CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-500

Statistical Review(s)

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

NEW DRUG APPLICATION

CLINICAL STUDIES

NDA/Serial Number:	21-500/ N000
Drug Name:	Emtriva® (emtricitabine, 2'3'Dideoxy-5-fluoro-3'thiacytidine) 200 mg tablets
Indication(s):	Treatment of HIV infection
Applicant:	Gilead Sciences, Inc.
Dates:	Submitted: September 2, 2002 Received: September 5,2002
Review Priority:	Standard review
Biometrics Division:	Division of Biometrics III (HFD-725)
Statistics Reviewer:	Susan Zhou, Ph.D.
Concurring Reviewers:	Guoxing Soon, Ph.D.
Medical Division:	Division of Antiviral Drug Products (HFD-530)
Clinical Team:	Russell Fleischer, PA-C, MPH.
Project Manager:	Virginia Yoerg
Keywords:	HIV infection, antiretroviral naïve and experienced, 48-week data, Time to Loss-of-Virologic-Response (TLOVR), estimation of efficacy by emtricitabine

Document updated July 2, 2003.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Based on analyses of the Week 48 data and beyond, it is concluded:

(1) In Study FTC301A, of the two treatment arms the FTC arm had significantly better virologic suppression and immunologic responses than the d4T arm ($p<0.0001$).

(2) In Study FTC303 the FTC and 3TC treatment arms were numerically similar in maintaining HIV RNA below assay limit and immunologic responses. However, superiority of FTC over placebo in maintaining HIV RNA below assay limit could not be established without strong assumptions on the 3TC's contribution in the control regimen. The reviewer used the indinavir study 035 to derive the 3TC's contribution in efficacy in a combination regimen. Because Study 035 is small, only when we assume that more than 70% of the 3TC effect observed in Study 035 was preserved we can indirectly prove that FTC is superior to placebo at significant level 0.05.

- Since patients in FTC303 were already 3TC-experienced, it is unlikely that 3TC contribution in this population will be as large as the 3TC contribution in the 3TC-naïve population in the Study 035. The combination regimen used in the historical trial also differed from the combination in FTC303 study. Given these differences the 70% preservation rate may be too high. The 70% preservation rate is also higher than the more commonly used 50% preservation rate. Therefore, we can not firmly establish the superiority of FTC over placebo.

Overall, there is enough evidence to show that FTC is efficacious.

1.2 Brief Overview of Clinical Studies

It was decided by the FDA review team to use data from two phase III studies FTC303 and FTC301A for efficacy evaluation. Study FTC302 was excluded because violation of regulatory rules.

Study FTC303 was a phase III, randomized, open-label, non-inferiority study of emtricitabine versus lamivudine in patients who were previously on a stable triple antiretroviral therapy (ART) containing lamivudine, stavudine, or zidovudine, and a protease inhibitor (PI) or non-nucleoside reverse transcriptase inhibitor (NNRTI) had HIV-1 RNA $\leq$ 400 copies/mL. Patients were randomized 1:2 to emtricitabine 200 mg q.d. +ZDV/d4T+ NNRTI/PI (FTC treatment arm), and lamivudine 150 mg bid +ZDV/d4T+ NNRTI/PI (3TC treatment arm). Patients were stratified based on plasma HIV-1 RNA levels and the background therapy. The primary objective was to establish the non-

inferiority of antiviral activity between emtricitabine and lamivudine by comparing the proportion of randomized subjects at weeks 48 whose plasma HIV-1 RNA level remains at or below 400 copies/mL.

The analysis population is the treated population consisting of 294 subjects in the FTC treatment arm, and 146 subjects in the 3TC treatment arm. The two treatment arms in FTC303 were well balanced with respect to demographic variables and baseline characteristics, except for duration of ART prior to entry. The treated population was 86% male, 64% Caucasian, 21% Black and 13% Hispanic. The mean age was 42 years with the median age 41 and range 22-80 years. The mean baseline CD4+ cell counts was 527, with a median 488 and range 37-1909 cells/mm^3. The mean baseline HIV RNA was 1.8 log_{10} copies/mL with a median 1.7 and the range 1.7-4.0 log_{10} copies/mL. Overall, the median duration of prior antiretroviral therapy was 27.6 months: 29.5 months for the FTC arm and 23.9 months for the 3TC arm respectively. The completion rate was 78.6% with a mean duration of 42.6 weeks in the treated population and the proportion of completion was slightly lower in the FTC treatment arm.

