-1.1	0.2	-1.8	-0.4
	Placebo	84	-0.8	0.1				
No	MVC QD	239	-1.9	0.1	-0.8	0.1	-1.3	-0.3
	MVC BID	247	-2.0	0.1	-0.9	0.1	-1.4	-0.4
	Placebo	125	-1.1	0.1				
4. CD4+ at Baseline								
	MVC QD	139	-1.4	0.1	-0.8	0.2	-1.5	-0.1
	MVC BID	142	-1.5	0.1	-0.9	0.2	-1.6	-0.2
	Placebo	62	-0.6	0.1				
	MVC QD	96	-1.9	0.1	-0.8	0.2	-1.6	0.0
	MVC BID	108	-2.2	0.1	-1.0	0.2	-1.8	-0.3
	Placebo	56	-1.1	0.2				
	MVC QD	178	-2.2	0.1	-1.1	0.2	-1.7	-0.5
	MVC BID	176	-2.2	0.1	-1.1	0.2	-1.7	-0.5
	Placebo	90	-1.1	0.1				
5. Baseline PSS								
0-2	MVC QD	254	-1.6	0.1	-1.2	0.1	-1.6	-0.7
	MVC BID	272	-1.8	0.1	-1.3	0.1	-1.7	-0.9
	Placebo	125	-0.5	0.1				
≥3	MVC QD	154	-2.3	0.1	-0.4	0.2	-1.1	0.3
	MVC BID	150	-2.3	0.1	-0.4	0.2	-1.1	0.3
	Placebo	79	-1.8	0.2				
6. Baseline GSS								
0-2	MVC QD	300	-1.7	0.1	-1.1	0.1	-1.5	-0.7

NDA 22-128, N000; Selzentry® (Maraviroc)
 Statistical Review and Evaluation

	MVC BID	320	-1.8	0.1	-1.3	0.1	-1.7	-0.9
	Placebo	145	-0.6	0.1				
≥3	MVC QD	109	-2.4	0.1	-0.4	0.2	-1.2	0.4
	MVC BID	104	-2.3	0.1	-0.3	0.2	-1.1	0.5
	Placebo	59	-2	0.2				

Appears This Way
 On Original

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Statistical Issues and Collective Evidence

During the review, several statistical issues in estimating the treatment effects maraviroc (MVC) versus Placebo group have been addressed.

5.1.1 Significant Heterogeneity in Discontinuation Status

In clinical trials, it is common to observe the heterogeneities in proportions of subjects who discontinued from the study, in the censorship violating the non-informative assumptions, and in different durations of treatment, between treatment groups. It is known that significant distributional differences in discontinuation between treatment groups may be associated with biased estimated treatment effects both in efficacy and safety.

Unfortunately, the discontinuation rates were much higher in the placebo treatment group than in the MVC treatment groups. As of the date of the Week 24 cut-off, for data combining Studies A4001027 and A4001028, 133 (63.6%) of the subjects in the placebo group discontinued from study, a significantly higher proportion than the MVC QD 143 (34.5%) and MVC BID 138 (32.4%) groups. The main reason was due to lack of clinical response or efficacy, found in 106 (50.7%) of the subjects in the placebo group, significantly higher than MVC QD 81 (19.6%) and MVC BID 91 (21.4%) groups. These calculations are based on sample sizes 209, 414 and 426, respectively, for the placebo, MVC QD and MVC BID groups.

Hence, time to discontinuation is much shorter in the placebo treatment group compared to the two MVC groups. Figures 4 and 5 show Kaplan-Meier (K-M) curves of time to discontinuation by A4001027 and A4001028, respectively. It appears that the K-M curves in the MVC groups are significantly separate from that in the placebo group, as early as Week 12.

As a result, there has been a great concern that the extreme unbalanced discontinuation rates between MVC and Placebo groups may result biased estimation in maraviroc efficacy.

5.1.2 Sensitivity Analyses of the Primary Efficacy Endpoint

To verify whether the estimated efficacy sizes by the sponsor are representative of the true effects of maraviroc, and to examining the potential effects of substantial discrepancies in the discontinuation status between placebo and maraviroc groups, this reviewer conducted three

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

types of sensitivity analyses on primary efficacy endpoint:

- Analysis using all available HIV-1 VL data regardless whether a subject was on study drug or not.
- Analysis on Week 24 completer.
- Imputation of missing by -0.4 to 0.3 log₁₀ copies/mL with an increment of 0.1 log₁₀ copies/mL.

The methodology used in the sensitivity analyses was as follows;

- The treatment difference on the primary endpoint was evaluated using the mean and median difference. The later is known as Hodges-Lehman approach, where the median of n_1 by n_2 pairs in treatment difference should be obtained.
- Type I error was adjusted when data were combined for A4001027 and A4001028 ($\alpha=0.001$), and also the two comparisons MVC QD or MVC BID versus Placebo ($\alpha =0.0005$). Hence, 99.95% confidence intervals of the mean and median treatment differences were estimated for the primary efficacy endpoint.
- The mean or median treatment differences are sample summary statistics without adjusting for their randomization strata.
- All the sensitivity analyses used Day 1 HIV-1 VL as baseline. If there are more than one value in the Week 24 time window (from Week 22 to Week 28), the one closest to Day 168 was selected.

The results of these sensitivity analyses are summarized in Tables 4 and 5.

The Sensitivity Analysis 1 used all available HIV-1 VL data regardless whether a subject was on study drug or not, meaning that those VL data after the subjects in the placebo group had switched to MVC+ OBT were included in the analysis as if they had been treated with Placebo+OBT. We obtained an extra mean reduction of 0.5 log₁₀ copies/mL and median reductions of 0.45~0.46 log₁₀ copies/mL in the two MVC regimens compared to the Placebo. The 99.95% CIs were all excluding zeros, indicating the superiority of MVC compared to Placebo. The treatment difference between the two MVC regimens is within 0.01 log₁₀ copies/mL.

The Sensitivity Analysis 2 was conducted among Week 24 completers, defined as the *earliest date of two events* exceeded Day 155 since Day 1: (1) the date of discontinuation from study and (2) the last date of treatment with study regimen. HIV-1 VL data were extended to one week from the date of discontinuation or the date of stopping treatment with study regimen.

- Subjects in the MVC regimens showed extra mean reductions 0.48 and 0.54 log₁₀ copies/mL in the MVC-QD and MVC-BID regimen, compared to Placebo. The extra median reductions were 0.38 log₁₀ and 0.44 log₁₀ copies/mL respectively in the MVC-

QD and MVC-BID regimen, compared to Placebo. The superiority of MVC compared to Placebo was confirmed using mean comparisons. Results by the Hodges-Lehman approach support the superiority of MVC-BID ($p < 0.0005$), not MVC-QD ($p > 0.0005$), at the type I error of 0.0005 level.

The third sensitivity analysis contained eight imputations (Analysis 3-10) to impute the missing of Week 24 HIV-1 VL values from -0.4 to 0.3 \log_{10} copies/mL with an increment of 0.1 \log_{10} copies/mL.

- As the imputed value increasing from -0.4 to 0.3 \log_{10} copies/mL, the mean treatment difference in change from baseline to Week 24 ranging -0.79 to -0.98 \log_{10} copies/mL for (MVC QD-Placebo), and -0.88 to -1.07 \log_{10} copies/mL for (MVC BID-Placebo).
- The estimated median treatment differences were slightly less than the mean treatment differences. As the imputed value increasing from -0.4 to 0.3 \log_{10} copies/mL, the median treatment difference in change from baseline to Week 24 ranging -0.69 to -0.88 \log_{10} copies/mL for MVC QD-Placebo, -0.80 to -0.99 \log_{10} copies/mL for MVC BID-Placebo.
- More than 50% of the subjects in the placebo regimen had discontinued from study by Week 24. Hence, any imputation approach that involved imputing of a single value to the missing in the placebo regimen should make this imputation value a median of the sample after imputation. For example, after imputing missing with -0.3 \log_{10} copies/mL, the median change from baseline to Week 24 should be -0.3 \log_{10} copies/mL in placebo regimen.

We conclude that the superiority of MVC versus placebo in estimating primary efficacy endpoint is essentially well maintained by these sensitivity analyses, even though some of the sensitivity analyses were rather conservative and were designed *not favoring the maraviroc treatment*.

5.1.3 Baseline HIV-1 Viral Load: Average Value or Day 1 Value

In evaluating the primary efficacy endpoint, the sponsor used average of HIV-1 viral load (Average-VL) at screening, at randomization and Day 1 (Day 1-VL) prior to treatment with study drugs as baseline VL. It was noticed that there was a mean time window of 5.9 weeks (mean=41 days, range -86 to -5 days prior to Day 1) between screening and Day 1. In addition, the study population in the A4001027 and A4001028 had a mean of 14 years of HIV-1 infections, and most of the subjects were on stable ART for at least 4 weeks. There was a concern whether the baseline VL may be influenced by the ART prior to Day 1.

Figure 3 shows the histogram of difference between the Day 1-VL and Average-VL. Overall, the Day 1-VL appears below the Average-VL, $p=0.053$ by the signed rank test. The mean (median) difference (Day 1-VL - Average-VL) is -0.003 (-0.013) \log_{10} copies/mL, with a range of $(-1.418, 0.773)$ \log_{10} copies/mL. Percentages of subjects with such difference < -0.3 , -0.2 , and -0.1 \log_{10} copies/mL are 5.3%, 9.7% and 23.9% respectively.

This reviewer conducted sensitivity analysis in primary efficacy endpoint using the Day 1-VL as baseline. Change from baseline at Week 24 data in VL were fitted to ANCOVA models which include treatment regimens, screening HIV-1 VL strata ($<$ or $\geq 100,000$ copies/mL), and enfuvirtide use in the ART.