Study FTC301A was a phase III, randomized, double-blind non-inferiority trial comparing emtricitabine to stavudine within a triple drug combination containing didanosine plus efavirenz in antiretroviral-drug naïve HIV-1 infected patients. The primary objective was to assess the safety and efficacy of emtricitabine versus stavudine. Patients were randomized 1:1 to emtricitabine 200 mg QD +stavudine placebo BID +didanosine (250 mg QD if <60kg, or otherwise 400 mg QD) +efavirenz (600 mg QD) (FTC treatment arm), and emtricitabine placebo QD +stavudine (30 mg BID if < 60kg, or otherwise 40 mg BID) +didanosine (250 mg QD if <60kg, or otherwise 400 mg QD) +efavirenz (600 mg QD) (d4T treatment arm). Patients were stratified based on plasma HIV-1 RNA levels and geographic regions.

The analysis population is the treated population consisting 286 subjects in the FTC treatment arm and 285 subjects in the d4T treatment arm. The two treatment arms in FTC301A were well balanced with respect to demographic variables and baseline characteristics. The treated population had a mean age of 36 years. 85% was male, was 52% Caucasian, 26% Hispanic and 16% Black, 45% from North America, 33% from South American, and 22% from European counties. The mean and median baseline CD4+ cell count were 318 and 288 cells/mm^3, respectively. The mean baseline HIV RNA was 4.8 in log_{10} copies/mL. The treated population included 57% subjects with screening HIV RNA between 5000 and 100,000 copies/mL, 40% with screening HIV RNA above 100,000 copies/mL and 3% previously treated under the FTC301-protocol. Prior to entry, 21% of the subjects in the treated population had history of HIV-1 related events. The overall completion rate was 77% and the proportion of completion was 11% higher in the FTC treatment arm.

1.3 Statistical Issues and Findings

1.3.1 Study FTC303

1.3.1.1 Similarities and discrepancies in the efficacy of the two study arms

Sensitivity analyses of efficacy endpoints based on Time to Loss-of-Virologic Response (TLOVR) data were obtained according to the definition of Loss-of-Virologic Response (LOVR) by the DAVDP, and the results were summarized in 1 and 2 below. Comparisons of longitudinal Kaplan-Meier curves for TLOVR in subgroups of gender, age, ethnic and other risk factors were performed using the entire TLOVR dataset at Week 48 and beyond based on LOQ=400 copies/mL. The findings are listed in 4 and 5 below.

1. The proportion of subjects with HIV RNA below 400 copies/mL at Week 48 between the treatment arms was 75.5% in the FTC treatment arm and 80.8% in the 3TC treatment arm. The stratum-adjusted difference (FTC treatment arm - 3TC treatment arm) was –4.8% which favors the 3TC treatment arm with the lower 95% confidence limit of -12.8%, meaning the FTC arm is at most 12.8% worse than the 3TC arm with 95% confidence.

2. The proportion of subjects with HIV RNA below 50 copies/mL at Week 48 was 66.3% in the FTC treatment arm and 69.9% in the 3TC treatment arm, respectively. The stratum-adjusted difference (FTC treatment arm - 3TC treatment arm) was –3.6% which favors the 3TC treatment arm with the lower confidence limit of –11.7%.

3. Immunological responses measured by mean increase from baseline in CD4+ cell count and CD4% at Week 48 were numerically not very different in both treatment arms. The mean increase from baseline in CD4+ cell count was 29 cells/mm^3 in the FTC treatment arm and 61 cells/mm^3 in the 3TC treatment arm. The mean increase from baseline in CD4% was 1.6% in the 3TC treatment arm and 2.4% in the FTC treatment arm. The mean difference in CD4+ cell increase was 32 cells/mm^3 which favors the 3TC treatment arm p=0.047 by the Wilcoxon test, and the mean difference in CD4% was 0.8% which favors the FTC treatment arm with p=0.045 by the Wilcoxon test.

   - Significant interactions in baseline CD4+ (p=0.0110, F-test) and change from baseline in CD4+ cell count (p=0.0062, F-test) between treatment arms and randomization strata were observed in a mixed-effect model with compound symmetry. Further analyses showed that the significant differences in CD4+ temporal trend were associated with differences in randomization strata of the 3TC arm, not those in the FTC arm. There is no evidence of treatment difference among each stratum regarding the CD4+ temporal trend.

4. Comparisons of longitudinal Kaplan-Meier curves for TLOVR dataset with LOQ=400

copies/mL were conducted to investigate associations between LOVR and demographic factors such as gender, race and age, and to identify treatment differences in TLOVR. The results show that gender and age were significantly associated with TLOVR, and differences in K-M curves between subgroups were observed. No association between TLOVR and ethnic groups was identified. Significance was measured by the log-rank test and significant findings ($p<0.05$) in the subgroups are as follows:

- Female subjects in the 3TC treatment arm had maintained an undetected HIV-1 RNA better than those in the FTC treatment arm.
- In the FTC treatment arm, male subjects had maintained an undetected HIV-1 RNA better than female subjects.
- Younger subjects below 42 years in the 3TC treatment arm had maintained an undetected HIV-1 RNA better than those in the FTC treatment arm.
- In the FTC treatment arm, younger subjects below 42 years had maintained an undetected HIV-1 RNA better than older subjects.
- When stratifying by randomization stratum for age subgroups, the significance remained for Randomization Stratum 1 (not other strata) where subjects had screening HIV-1 RNA below 50 copies/mL and were treated with PI.

5. Comparisons of longitudinal Kaplan-Meier curves for the TLOVR dataset with LOQ=400 copies/mL were conducted to investigate associations between LOVR and other risk factors such as randomization stratum, history of Combivir use, history of mono or dual ART with lamivudine, history of HIV-1 related event, and duration of previous ART, and to identify treatment differences in subgroups. The results show that randomization stratum and history of Combivir use were significantly associated with TLOVR, and significant differences in K-M curves between subgroups were observed. No association between TLOVR and other risk factors was identified. Significance was measured by the log-rank test and significant findings ($p<0.05$) in the subgroup analysis are listed below.

   - Subjects in the 3TC treatment arm with screening HIV RNA $\leq$ 50 copies/mL had maintained an undetected HIV-1 RNA better than those with screening HIV RNA between 50 and 400 copies/mL. In the FTC arm, no significant difference in K-M curves was identified between randomization strata.
   - Subjects treated with combivir (AZT+3TC) prior to entry of the study had maintained an undetected HIV-1 RNA better than those subjects with other treatment. Consistent results on subjects with screening HIV-1 RNA below 50 copies/mL and treated with PI (Randomization Stratum 1) were observed.

1.3.1.2 Efficacy contributed by emtricitabine in this study design is unclear

The stratum-adjusted difference (FTC treatment arm - 3TC treatment arm) in percentage of subjects with HIV RNA below 400 copies/mL was –4.8% with the stratum-adjusted lower 95% confidence limit of –12.8%, in favor of the 3TC treatment arm. Based on this result, this reviewer can not judge the efficacy contributed by emtricitabine in this triple-drug regimen. The reasons are as follows.

(1) Although a non-inferiority margin of –15% was pre-specified in the protocol, and the lower 95% CI limit of the primary endpoint was greater than this value. However, this margin was only appropriate for sample size calculation, not an accepted criterion for non-inferiority judgement.

(2) One can not directly infer efficacy of emtricitabine based on the FTC303 study design. Though the two treatment arms in Study 303 were numerically similar in efficacy in maintaining an undetectable HIV RNA and immunologic responses, the lamivudine contribution in a triple-drug regimen has not been well established. Therefore, the non-inferiority of emtricitabine contribution in this Study FTC303 can not be inferred directly without further analysis and assumptions.