Table 4 summarizes the results using different baseline VL. The estimated mean reduction from baseline to Week 24 is about 0.14 \log_{10} copies/mL more using the Day 1-VL as baseline, compared to Average-VL as baseline. Only slight differences were found among three treatment groups. Hence, the estimated treatment differences MVC-Placebo using different baseline VL were similar. However, the lengths of the estimated 97.5% CI using Day 1-VL appear to be 0.15 \log_{10} copies/mL wider than those using the Average-VL as baseline.

- Please note that the 99.95% CIs for the two studies combined should be used for multiplicity concern. This reviewer used 97.5% CIs so that the results should be comparable between the two methods.

Using Day 1-VL as baseline should be more appropriate for endpoints such as change from baseline in VL and time average difference (TAD) for this ART experienced study populations. However, the treatment differences regarding change from baseline in VL appear to be robust due to the double-blind and randomized study design.

5.1.4 Multiplicity

Multiplicity is unique problem for trials such as the two maraviroc trials A4001027 and A4001028 with phase 2b/3 design. In the maraviroc label, the sponsor decided to use the twice-daily regimen (MVC BID) for maraviroc as recommended regimen. As a result, data in the MVC QD groups were excluded; only data in the MVC BID and placebo groups were used for summarizing the clinical studies and maraviroc safety and efficacy. All results were corresponding to data pooling the two maraviroc trials A4001027 and A4001028. The sponsor's presentations in the Advisory Committee Meeting (April 24, 2007) and maraviroc labeling, the sponsor kept the 97.5% CI or a type I error of 0.025 unchanged for the estimating of primary efficacy endpoint.

We insisted to adjust for the multiplicity and suggested using 99.95% CI or a type I error of 0.0005 for such entity. It appears to be correct to use 97.5% CI for the treatment differences in primary efficacy endpoint obtained by individual study A4001027 or A1004028 respectively. When pooling the two studies together, it is implied that one study has been conducted, then the type I error should be 0.0005 (two treatment arms versus one placebo arm) to account for multiple comparisons, or 0.001 (one treatment arm versus one placebo arm) with no need to adjust for multiple comparison. One suggestion to solve this problem was to change the confidence interval to p-value.

The sponsor claimed that the change of the protocol-defined level of significance in type I error from 0.025 to 0.0005 would cause confusions and hence was not acceptable for the medical communities. As a result, reporting 97.5% CI of the treatment difference in primary efficacy endpoint remained unchanged in maraviroc label.

In the future, the adjustment of multiplicity in a phase 2b/3 design should be discussed in the protocol and the SAP reviews to avoid the potential confusions.

5.1.5 Different Matrices in Measuring Efficacy

The assessment of treatment effect size maraviroc versus placebo by the mean approach could be greatly influenced by outliers in the data since the distributions of change from baseline to Week 24 in VL were all skewed thus not Normal distributed. The problem should be worse when the sample size is small. Different matrices in measuring treatment difference should be considered for the evaluations. Here we used Hodges-Lehman's median approach as an alternative. Similar findings and conclusions were obtained although the estimated median treatment differences MVC versus placebo were differed from those using the mean approach.

5.1.6 Time to Discontinuation Due to Competing Risks

By the study design, subjects who had AE, pregnancy, etc, and treatment failure, will be discontinued from study. The treatment failure is defined as:

1) An increase to at least 3x the baseline (mean of all 3 values before start of dosing; at screening, randomization and baseline (Day 1)) viral load at the Week 2 visit or thereafter (confirmed by a second measurement taken no more than 14 days after the first measurement);

NDA 22-128, N000; Selzentry® (Maraviroc)
Statistical Review and Evaluation

- 2) Viral load $<0.5 \log_{10}$ decrease from baseline on 2 consecutive measurements starting at Week 8 (second measurement taken no more than 14 days after the first measurement);
- 3) Viral load $<1.0 \log_{10}$ decrease from baseline on 2 consecutive measurements starting at Week 8 (second measurement taken no more than 14 days after the first measurement), in a subject who had previously achieved a $\geq 2.0 \log_{10}$ decrease from baseline; or
- 4) An increase in viral load to $\geq 5,000$ copies/mL on 2 consecutive measurements taken no more than 14 days apart, in subjects previously confirmed to have undetectable levels of <400 copies/mL on 2 consecutive visits.

As seen in the study reports and data, the main reason of the subjects discontinued from studies was related to the lack of efficacy. In the two studies combined, there were approximately 20% of the subjects in each of the MVC groups and 50% in the placebo group discontinued due to virologic failure.

Time to discontinuation due to different reasons should be analyzed appropriately, especially when substantial number of subjects would discontinued from the study due to the study design. This reviewer suggests using the competing risk models⁴⁻⁸ to conduct analyses for such discontinuation data in the presence of multiple reasons such as 1) adverse events; 2) virologic failure by the study design; and 3) other reasons such as patient withdrawn consent, lost to follow-up, etc. Statistical methodology should be discussed with the sponsor in protocol review as well as the review of statistical analysis plan (SAP).

5.1.7 DAVP's Suggestions in Requesting SAS Programs

The sponsor failed to submit SAS programs on time as the review team requested.

SAS programs are needed for reviewing a NDA submission in order to assure the quality and timing of review. We need the SAS program to understand how data sets were compiled during the early phase of study such as from patients' enrollment to randomization (inclusion and exclusion), and from randomization to study, etc., so that how ITT populations were obtained, and reasons for patients' inclusions or exclusions to the study can be verified. Although the sponsor provided definitions for each variable in the data sets, it would be a great help to show how the important variables such as primary and secondary endpoints were created from the original longitudinal data such as HIV-1 viral load or CD4+ cell count data, patients' baseline data, the on study contaminant medication, patients' discontinuation, etc.

As suggested by the DAVP, the statistical review team should request the SAS programs before a NDA filing meeting and the statistical reviewer should refuse to recommend filing if the sponsor fails to submit SAS programs along with the SAS *.xpt files.

5.2 Conclusions and Recommendations

Maraviroc, in combination with other antiretroviral agents, is intended to be used for treatment-experienced adult patients infected with CCR5-tropic HIV-1, ~~_____~~

The applicant demonstrated statistically significant differences in mean reductions from baseline to Week 24 in HIV-1 viral load, the primary efficacy endpoint, in both maraviroc (MVC)+optimized background therapy (OBT) regimens as compared to Placebo+OBT for Studies A4001027 and A4001028 respectively. Thus it was indicated that maraviroc QD and maraviroc BID added to OBT were superior to OBT alone in the primary efficacy endpoint. All the secondary efficacy endpoint results at Week 24 measuring the virologic and immunologic responses were consistent with the primary endpoint and support the superior efficacy of both maraviroc treatment groups over placebo.

This reviewer conducted several sensitivity analyses of the primary and key secondary efficacy endpoints. The methods used favored placebo rather than the MVC regimens. Superior efficacy of both maraviroc treatment groups over placebo was confirmed.

Treatment of CCR5-tropic HIV-1 infected, ART experienced adult subjects with HIV-1 >5000 copies/mL with maraviroc adding to OBT was more effective than treatment with OBT alone in reducing viral load. For the two studies combined, the mean changes from baseline to Week 24 in HIV-1 viral load among subjects treated with maraviroc QD and BID +OBT regimens (n=414 and 426) were -1.9 and -2.0 log₁₀ copies/mL as compared to -1.0 log₁₀ copies/mL in placebo+OBT treated patients (n=209). The mean treatment differences and their 99.95% confidence intervals between maraviroc regimens and placebo regimen were -0.9 (-1.2,-0.6) and -1.0 (-1.2,-0.7) log₁₀ copies/mL respectively.

Subgroup analyses of the primary efficacy endpoint on selected baseline characteristics were conducted. This included the number of overall susceptibility score (OSS: 0-2, ≥3), the number of overall phenotypic score (PSS: 0-2, ≥3), the number of genotypic susceptibility score (GSS: 0-2, ≥3), previous enfuvirtide use in the ART (user and non-user), subjects' screening VL (< or ≥100,000 copies/mL) and baseline CD4+ cell count (<100, 100-200, >200 cells/mm³).

- The criteria for grouping the screening OSS, PSS and GSS were based on homogeneity of the outcome within a subgroup and sample sizes after grouping. For example, grouping screening OSS into 0-2 and ≥3 were based on initial subgroup analyses of OSS=0,1,2 and ≥3 on primary efficacy endpoint. The results showed that the two subgroups of OSS 0-2 and ≥3 provided two levels of virologic responses without significant differences within a subgroup after grouping.

Subgroups of baseline OSS, GSS and PSS showed significant treatment differences (MVC versus placebo) regarding the primary efficacy endpoint, while subgroups of previous enfuvirtide use in the ART, subjects' screening VL and baseline CD4+ cell count, respectively, did not appear to be any clinically important treatment differences. Subjects with baseline OSS ≥ 3 had significantly increased mean reductions of Week 24 VL from baseline, 2.3 \log_{10} copies/mL in MVC+OBT groups and 2.0 \log_{10} copies/mL in placebo+OBT group. This resulted almost no meaningful treatment benefit (-0.3 \log_{10} copies/mL) maraviroc over placebo. Conversely, subjects receiving maraviroc and with baseline OSS 0-2 had at least -1.1 \log_{10} copies/mL treatment benefit in mean change from baseline in viral load at Week 24. The results of subgroup analyses based on screening PSS (0-2, ≥ 3) and GSS (0-2, ≥ 3) appeared to be similar to those of the OSS subgroup analyses.

Subgroup analyses of the primary efficacy endpoint on selected demographic characteristics were conducted, including race (white versus black), age (≤ 42 , 43-48, ≥ 49), gender and region (USA versus non-USA). These three age subgroups (≤ 42 , 43-48, ≥ 49) were obtained using the 33% (age=42) and 67% (age=49) percentiles as cut points for the baseline age category, so that the sample sizes among age strata were similar.

There did not appear to be any clinically important gender or region differences in the mean change from baseline to Week 24 in VL.

As to race (black and white), the following results were observed.

- Using an ANCOVA model with treatment groups, baseline CD4+ cell count in square root transformation, baseline VL level, previous enfuvirtide use in ART, baseline OSS (0-2, ≥ 3), black or white and the interactions to treatment as explanatory variables (See Section 4.2), race (black v. white) was significantly associated with mean reduction from baseline to Week 24 in VL among subjects receiving MVC BID ($p < 0.01$), MVC QD ($p = 0.06$), but not in placebo ($p = 0.43$). The race by (MVC BID-placebo) treatment interaction was estimated as -0.80 (se=0.24) \log_{10} copies/mL, resulting $p = 0.018$ by the Wald-t test.
- The ANCOVA model also indicated that black subjects receiving maraviroc had only 0.3~0.4 \log_{10} copies/mL more reductions in Week 24 VL compared to placebo. However, white subjects in the MVC regimens showed at least -0.9 \log_{10} copies/mL treatment benefit over placebo.
- Subgroup analyses for white and black respectively showed similar results. However, the analyses were based on univariate analyses where other baseline factors such as baseline CD4+, screening VL, previous use of enfuvirtide, and baseline OSS were not controlled. Please note that the sample size ($n = 26$) in the black placebo subgroup is small. The small number will produce relatively coarse estimates of the efficacy.

Hence the observed treatment differences between white and black subjects remained unclear.

As to age effects, the following results were observed.

- The ANCOVA model with treatment groups, baseline CD4+ cell count in square root transformation, baseline VL level, enfuvirtide use in ART, baseline OSS, age and age by treatment interactions as explanatory variables (Section 4.3. ANCOVA Model 2) indicated that overall, age was significantly associated with mean reduction in Week 24 VL from baseline among subjects receiving MVC BID ($p=0.027$). The older the age, the more reduction in Week 24 VL. The estimated age by (MVC BID – placebo) interaction was -0.020 ($se=0.014$) \log_{10} copies/mL, resulting $p=0.16$, by the Wald-t test. Thus, controlling for the baseline characteristics in the model, age by MVC QD treatment effect versus placebo was no longer significant at $\alpha=0.05$ level.
- Subgroup analyses using three age subgroups (≤ 42 , 43-48, ≥ 49) showed those receiving MVC BID with age 49 or older had an extra mean reduction ($0.5 \log_{10}$ copies/mL) in Week 24 VL than those with age 42 or younger, compared with placebo subjects ($p=0.08$). No such findings were observed among subjects receiving MVC QD or placebo groups.

This reviewed conducted alternative analyses on selected key secondary efficacy endpoints including as time-average difference (TAD) in VL from baseline to Week 24 and change from baseline to Week 24 in CD4+ cell count using data pooling the two studies A4001027 and A4001028. In addition, time to discontinuation using the Kaplan-Meier approach was conducted respectively for A4001027 and A4001028 in the early phase of review.

- The sensitivity analyses on TAD in VL from baseline to Week 24 showed robustness in estimated mean treatment differences, regardless of different baseline VL (Day 1 or average VL) used in calculation of TAD, different cut points (Day 154, Day 168, Day 196) used to define the discontinuation and imputation. The mean TAD were $0.76\sim 0.77$ and $0.80\sim 0.84 \log_{10}$ copies/mL, respectively in subjects receiving MVC QD and MVC BID, compared with placebo. All the 99.9% CIs of the mean treatment differences in VL exclude zeros, indicating the supportive evidence of the superiority of MVC in VL reductions, compared to Placebo. MVC BID regimen appears to have slightly better results ($<0.07 \log_{10}$ copies/mL) than the MVC QD regimen in VL reductions.

The results of the alternative analyses of treatment difference in CD4+ cell count increase from baseline to Week 24 are as follows.

- When Day 1 CD4+ cell count was used as baseline, the two analyses using last observation carry forward (LOCF) with or without imputing zero to missing baseline

(or on study) showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 57~58, and 50~52 cells/ μ L, respectively for MVC QD and MVC BID groups, compare with placebo. The estimated mean treatment differences results using average CD4+ prior to treatment with study drugs as baseline were 2~6 cells/ μ L lower. All the 99.9% CIs excluded zero, supporting the superiority of maraviroc, compare with placebo in CD4+ cell count increase.

- The Week 24 completers (n=701) were defined as subjects who had on study CD4+ at Week 24 time window. Analyses of the Week 24 completers using Day 1 CD4 as baseline showed that the mean treatment differences in CD4+ increase from baseline to Week 24 were 38 and 25 cells/ μ L, respectively for MVC QD and MVC BID, compare with placebo. The results using average CD4+ prior to treatment with study drugs as baseline were 5~8 cells/ μ L lower. All the 99.9% CIs included zero, indicating a reduced significant level.
- Different from change from baseline to Week 24 in VL, 1) It appeared that subjects receiving MVC QD had slightly more increase (<13 cells/ μ L) in Week 24 CD4+ cell count than those who received MVC BID, compare with placebo; and 2) the treatment difference (MVC versus placebo) using median (Hodges-Lehman) may or may not be smaller than the mean for CD4+ cell count.

Using the Kaplan-Meier method (K-M), time to discontinuation was significantly longer in the maraviroc QD and BID regimens compared with the placebo regimen, $p < 0.0001$ by the log-rank test for A4001027 and A4001028, respectively. Subgroup analyses suggest that time to discontinuation may be associated with the previous use of enfuvirtide in ART but not associated with screening HIV-1 VL level at a significant level of $p = 0.20$. Different temporal patterns in different studies were observed. In A4001027 among those who used enfuvirtide in ART, the subjects receiving MVC BID were doing somewhat better than those receiving MVC QD. Conversely, in A4001028 among the non-enfuvirtide users in ART, the subjects in the MVC QD group were doing somewhat better than the MVC BID group. However, these qualitative interactions were based on the univariate analyses (K-M) and at the significance level of 0.2.

The strong maraviroc efficacy and safety outcomes observed in Studies A4001027 and A4001028 was mainly contributed by male and Caucasian, since majority of the population was male and Caucasian. The maraviroc effect in treatment of other demographical groups such as African American, or female subjects should be evaluated in the future¹¹.

Appears This Way
On Original

6. APPENDICES

6.1 References

1. Myles Hollander and Douglas A. Wolfe (1999). Chapter 4. The Two-Sample Location Problem, Nonparametric Statistical Methods. 106-140. New York: John Wiley & Sons, Second Edition.
2. Chris Decker. Calculating a Nonparametric Estimate and Confidence Interval using SAS Software.
3. John Klein and Melvin Moeschberger. Survival Analysis, Springer, 1997.
4. Lawless, J.F., *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York, Second Edition, 2002.
5. Prentice, R.L and Kalbfleisch, J.D. *The Analysis of Failure Times in the Presence of Competing Risks. Biometrics* 34, 541-554, December, 1978.
6. Lunn, M. and McNeil, D. *Applying Cox Regression to Competing Risks. Biometrics* 51, June 1995.
7. Gooley, T.A., Leisenring, W., Crowley, J. and Storer, B.E. *Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat. in Medicine*, 18, 1999.
8. Tai, B.C., Machin, D., White, I and GebSKI, V., *Competing risks analysis of patients with osteosarcoma: a comparison of four different approaches. Stat. in Medicine*, 20, 2001.
9. Moye, A.A. Multiple Analyses in Clinical Trials: Fundamentals for Investigators. Springer-Verlag, 2003.
10. CDC. HIV/AIDS Surveillance Report, 2005. Vol. 17. Atlanta: US department of Health and Human Services, CDC: 2006:1-46. Accessed March 1, 2007.
11. Daily HIV/AIDS Report: Politics and Policy: FDA Panel Recommends Agency Approve Pfizer's Antiretroviral Maraviroc. April 25, 2007.
12. Cohen J., Cohen P., West S.G. and Aiken L.S. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Third Edition. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey. 2003.

Appears This Way
On Original

2 Page(s) Withheld

2 Trade Secret / Confidential

 Draft Labeling

 Deliberative Process

SIGNATURES/DISTRIBUTION LIST (Optional)

The reviewer may officially identify the primary and concurring reviewer(s), as well as the Biometrics Deputy Division Director and Office Associate Director who will oversee adherence to template format. The distribution list provides reference to medical review division, medical officer, and project manager. Even though the review is stored in the Division File System (DFS), the reviewer may also include as an additional reference marker the final identifier of actual physical review file name. An example of this optional documentation is as follows:

Primary Statistical Reviewer: Susan Zhou, Ph.D.
Date: June 7, 2007

Concurring Reviewer(s): Greg Soon, Ph.D.

Statistical Team Leader: Greg Soon, Ph.D.