1.3.1.3 Estimation of efficacy contributed by emtricitabine via statistical modeling

Statistical modeling was carried out to estimate efficacy of emtricitabline indirectly using information from a historical lamivudine study (035 in NDA 20-683) and the current Study FTC303. In NDA 20-685, the applicant was seeking the approval of indinavir sulfate. In Study 035, approximately 80% subjects in the indinavir +ZDV+3TC arm (n=40) and 39.2% subjects in indinavir arm (n=28) had plasma HIV-1 RNA below 500 copies/mL at Week 24. Note subjects in indinavir arm were ZDV-experienced (median exposure 30 months), protease-inhibitor- and lamivudine-naïve. Because of the extensive ZDV exposure, the reviewer assumes that the difference between the two arms in the 035 can be attributed mainly to 3TC.

Assuming drug efficacy to be ε, then the efficacy of emtricitabine relative to placebo is:

$$\varepsilon(\text{FTC-Placebo}) = \varepsilon(\text{FTC-3TC}) + \varepsilon(\text{3TC-Placebo}) \qquad \text{(Eq.1)}$$

The estimate of ε (3TC-Placebo) obtained from Study 035 may need to be discounted by a factor $1-\lambda$, $(0 \leq \lambda \leq 1)$, which reflects the uncertainty associated with using data from a historical trial, possible contribution of the ZDV, and dissimilarity of the Study FTC303 and Study 035.

In the following, the FDA reviewer first obtained the efficacy of emtricitabine relative to placebo at Week 24.

- In Study FTC303, ε (FTC-3TC) can be estimated as difference in proportion of subjects

with plasma HIV-1 RNA below LOQ=400 copies/mL at Week 24 between FTC and 3TC arms. The estimate of [ε(FTC-3TC)] was -3.3%, which favors the 3TC arm[1]. The standard error of [ε(FTC-3TC)] was 3.6%. Using HIV-1 RNA data in Study FTC303, at Week 24, the difference of ε(FTC-3TC) by LOQ=400 copies/mL and ε (FTC-3TC) by LOQ=500 copies/mL = -0.5%. If one used LOQ=500 copies/mL, the estimate of ε (FTC-3TC) would be –2.8% with the adjusted standard error of ε(FTC-3TC) by LOQ=500 copies/mL to be 3.6% ignoring the variance by difference in LOQ.

- The second term in the right side of Eq. 1 can be estimated from Study 035 in NDA 20-685. The mean difference (3TC-Placebo) in proportion of subjects with plasma HIV-1 RNA below 500 copies/mL at Week 24 was 40.8% and standard error of the difference was 11% using an exact method and 12% using Normal approximation.

- The overall standard error for the right side of Eq. 1 would be 11.6% to 12.5%, if no inflation of the variance is assumed. Assuming one can replicate the lamivudine efficacy in a clinical trial like Study 035 with limited effect attributed to ZDV and a study population similar to FTC303, with a discount 5%-30% or at least 70% of the effects in the historical trial would be applicable to the FTC303, the estimated efficacy of emtricitabine would be between 25%-35% at Week 24, with the lower 95% CI limit ranging 1.3%-11.1% if the standard error was 11.6%. These indicate that emtricitabline is at least 1.3% better than placebo with 95% confidence even with a 30% discount of the results from Study 035.

Since Study 035 was a 24-week study, in order to use it for FTC303 Week 48 evaluation, additional assumptions would be needed to bridge the virologic response (1-LOVR) at Week 24 and the virologic response at Week 48. Denoting the probability of LOVR-free at time t is $P(\text{FTC-Placebo})_t$, the efficacy at Week 48 can be estimated:

$$P(\text{FTC-Placebo})_{48} = \kappa\, P(\text{FTC-Placebo})_{24} \qquad \text{(Eq.2)}$$

where κ is a ratio of the two probabilities. Using the FTC303 TLOVR data, the ratio of 1-LOVR at Week 48 vs. that at Week 24 for the FTC arm is 0.59. Assuming that this ratio is similar to FTC-Placebo, then the effective size ε(FTC-Placebo) at Week 48 would be 15% to 21%.