Biometrics Deputy Division Director: Daphne Lin, Ph.D.

cc:

HFD-Number/Project Manager
HFD-Number/Medical Officer
HFD-Number/Medical Team Leader
HFD-Number/Primary Statistical Reviewer
HFD-Number/Statistical Team Leader
HFD-Number/Biometrics Deputy Division Director
HFD-Number/Biometrics Division Director
HFD-Number/Office of Biostatistics Associate Director
HFD-Number/Office of Biostatistics Deputy Director

c:\NDA\statreview.doc

**Appears This Way
On Original**

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

/s/

Susan Zhou
6/20/2007 10:27:59 AM
BIOMETRICS

Greg Soon
6/20/2007 11:25:18 AM
BIOMETRICS

STATISTICAL REVIEW AND EVALUATION

NDA NUMBER	22128
SERIAL NUMBER	N000
SPONSOR	Pfizer Global Research & Development
NAME OF DRUG	Maraviroc (UK-427,857) 100 or 300 mg
INDICATION	Treatment of HIV-1 Infection
DOCUMENT REVIEWED	February 7, 2007
STATISTICAL REVIEWER	Susan Y. Zhou, Ph.D. (HFD-725)
STATISTICAL TEAM LEADER	Greg Soon, Ph.D. (HFD-725)
BIOMETRICS DIVISION DIRECTOR	Mohammad Huque, Ph.D. (HFD-725)
CLINICAL REVIEWER	Scott Proestel, M.D. (HFD-530)
PROJECT MANAGER	Kenny Shade, JD, BSN (HFD-530)
STATISTICAL KEY WORDS	SAS programs, Studies A4001027 and A4001028

1. Introduction

The review team had requested the submission of SAS programs via Facsimile dated 01 February 2007. The sponsor submitted two SAS programs 'analpop1027_28.sas', for Studies A4001027 and A4001028, 'analpop1029.sas' for A4001029. These SAS programs are SAS macros to create two SAS output files: '~.exclus' and '~.analpop', the former is a dataset to include randomized subjects who were excluded from the analysis datasets, and the latter is an analysis dataset.

This reviewer had verified the 'analpop1027_28.sas' using *.xpt files for Study A4001027 and has the following comments and requests.

2. Statistical Reviewer's Comments

2.1 Submitting Datasets not sent to FDA EDR

You failed to submit *.csv files used to compile the analpop.xpt and exclus.xpt files, and the output files exclus.xpt. Please submit all datasets including the *.csv files that were used to create *.xpt files such as analpop.xpt and exclus.xpt files, and the output files exclus.xpt.

There are 18 input datasets specified in the SAS Macro analpop. In addition, there are two datasets 'pidlst.csv' and 'pv.csv' that were repeatedly used but not specified in the SAS Macro parameters for SAS Macro analpop.

- The 'pidlst.csv' is a 'PREVIOUS PIDLIST' file (See analpop1027_28.sas).
- The 'pv.csv' is an "EXCLUS dataset to obtain additional PVs from Clinical Review" (See analpop1027_28.sas).
- Usually, a '*.csv' file is a Microsoft Excel file.

2.2 Submitting the revised SAS Programs Creating analpop.xpt and exclus.xpt Files

Please revise /debug the two SAS programs and submit to us ASAP. The current version could not be compiled due to miss specifications of input filename, missing in input datasets, miss specifications of variable name, miss use of SAS Functions, etc.. If a variable was coded differently between A4001027 and A4001028, you should consider using a Macro variable to specify it or to compile two SAS programs, one for A4001027, one for A4001028. The names for input files (macro variables) should be checked and to use the names submitted to the FDA EDR.

2.3 Submitting SAS Programs Creating Efficacy Data

Please send us SAS programs to create efficacy data such as hivr.xpt and vir*.xpt ASAP. The criterion and principles of such a submission should be the same as specified in comments 2.1-2.2.

2.4 Please submit the update *.xpt data for all studies.

2.5 Please submit SAS Programs for A4001026.

Susan Zhou, Ph.D.

Mathematical Statistician

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

/s/

Susan Zhou
2/8/2007 05:35:19 PM
BIOMETRICS

Greg Soon
2/12/2007 05:40:04 PM
BIOMETRICS

**MEMORANDUM OF FACSIMILE CORRESPONDENCE**

IND: 22-128

Drug: CELSENTRI® (maraviroc)

Date: February 1, 2007

To: Leilani Kapili

Sponsor: Pfizer Inc.

From: Kenny Shade, JD, BSN

Through: Scott Proestel, MD
Susan Zhou, PhD
Pravin Jadhav, PhD

Concur: Katherine Laessig, MD
Guoxing Soon, PhD
Jogarao Gobburu, PhD

Subject: Review Team Comments

The following comments are being conveyed to you on behalf of the review team. Please refer to your new drug application (NDA) 22-128 for CELSENTRI® (maraviroc) submitted December 19, 2006.

Review Team Comments

In the pre-NDA meeting on November 28, 2006, the statistical reviewer requested the SAS XPT files be submitted to the FDA electronic document room (EDR) with SAS programs. To date we have not received any SAS programs. Please submit key SAS programs with appropriate instructions for Studies A4001027, A4001028 and A4001029 as soon as possible. In addition, we have the following additional questions and request:

1. It is unclear how the Full Analysis Sets (FAS) and Per Protocol (PP) datasets were created. Please submit the key SAS programs to summarize subjects' status in screening, inclusion/exclusion, randomization, and treatment, so that the indicators such as FAS, PP, etc., in datasets "anlpop.xpt" can be clearly understood.
 - Some datasets for Studies A4001027 and A4001028 had discrepancies in their variables. For example, the variable "Randtst" is in "anlpop.xpt" under A4001027 but not in A4001028. Please comment.

March 14, 2006

2. Please submit SAS programs for creating the efficacy datasets such as vir27.xpt, vir28.xpt and vir29.xpt under the subdirectory of /datasets/virology.

- It appears that the number of randomized subjects in your report are different from the numbers of distinguished PID in vir** .xpt. Please clarify

We are providing this above information via telephone facsimile for your convenience. **THIS MATERIAL SHOULD BE VIEWED AS UNOFFICIAL CORRESPONDENCE.** Please feel free to contact me at 301-796-0807 if you have any questions regarding the contents of this transmission.

Kenny Shade, JD, BSN
Regulatory Project Manager
Division of Antiviral Drug Products
Center for Drug Evaluation and Research
Food and Drug Administration

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

/s/

Susan Zhou
2/5/2007 09:26:35 AM
BIOMETRICS



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

Statistical Review and Evaluation

CARCINOGENICITY STUDIES

IND/NDA Number: NDA 22-128

Drug Name: Maraviroc (UK-427,857)

Indication(s): 104 Week Carcinogenicity in Rats

Applicant: Pfizer
Testing Facility: Pfizer Global Research & Development,
333 Portage Street, Kalamazoo, MI

Documents Reviewed: Electronic submission, Dated December 20, 2006, and
data submitted electronically, received on December 20, 2006

Review Priority: Standard

Biometrics Division: Division of Biometrics -6

Statistical Reviewer: Mohammad Atiar Rahman, Ph.D.

Concurring Reviewer: Karl Lin, Ph.D.

Medical Division: Division of Anti Viral Products

Reviewing Pharmacologist: Pritam Verma, Ph.D.

Project Manager: Kenny Shade

Keywords: Carcinogenicity, Dose-Response

Table of Contents

1.....	Background	3
2.....	The Study	3
2.1. Sponsor's analyses.....		3
2.1.1. Survival analysis.....		3
2.1.2. Tumor data analysis.....		4
2.2. Reviewer's analyses.....		5
2.2.1. Survival analysis.....		5
2.2.2. Tumor data analysis.....		5
3.....	Summary	6
4.....	Appendix: Tables and Graphs	8
5.....	References:	18

1. Background

In this submission the sponsor included reports of an animal carcinogenicity study in rats. A short term carcinogenic study on transgenic mice for the same compound was submitted to the agency and was reviewed earlier in March, 2006. The present rat study is submitted as the complement of the required two studies, one in rat and one in mouse. This study was intended to assess the carcinogenic potential of maraviroc in rats when administered orally through gavage once daily at appropriate drug levels for about 104 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Verma.

2. The Study

Two separate experiments, one in males and one in females were conducted. In each of these two experiments there were four treated groups and one control group. Three hundred Sprague-Dawley rats of each sex were randomly allocated to treated and control groups in equal size of 60 animals. The dose levels for treated groups were 50, 100, 500, and 900 mg/kg/day. The control received vehicle [0.5% methylcellulose (w/v)/0.1% Tween 80 (w/w)] through gavage. Dosing was stopped in females after 96 weeks due to the high mortality in its control group. All surviving females from each group were necropsied.

Reviewer's comment: This reviewer discussed the issue early stopping of dosing with the reviewing pharmacologist. He mentioned that the sponsor discussed with the agency prior to stopping dosing and the agency agreed with this plan. The reviewing pharmacologist believes that this action should not have any significant effects on the outcomes of the study.

Animals were checked twice daily for mortality/viability and once daily for clinical signs. Mass tracking was performed once every two weeks following 6 months of treatment. Body weights were measured twice pretest; once weekly during the first 6 months and monthly, thereafter, during treatment. A complete histopathological examination was performed on all animals found dead, killed moribund, or sacrificed during or at the end of the experiment.

2.1. Sponsor's analyses

2.1.1. Survival analysis

Survival data were analyzed for the evidence of a dose-response relationship via sequential trend tests. That is, the trend test over all doses was performed first. If this test was found significant, the high dose (900 mg/kg/day) was removed and the trend test through the upper mid dose (500 mg/kg/day) was performed. If that test was significant, then the upper mid dose was removed and the trend test through the lower mid dose (100 mg/kg/day) was performed. If that test was significant, then the lower mid dose was removed and the trend test through the low dose (50 mg/kg/day) was performed. In keeping with the monotonicity assumption inherent in the trend tests, the p-value for any follow-up trend test was adjusted upward if necessary, so that it was never reported lower than the p-value for the trend through the next higher dose.

In all trend tests, ordinal doses were used as the independent variable (i.e., the weight used were 0, 1, 2, 3, and 4 for the control through the high-dose group, respectively). Time to event (death) was analyzed in days.

The two-tailed Tarone trend test¹ using all dose groups was performed at the 0.05 level of significance. If the initial test showed a significant trend, then one-tailed Tarone tests in the same direction of that trend were performed in a sequential manner at the 0.05 level. In addition, the survival percentage for each group was estimated using the product-limit method and Kaplan-Meier plots were produced. One 100 mg/kg/day dose

female animal and two 900 mg/kg/day dose female animals were removed from the study and treated as censored observations in these analyses.

Reviewer's comment: The sponsor did not explain the reason why these three animals were excluded from their analysis. However, the submitted data includes all animals (60 per group), and this reviewer's analysis (given in Section 2.2) includes all animals.

Sponsor's findings: Sponsor's analysis showed that there was a positive trend in female survival (increased survival) through the high dose group, $p=0.043$. No such statistically significant trend or differences in survival time were found in males.

2.1.2. Tumor data analysis

Tumor data were also analyzed for the evidence of a dose-response relationship via sequential trend tests, described in the section of survival analysis. The Peto method (Peto et al., 1980) was the primary method for analyzing the neoplastic lesions. For any lesion found in 12 or fewer animals (over all dose groups and regardless of cause of death status), the method was implemented through an exact trend test as described in Lin and Ali (Lin and Ali, 1994). Sequential dose-response tests were one-tailed for a positive relationship between tumor incidence and dose.

Tumors observed in a mortality-independent site (e.g., skin and mammary gland tumors) were analyzed using the onset-rate method. Any tumors of this type that were not detected prior to necropsy were assigned an onset time equal to time of death/sacrifice. For the analysis of these tumor types, the terminal sacrifice for each sex was regarded as just a single time point. Tumors observed at terminal sacrifice in sites that are not mortality-independent were considered incidental.

Fatal tumors were analyzed using the death-rate method. For the analysis of fatal tumors, actual death dates were used, even during the terminal sacrifice period, both for tumor-bearing animals and tumor-free animals. Animals with incidental tumors were regarded as censored observations in the corresponding fatal-tumor analysis.

Incidental tumors were analyzed using the prevalence method. Animals with fatal tumors were omitted from the corresponding incidental tumor analysis. In the analysis of these lesions, the following time intervals, expressed in days, were used: 1-365, 366-516, 517-591, 592-666, 667-674 (terminal sacrifice) for females, and 1-365, 366-545, 546-636, 637-727, 728-731 (terminal sacrifice) for males. Tumors of the duodenum, ileum, and jejunum were analyzed only at the "small intestine" level, and tumors of the colon, cecum, and rectum were analyzed only at the "large intestine" level. An intestine was considered missing only if all three of its components were missing.

Adjustment for multiple testing: Due to the large number of tests performed, a method of adjustment of p-values described in Mantel (1980) was used. Multiple testing adjustment method, outlined in the FDA guidance, was also applied.

Sponsor's findings: The following table shows sponsor's results for tumors with unadjusted p -value < 0.05 .

Appears This Way
On Original

Sponsor's Analyses Results

Organ and Lesion	Sex	Trend	MATPV
Liver Cholangiocarcinoma	M	0.028*	0.414
Parathyroid Adenoma	M	0.038*	0.558
Skin and Adnexa Basal Cell Neoplasm Benign	M	0.031*	0.457
Testis Leydig Cell Adenoma	M	0.044*	0.650
Thyroid Follicular Cell Adenoma	M	0.031*	0.457
Thyroid Follicular Cell Adenoma	F	0.038*	0.520

MATPV - Multiplicity-adjusted trend p-value based on Mantel
Source: Text table on Page 303 of 2639

The sponsor made the conclusion that all these p-values would be deemed significant based on applicable FDA guidance only if these tumor types were considered rare, however would not be deemed significant if they were considered common. No tumor type in either sex had either a trend through the 900 mg/kg/day group with $p < 0.01$ or a trend through the 500 mg/kg/day group with $p < 0.05$.

2.2. Reviewer's analyses

To verify sponsor's analyses and to perform additional analysis suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically.

2.2.1. Survival analysis

The survival distributions of animals in all five treated groups were estimated by the Kaplan-Meier product limit method. The homogeneity of survival distributions was tested using the Cox test (Cox, 1972) and the Generalized Wilcoxon test (Gehan, 1965). The intercurrent mortality data are given in Tables 1A and 1B for males and females, respectively. The Kaplan-Meier curves for survival rate are given in Figures 1A and 1B for males and females, respectively. Results of the tests are given in Tables 2A and 2B for males and females, respectively.

Reviewer's findings: The tests showed no statistically significant differences in survivals across treatment groups in either sex. Pairwise comparisons also showed no statistically significant difference between control and any of the treated groups.

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response using the methods described in the paper of Peto et al. (1980). Pairwise comparisons between treated groups and control were also performed using the age adjusted Fisher exact test. Since the sponsor classified the tumor types as 'cause of death' and 'not a cause of death', following Peto et al, this reviewer applied the 'death rate method' and the 'prevalence method' for these two categories of tumors respectively, to test the dose-response¹ relationship. For tumor types occurring in both categories a

¹ In this reviewer's analysis the phrase "Dose-response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor rate as dose increases.

combined test of 'death rate method' and the 'prevalence method' was performed. For the calculation of p-values, the Exact Permutation method was used. The actual dose levels of treatment groups were used as the weight for the dose response analysis. The time intervals used were 0 - 52, 53 - 78, 79 - 91, 92 - 103 weeks, and terminal sacrifice for males and 0 - 52, 53 - 78, 79 - 91, 92 - 94 weeks, and terminal sacrifice females. The tumor rates and the p-values of the tumor types tested for dose-response relationship are listed in Table 3A and 3B for males and females, respectively. The p-values for pairwise comparisons between the control and treated groups are given in Table 4A and 4B for males and females, respectively.

Multiple testing adjustment: Adjustment for the multiple dose response testing was done using the results of Lin and Rahman (1998), which recommends for a single study the use of a significance level $\alpha=0.05$ for rare tumors and $\alpha=0.01$ for common tumors, in order to keep the false-positive rate at the nominal level of approximately 10%. A rare tumor is defined as one in which the published spontaneous tumor rate is less than 1%. Adjustment for multiple pairwise comparisons was done using the results of Haseman (1983), which recommends the use of a significance level $\alpha=0.05$ for rare tumors and $\alpha=0.01$ for common tumors, in order to keep the false-positive rate at the nominal level of approximately 10%.

Reviewer's findings: The following tumor types showed dose response p-values less than or equal to 0.05.

Sex	Organ	Tumor	Cont	50mg	100mg	500mg	900mg	P-value
Male	Liver	M-CHOLANGIOCARCINOMA	0	0	0	0	2	0.0280
	Parathyroid(s)	B-ADENOMA	0	0	0	0	2	0.0350
	Skin and adnexa	B-BASAL CELL TUMOR, BENIG	0	0	0	0	2	0.0285
	Testis(es)	B-ADENOMA, LEYDIG CELL	0	0	0	0	2	0.0435
	Thyroid	B-ADENOMA, FOLLICULAR CEL	0	2	1	1	5	0.0221
Female	Thyroid	B-ADENOMA, FOLLICULAR CEL	0	1	0	2	3	0.0324

Based on the results of Lin and Rahman, the incidence of all the above listed tumors were considered to have statistically significant dose-response relationships. Also based on the results of Haseman, the pairwise comparison of high dose group and control for the incidence of follicular cell b-adenoma in thyroid in males was considered to be statistically significant ($p=0.0307$). No other dose response or pairwise comparisons on any tumor type was considered to be statistically significant.

3. Summary

In this submission the sponsor included reports of an animal carcinogenicity study in rats. A short term carcinogenic study on transgenic mice for the same compound was submitted to the agency and was reviewed earlier in March, 2006. The present rat study is submitted as the complement of the required two studies, one in rat and one in mouse. This study was intended to assess the carcinogenic potential of maraviroc in rats when administered orally through gavage once daily at appropriate drug levels for about 104 weeks. However, due to the high mortality in its vehicle control dosing of females was stopped after 96 weeks.

In this review, the phrase "dose-response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor rate as dose increases.

This study had four treated groups namely, 50, 100, 500, and 900 mg/kg/day dose levels along with one control

group. The tests showed no statistically significant differences in survivals across treatment groups in either sex. Tests showed statistically significant dose-response relationships in the incidence of liver/m-cholangiocarcinoma, parathyroid(s)/b-adenoma, skin and adnexa/benign b-basal cell tumor, testis(es)/ leydig cell b-adenoma, and thyroid/ follicular cell b-adenoma in males, and thyroid/ follicular cell b-adenoma, in females. The pairwise comparison of high dose group with control showed a statistically significant increase in the incidence of thyroid follicular cell b-adenoma in males.

M. Atiar Rahman, Ph.D.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, Biometrics-6

cc:

Archival NDA 22-128

Dr. Birnkrant

Dr. Proestel

Dr. Verma

Dr. Zheng

Dr. Naeger

Ms. Shade

Dr. Machado

Dr. Lin

Dr. Rahman

Dr. Zhou

Dr. ONeil

Ms. Patricianl

**Appears This Way
On Original**

4. Appendix: Tables and Graphs

**Table 1A: Intercurrent Mortality Rate
Male Rats**

Week	Control		50 mg/kg/day		100 mg/kg/day		500 mg/kg/day		900 mg/kg/day	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	4	6.9	4	6.7	5	8.3	2	3.3	3	5.0
53 - 78	7	18.3	13	28.3	14	31.7	8	16.7	12	25.0
79 - 91	15	43.3	12	48.3	8	45.0	7	28.3	10	41.7
92 - 104	11	61.7	11	66.7	8	58.3	14	51.7	15	66.7
Term. Sac.	23	38.3	20	33.3	25	41.7	29	48.3	20	33.3

**Table 1B: Intercurrent Mortality Rate
Female Rats**

Week	Control		50 mg/kg/day		100 mg/kg/day		500 mg/kg/day		900 mg/kg/day	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	2	3.3	3	5.0	6	10.0	2	3.3	6	10.0
53 - 78	17	31.7	13	26.7	12	30.0	10	20.0	7	21.7
79 - 91	14	55.0	16	53.3	14	53.3	14	43.3	13	43.3
92 - 103	4	61.7	1	55.0	3	58.3	4	50.0	4	50.0
Term. Sac.	23	38.3	27	45.0	25	41.7	30	50.0	30	50.0

**Table 2A: Intercurrent Mortality Comparison
Male Rats**

Method	Test	Statistic	P-value
Cox	Homogeneity	4.34	0.3617
Kruskal-Wallis	Homogeneity	4.58	0.3326

**Table 2B: Intercurrent Mortality Comparison
Female Rats**

Method	Time adjusted Trend test	Statistic	P-value
Cox	Homogeneity	2.83	0.5863
Kruskal-Wallis	Homogeneity	3.09	0.5435

Table 3A

**Tumor Rates and Dose Response p-values of Tested Tumors
Male Rats - Treated Over 104 Weeks**

Organ	Tumor	Cont	50mg	100mg	500mg	900mg	P-value
Abdomen	B-LIPOMA	0	0	1	0	0	0.5556
	M-CHONDROSARCOMA	0	1	0	0	0	0.8049
	M-LIPOSARCOMA	0	0	1	0	0	0.6173
Adrenal(s)	B-ADENOMA, CORTICAL	2	6	5	3	3	0.7616
	B-NEOPLASM, MEDULLARY, BE	11	3	3	4	5	0.7795
	M-CARCINOMA, CORTICAL	1	1	0	0	0	0.9633
Bone, unspec.	M-OSTEOSARCOMA	1	0	0	1	2	0.1032
Brain	B-ASTROCYTOMA, BENIGN	0	1	0	0	0	0.7115
	M-ASTROCYTOMA, MALIGNANT	0	1	0	1	0	0.5784
	M-GLIOMA, MIXED, MALIGNANT	1	0	1	1	0	0.6836
Eye(s)	M-SCHWANNOMA, INTRAOCUL,	0	1	0	0	0	0.8034
Heart	B-PARANGLIOMA	0	0	1	0	0	0.6083
	M-SCHWANNOMA, ENDOCARD, M	1	1	0	0	0	0.9573
Kidney(s)	B-ADENOMA	1	0	1	1	0	0.7241
	B-LIPOMA	0	1	0	0	0	0.8034
	B-TERATOMA	0	0	1	0	0	0.6325
	M-CARCINOMA	1	0	5	0	1	0.8117
Large Intestine	M-HISTIOCYTOMA, FIBROUS	1	0	0	0	0	1.0000
	M-LEIOMYOSARCOMA	1	0	0	0	0	1.0000
Liver	B-ADENOMA, HEPATOCELLULAR	0	1	0	0	0	0.8034
	M-CARCINOMA, HEPATOCELLUL	3	1	4	2	1	0.8395
	M-CHOLANGIOCARCINOMA	0	0	0	0	2	0.0280
Lymph node, uns.	B-LYMPHANGIOMA	0	0	0	1	0	0.4188
Lymphoreticular	M-LYMPHOMA, MALIGNANT	0	0	3	1	0	0.7106
	M-SARCOMA, HISTIOCYTTIC	3	2	2	0	0	0.9942
Mammary gland	B-FIBROADENOMA	1	0	1	1	0	0.6862
Mesenteric node	M-HEMANGIOSARCOMA	0	0	0	1	0	0.3333
Pancreas	B-ADENOMA, ISLET CELL	4	6	8	1	5	0.8365
	B-LIPOMA	0	0	0	1	0	0.4188
	M-ADENOCARCINOMA - MIXED	0	0	1	0	0	0.6271
	M-ADENOCARCINOMA, ACINAR	0	0	0	0	1	0.1923
	M-CARCINOMA, ISLET CELL	3	3	1	2	3	0.4562
Parathyroid(s)	B-ADENOMA	0	0	0	0	2	0.0350

Appears This Way
On Original

Table 3A (Continued)

**Tumor Rates and Dose Response p-values of Tested Tumors
Male Rats - Treated Over 104 Weeks**

Organ	Tumor	Cont	50mg	100mg	500mg	900mg	P-value
Pituitary	B-ADENOMA, PARS DISTALIS	41	39	42	41	42	0.2966
	B-ADENOMA, PARS INTERMEDI	4	0	0	0	0	1.0000
	M-CARCINOMA, PARS DISTA	2	2	2	1	2	0.5978
Prostate	M-ADENOCARCINOMA	0	0	1	0	0	0.6046
Salivary gland	M-ADENOCARCINOMA	0	1	0	0	0	0.7867
Seminal vesicle	B-ADENOMA	0	0	0	0	1	0.2222
	B-LEIOMYOMA	0	0	1	0	0	0.6271
Skin and adnexa	B-BASAL CELL TUMOR, BENIG	0	0	0	0	2	0.0285
	B-FIBROMA	1	1	1	0	1	0.6306
	B-HAIR FOLLICLE TUMOR-BEN	1	1	3	3	2	0.2887
	B-HISTIOCYTOMA, FIBROUS,	1	1	0	0	0	0.9428
	B-LIPOMA	0	1	1	1	0	0.6410
	B-PAPILLOMA, SQUAMOUS CEL	1	0	0	1	1	0.2714
	B-SEBACEOUS ADENOMA	1	0	0	0	1	0.4296
	M-FIBROSARCOMA	0	0	0	1	0	0.3704
	M-FIBROUS HISTIOCYTOMA	0	2	3	0	0	0.9193
	M-HEMANGIOPERICYTOMA, MAL	0	0	0	1	0	0.4915
	M-SCHWANNOMA, MALIGNANT	1	0	0	0	0	1.0000
	M-SQUAMOUS CELL CARCINOMA	0	1	0	0	0	0.8136
Small Intestine	B-ADENOMA	0	0	0	0	1	0.1709
	M-ADENOCARCINOMA	1	0	1	0	0	0.8746
Spleen	M-HEMANGIOSARCOMA	0	1	1	1	1	0.3789
Testis(es)	B-ADENOMA, LEYDIG CELL	0	0	0	0	2	0.0435
Thymus	B-THYMOMA, BENIGN	0	1	0	0	0	0.8070
Thyroid	B-ADENOMA, C-CELL	5	2	10	7	5	0.4976
	B-ADENOMA, FOLLICULAR CEL	0	2	1	1	5	0.0221
	M-CARCINOMA, C-CELL	1	2	0	0	0	0.9719
Tongue	M-SCHWANNOMA	0	0	1	0	0	0.6325
Urinary bladder	B-LEIOMYOMA	0	0	0	0	1	0.1923
	B-PAPILLOMA, TRANSITIONAL	0	1	0	0	0	0.8704

Appears This Way
On Original

Table 3B

Tumor Rates and Dose Response p-values of Tested Tumors
Female Rats - Treated Over 104 Weeks

Organ	Tumor	Cont	50mg	100mg	500mg	900mg	P-value
Adrenal(s)	B-ADENOMA, CORTICAL	3	1	3	4	2	0.4855
	B-NEOPLASM, MEDULLARY, BE	0	3	0	2	1	0.4346
	M-MEDULLARY CARCINOMA	0	1	1	0	0	0.8010
Brain	B-NEOPLASM, GRANULAR CE,	1	0	0	0	0	1.0000
	B-PAPILLOMA, CHOROID PLEX	1	0	0	0	0	1.0000
	M-ASTROCYTOMA, MALIGNANT	0	3	0	1	1	0.5841
	M-GLIOMA, MIXED, MALIGNAN	0	0	0	0	1	0.1186
	M-SARCOMA, MENINGEAL	1	0	0	0	0	1.0000
Cervix	B-POLYP, ENDOMETRIAL STRO	0	0	0	1	0	0.3824
	M-SCHWANNOMA	0	0	0	1	1	0.1288
Clitoral Gland	B-PAPILLOMA, SQUAMOUS CEL	0	1	0	0	0	0.7119
Heart	B-SCHWANNOMA, ENDOCARDIAL	0	0	0	1	0	0.3803
	M-MESOTHEL, ATRIOCAVAL, M	1	0	0	0	0	1.0000
Liver	B-ADENOMA, HEPATOCELLULAR	1	0	1	0	0	0.8557
	B-CHOLANGIOMA	0	0	0	1	0	0.4444
	M-CARCINOMA, HEPATOCELLUL	0	1	0	0	0	0.8296
Lymphoreticular	M-LYMPHOMA, MALIGNANT	1	1	0	1	2	0.2314
	M-SARCOMA, HISTIOCYTIC	0	0	2	0	0	0.7311
Mammary gland	B-ADENOMA	1	3	3	2	3	0.3975
	B-FIBROADENOMA	20	17	20	14	17	0.7899
	M-ADENOC IN FIBROADENOMA	2	5	6	3	4	0.4795
	M-ADENOCARCINOMA	27	31	19	25	17	0.9802
	M-NEOPLASM, MIXED, MALIGN	0	1	0	1	1	0.2352
Mesenteric node	B-HEMANGIOMA	0	0	0	1	0	0.4444
Ovary(ies)	B-CYSTADENOMA	0	0	1	0	0	0.6269
	B-LUTEOMA, BENIGN	0	0	1	0	0	0.6269
	B-NEOPL, GRANULOSA CELL,	1	1	1	1	0	0.8112
Pancreas	B-ADENOMA, ISLET CELL	1	1	0	5	0	0.4204
	M-CARCINOMA, ISLET CELL	1	0	2	2	2	0.2374
Parathyroid(s)	B-ADENOMA	0	1	1	1	0	0.6586
Pituitary	B-ADENOMA, PARS DISTALIS	50	50	44	48	44	0.9306
	B-ADENOMA, PARS INTERMEDI	0	0	0	2	0	0.1876
	M-CARCINOMA, PARS DISTA	4	5	5	6	6	0.4085
Skin and adnexa	B-BASAL CELL TUMOR, BENIG	0	1	0	0	0	0.8296

Appears This Way
On Original

Table 3B (Continued)

Tumor Rates and Dose Response p-values of Tested Tumors
Female Rats - Treated Over 104 Weeks

Organ	Tumor	Cont	50mg	100mg	500mg	900mg	P-value
Skin and adnexa	B-HAIR FOLLICLE TUMOR-BEN	0	1	0	0	0	0.8028
	B-LIPOMA	1	1	1	0	0	0.9165
	M-FIBROSARCOMA	0	0	2	0	0	0.5354
	M-FIBROUS HISTIOCYTOMA	0	1	0	0	0	0.8296
	M-HEMANGIOPERICYTOMA, MAL	1	0	0	0	0	1.0000
Small Intestine	B-LEIOMYOMA	1	0	0	0	0	1.0000
	M-ADENOCARCINOMA	0	0	1	0	0	0.6296
Thymus	M-THYMOMA, MALIGNANT	0	1	0	0	0	0.8026
Thyroid	B-ADENOMA, C-CELL	4	6	6	4	7	0.4016
	B-ADENOMA, FOLLICULAR CEL	0	1	0	2	3	0.0324
	M-CARCINOMA, C-CELL	1	0	0	0	0	1.0000
Turbinate	M-CARCINOMA, NOS	0	0	0	1	0	0.3965
Urinary bladder	B-LIPOMA	0	0	0	1	0	0.4478
	B-PAPILLOMA, TRANSITIONAL	0	0	0	0	1	0.2239
Uterus	B-LIPOMA	1	0	0	0	0	1.0000
	B-NEOPL, GRANULAR CELL, B	0	1	0	0	0	0.8296
	B-POLYP, ENDOMETRIAL STRO	3	2	3	3	5	0.1556
	M-ADENOCARCINOMA	0	0	0	1	1	0.1408
	M-FIBROUS HISTIOCYTOMA	1	0	0	0	0	1.0000
	M-SARCOMA, ENDOMETRIAL ST	0	0	0	1	0	0.3803
	M-SCHWANNOMA	0	1	0	0	0	0.8031
Vagina	B-PAPILLOMA, SQUAMOUS CEL	0	0	1	0	0	0.6343
	B-POLYP, STROMAL	0	0	1	0	0	0.4915
Zymbal's gland	M-CARCINOMA, SEBACEOUS CE	0	0	1	0	0	0.4915

Appears This Way
On Original

Table 4A

Pairwise Comparisons of Treated Groups with Control
Male Rat - Fed Over 104 Weeks

Organ	Tumor	Cont. vs 50mg	Cont. vs 100mg	Cont. vs 500mg	Cont. vs 900mg
Abdomen	B-LIPOMA	.	0.5556	.	.
	M-CHONDROSARCOMA	0.4667	.	.	.
	M-LIPOSARCOMA	.	0.4805	.	.
Adrenal(s)	B-ADENOMA, CORTICAL	0.0988	0.2210	0.6129	0.4512
	B-NEOPLASM, MEDULLARY, BE	0.9949	0.9964	0.9961	0.9675
	M-CARCINOMA, CORTICAL	0.7326	1.0000	1.0000	1.0000
Bone, unspec.	M-OSTEOSARCOMA	1.0000	1.0000	0.7596	0.4940
Brain	B-ASTROCYTOMA, BENIGN	0.4444	.	.	.
	M-ASTROCYTOMA, MALIGNANT	0.4651	.	0.5584	.
	M-GLIOMA, MIXED, MALIGNANT	1.0000	0.7826	0.6845	1.0000
Eye(s)	M-SCHWANNOMA, INTRAOCUL,	0.4651	.	.	.
Heart	B-PARAGANGLIOMA	.	0.4713	.	.
	M-SCHWANNOMA, ENDOCARD, M	0.7082	1.0000	1.0000	1.0000
Kidney(s)	B-ADENOMA	1.0000	0.9000	0.7936	1.0000
	B-LIPOMA	0.4651	.	.	.
	B-TERATOMA	.	0.5208	.	.
	M-CARCINOMA	1.0000	0.1086	1.0000	0.7737
Large Intestine	M-HISTIOCYTOMA, FIBROUS	1.0000	1.0000	1.0000	1.0000
	M-LEIOMYOSARCOMA	1.0000	1.0000	1.0000	1.0000
Liver	B-ADENOMA, HEPATOCELLULAR	0.4651	.	.	.
	M-CARCINOMA, HEPATOCELLULAR	0.9200	0.4502	0.7745	0.8934
	M-CHOLANGIOCARCINOMA	.	.	.	0.2104
Lymph node, uns.	B-LYMPHANGIOMA	.	.	0.5577	.
Lymphoreticular	M-LYMPHOMA, MALIGNANT	.	0.1330	0.5584	.
	M-SARCOMA, HISTIOCYTIC	0.7698	0.7919	1.0000	1.0000
Mammary gland	B-FIBROADENOMA	1.0000	0.7744	0.7969	1.0000
Mesenteric node	M-HEMANGIOSARCOMA	.	.	0.3182	.
Pancreas	B-ADENOMA, ISLET CELL	0.3139	0.1202	0.9853	0.4742
	B-LIPOMA	.	.	0.5577	.
	M-ADENOCARCINOMA - MIXED	.	0.4211	.	.
	M-ADENOCARCINOMA, ACINAR	.	.	.	0.4000
	M-CARCINOMA, ISLET CELL	0.6063	0.9545	0.8345	0.6086

Appears This Way
On Original

Table 4A (Continued)

Pairwise Comparisons of Treated Groups with Control
Male Rat - Fed Over 104 Weeks

Organ	Tumor	Cont. vs 50mg	Cont. vs 100mg	Cont. vs 500mg	Cont. vs 900mg
Parathyroid(s)	B-ADENOMA	.	.	.	0.2986
Pituitary	B-ADENOMA, PARS DISTALIS	0.6314	0.4319	0.7856	0.3220
	B-ADENOMA, PARS INTERMEDI	1.0000	1.0000	1.0000	1.0000
	M-CARCINOMA, PARS DISTA	0.6406	0.6886	0.9095	0.6503
Prostate	M-ADENOCARCINOMA	.	0.4904	.	.
Salivary gland	M-ADENOCARCINOMA	0.4607	.	.	.
Seminal vesicle	B-ADENOMA	.	.	.	0.6316
	B-LEIOMYOMA	.	0.4211	.	.
Skin and adnexa	B-BASAL CELL TUMOR, BENIG	.	.	.	0.2104
	B-FIBROMA	0.8056	0.6224	1.0000	0.6791
	B-HAIR FOLLICLE TUMOR-BEN	0.8056	0.3263	0.2901	0.4705
	B-HISTIOCYTOMA, FIBROUS,	0.7028	1.0000	1.0000	1.0000
	B-LIPOMA	0.5000	0.5208	0.3182	.
	B-PAPILLOMA, SQUAMOUS CEL	1.0000	1.0000	0.8016	0.7582
	B-SEBACEOUS ADENOMA	1.0000	1.0000	1.0000	0.7462
	M-FIBROSARCOMA	.	.	0.5333	.
	M-FIBROUS HISTIOCYTOMA	0.2889	0.1773	.	.
	M-HEMANGIOPERICYTOMA, MAL	.	.	0.5600	.
Small Intestine	M-SCHWANNOMA, MALIGNANT	1.0000	1.0000	1.0000	1.0000
	M-SQUAMOUS CELL CARCINOMA	0.5000	.	.	.
	B-ADENOMA	.	.	.	0.4651
Spleen	M-ADENOCARCINOMA	1.0000	0.7870	1.0000	1.0000
	M-HEMANGIOSARCOMA	0.4444	0.4211	0.5600	0.5769
Testis(es)	B-ADENOMA, LEYDIG CELL	.	.	.	0.2683
Thymus	B-THYMOMA, BENIGN	0.4634	.	.	.
Thyroid	B-ADENOMA, C-CELL	0.9554	0.1368	0.4147	0.5779
	B-ADENOMA, FOLLICULAR CEL	0.2381	0.4211	0.5577	0.0307
	M-CARCINOMA, C-CELL	0.4553	1.0000	1.0000	1.0000
Tongue	M-SCHWANNOMA	.	0.5208	.	.
Urinary bladder	B-LEIOMYOMA	.	.	.	0.4000
	B-PAPILLOMA, TRANSITIONAL	0.6500	.	.	.

Appears This Way
On Original

Table 4B

**Pairwise Comparisons of Treated Groups with Control
Female Rat - Fed Over 104 Weeks**

Organ	Tumor	Cont. vs 50mg	Cont. vs 100mg	Cont. vs 500mg	Cont. vs 900mg
Adrenal(s)	B-ADENOMA, CORTICAL	0.9488	0.6466	0.5643	0.8412
	B-NEOPLASM, MEDULLARY, BE	0.1490	.	0.3157	0.4815
	M-MEDULLARY CARCINOMA	0.5400	0.5208	.	.
Brain	B-NEOPLASM, GRANULAR CE,	1.0000	1.0000	1.0000	1.0000
	B-PAPILLOMA, CHOROID PLEX	1.0000	1.0000	1.0000	1.0000
	M-ASTROCYTOMA, MALIGNANT	0.1450	.	0.5660	0.5660
	M-GLIOMA, MIXED, MALIGNAN	.	.	.	0.2917
Cervix	M-SARCOMA, MENINGEAL	1.0000	1.0000	1.0000	1.0000
	B-POLYP, ENDOMETRIAL STRO	.	.	0.4815	.
Clitoral Gland	M-SCHWANNOMA	.	.	0.5385	0.5000
	B-PAPILLOMA, SQUAMOUS CEL	0.4333	.	.	.
Heart	B-SCHWANNOMA, ENDOCARDIAL	.	.	0.5000	.
	M-MESOTHEL, ATRIOCAVAL, M	1.0000	1.0000	1.0000	1.0000
Liver	B-ADENOMA, HEPATOCELLULAR	1.0000	0.7191	1.0000	1.0000
	B-CHOLANGIOMA	.	.	0.5660	.
	M-CARCINOMA, HEPATOCELLUL	0.5400	.	.	.
Lymphoreticular	M-LYMPHOMA, MALIGNANT	0.7935	1.0000	0.8164	0.5660
	M-SARCOMA, HISTIOCYTIC	.	0.2807	.	.
Mammary gland	B-ADENOMA	0.3813	0.3351	0.5528	0.3000
	B-FIBROADENOMA	0.8251	0.5611	0.9323	0.7746
	M-ADENOCA IN FIBROADENOMA	0.1581	0.1279	0.4576	0.2260
	M-ADENOCARCINOMA	0.3091	0.9558	0.7788	0.9796
	M-NEOPLASM, MIXED, MALIGN	0.5400	.	0.5185	0.5000
Mesenteric node	B-HEMANGIOMA	.	.	0.5660	.
Ovary(ies)	B-CYSTADENOMA	.	0.5208	.	.
	B-LUTEOMA, BENIGN	.	0.5208	.	.
	B-NEOPL, GRANULOSA CELL,	0.7462	0.7262	0.7329	1.0000
Pancreas	B-ADENOMA, ISLET CELL	0.7938	1.0000	0.1537	1.0000
	M-CARCINOMA, ISLET CELL	1.0000	0.5448	0.5947	0.5978
Parathyroid(s)	B-ADENOMA	0.5333	0.4000	0.4400	.
Pituitary	B-ADENOMA, PARS DISTALIS	0.6930	0.7935	0.8581	0.9502
	B-ADENOMA, PARS INTERMEDI	.	.	0.2137	.
	M-CARCINOMA, PARS DISTA	0.6142	0.5503	0.5601	0.5141

Appears This Way
On Original

Table 4B (Continued)

Pairwise Comparisons of Treated Groups with Control
Female Rat - Fed Over 104 Weeks

Organ	Tumor	Cont. vs 50mg	Cont. vs 100mg	Cont. vs 500mg	Cont. vs 900mg
Skin and adnexa	B-BASAL CELL TUMOR, BENIG	0.5400	.	.	.
	B-HAIR FOLLICLE TUMOR-BEN	0.5333	.	.	.
	B-LIPOMA	0.7908	0.7593	1.0000	1.0000
	M-FIBROSARCOMA	.	0.1626	.	.
	M-FIBROUS HISTIOCYTOMA	0.5400	.	.	.
	M-HEMANGIOPERICYTOMA, MAL	1.0000	1.0000	1.0000	1.0000
Small Intestine	B-LEIOMYOMA	1.0000	1.0000	1.0000	1.0000
	M-ADENOCARCINOMA	.	0.5208	.	.
Thymus	M-THYMOMA, MALIGNANT	0.5106	.	.	.
Thyroid	B-ADENOMA, C-CELL	0.4382	0.3437	0.7163	0.3035
	B-ADENOMA, FOLLICULAR CEL	0.5400	.	0.2830	0.1213
	M-CARCINOMA, C-CELL	1.0000	1.0000	1.0000	1.0000
Turbinate	M-CARCINOMA, NOS	.	.	0.5000	.
Urinary bladder	B-LIPOMA	.	.	0.5769	.
	B-PAPILLOMA, TRANSITIONAL	.	.	.	0.5769
Uterus	B-LIPOMA	1.0000	1.0000	1.0000	1.0000
	B-NEOPL, GRANULAR CELL, B	0.5400	.	.	.
	B-POLYP, ENDOMETRIAL STRO	0.8647	0.6523	0.7415	0.3851
	M-ADENOCARCINOMA	.	.	0.5000	0.4815
	M-FIBROUS HISTIOCYTOMA	1.0000	1.0000	1.0000	1.0000
	M-SARCOMA, ENDOMETRIAL ST	.	.	0.5000	.
	M-SCHWANNOMA	0.5096	.	.	.
Vagina	B-PAPILLOMA, SQUAMOUS CEL	.	0.5319	.	.
	B-POLYP, STROMAL	.	0.4138	.	.
Zymbal's gland	M-CARCINOMA, SEBACEOUS CE	.	0.4138	.	.

Appears This Way
On Original

Figure 1A: Kaplan-Meier Survival Functions for Male Rats
Species: Rat, Sex: Male, MDA 22128

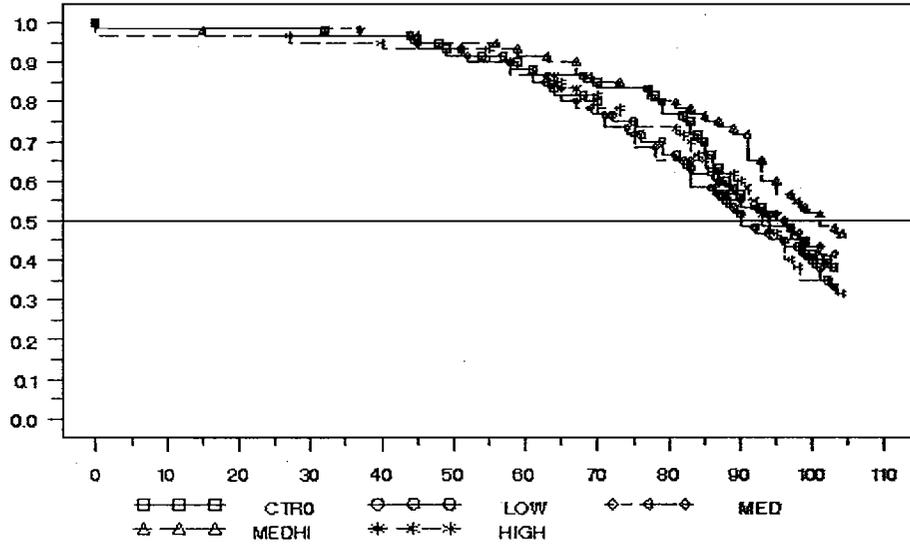
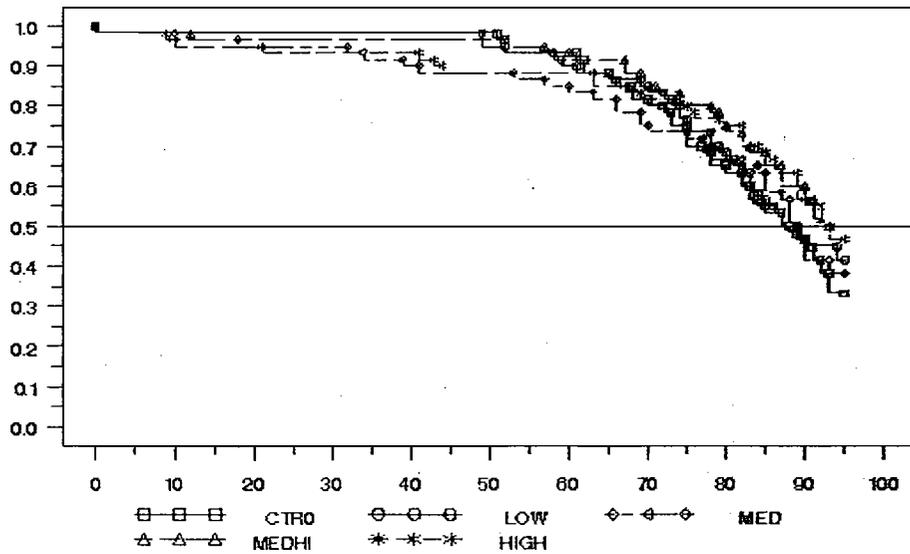


Figure 1B: Kaplan-Meier Survival Functions for Female Rats
Species: Rat, Sex: Female, MDA 22128



5. References:

- Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J. Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva*, 311-426, 1980.
- Cox D. R. "Regression models and life tables", *Journal of the Royal Statistical Society*, B, 34, 187-220, 1972.
- Gehan "A generalized Wilcoxon test for comparing arbitrarily singly censored samples", *Biometrika*, 52, 203-223, 1965.
- Lin K.K. and Rahman M.A., "Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics*, 8(1), 1-15, 1998.
- Haseman, J., "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339, 1983.
- Haseman J. "Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies", *Environmental Health Perspectives*, Vol. 58, pp 385-392, 1984.
- Haseman J. "Issues in carcinogenicity testing: Dose selection", *Fundamental and Applied Toxicology*, Vol. 5, pp 66-78, 1985.
- Chu, Cueto and Ward "Factors in the evaluation of 200 national cancer institute carcinogen bioassay", *Journal of Toxicology and environmental Health*. Vol. 8, pp 251-280, 1981.
- Lin KK, Ali MW. Statistical review and evaluation of animal tumorigenicity studies. In: Buchner CR, Tsay JY, editors. *Statistics in the pharmaceutical industry*, 2nd ed, revised and expanded. New York: Marcel Dekker, Inc; 1994.
- Mantel, N. Assessing laboratory evidence for neoplastic activity. *Biometrics* 1980; 36:381-400.

Appears This Way
On Original

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

/s/

Atiar Rahman
2/2/2007 10:57:46 AM
BIOMETRICS

Karl Lin
2/5/2007 10:03:21 AM
BIOMETRICS
Concur with review