Note that FTC303 was conducted in subjects who had exposed to 3TC prior to entry, while subjects in the Study 035 were 3TC naïve. Therefore, it is likely the 3TC would be less effective in FTC303 population than that in Study 035.

1.3.2 Study FTC301A

1.3.2.1 FTC Treatment Arm Showed Significant Benefit on Primary Efficacy Endpoint

The Study FTC301A demonstrated statistical superiority of the proportion of subjects with HIV RNA below 50 copies/mL at Week 48: 78% in the FTC treatment arm and 59% in the d4T treatment arm ($p<0.0001$). The stratum-adjusted difference (FTC treatment arm - d4T treatment arm) was 19.6% with the 95% CI of 12.1% to 27.1%, meaning the FTC arm is at least 12% better in achieving a virologic response than the d4T arm.

1.3.2.2 FTC Treatment Arm Showed Significant Benefit on Other Efficacy Endpoints

1) Superiority of the proportion of subjects with HIV RNA below 400 copies/mL at Week 48 was achieved ($p<0.0001$): 81% in the FTC treatment arm and 67% in the d4T treatment arm. The stratum-adjusted difference (FTC treatment arm - d4T treatment arm) was 14%, with the 95% CI of 7.2% to 21.4%.

2) Superiority of time-weighted average change (TAD) from baseline in plasma HIV RNA through Week 48 was achieved. The estimated difference (FTC-d4T) in TAD change from baseline in plasma HIV RNA through Week 48 adjusting for screening HIV RNA was 0.24 $\log_{10}$ copies/mL with a 95% confidence interval of (0.073,0.405) by the base value carry forward (BOCF) method, which significantly favors the FTC treatment arm, $p<0.05$.

3) Immunological responses at Week 48 were statistically significantly better for subjects in the FTC treatment arm than those in the d4T treatment arm. The mean increase from baseline in CD4+ cell count was 168 cells/mm^3 in the FTC treatment arm and 128 cells/mm^3 in the d4T treatment arm. The randomization stratum adjusted TAD difference (FTC-d4T) in CD4+ cell count through Week 48 by last value carry forward (LVCF) method was 29 cells/mm^3 with 95% CI of 10 to 47 cells/mm^3, $p<0.05$. The mean increase from baseline in CD4% was 5.8% in the d4T treatment arm and 9.3% in the FTC treatment arm. The randomization stratum adjusted TAD difference (FTC-d4T) in CD4% through Week 48 by the LVCF method was 2.5% with 95% CI of 1.7% to 3.2%, $p<0.05$. Both results showed that the immunologic responses were significantly in favor of the FTC treatment arm.

4) Comparisons of longitudinal K-M curves for the entire TLOVR dataset Week 48 and beyond with LOQ=400 copies/mL were conducted to investigate the associations between LOVR and subgroups of gender, race, age, randomization stratum, history of HIV-1 related event prior to entry and subjects who had higher tolerability. The results show significant differences in LOVR between subgroups of gender, race, age, randomization stratum and subjects who had higher tolerability. No association between LOVR and history of HIV-1 related event was observed. Significance was measured by the log-rank test and significant findings were based on $p<0.05$ in the subgroup comparisons. In the following, the main analysis was based on the

comparison of K-M between treatment arms, where subjects in the FTC arm had significant better virologic response.

- **Gender:** LOVR results among female subjects were consistent with the main analysis. However, in the male subgroups, no significant difference in K-M curves was observed between the two treatment arms. In the d4T treatment arm, male subjects had significant better virologic response than female subjects. No significant gender difference was observed in the FTC arm.

- **Age:** LOVR results among subjects 35 years or younger were consistent with main analysis. However, no significant difference in K-M curves was observed between the two treatment arms for older subjects (>35 year). No significant difference in K-M curves was observed between older and younger subgroups in the FTC arm or in the d4T arm.

- **Race:** LOVR results in Hispanic subjects consistent with main analysis. However, in the White or Black subgroups, no significant difference in K-M curves was observed between the two treatment arms. In the d4T treatment arm, the virologic response for the Caucasians was the best, followed by Hispanic subjects, and Black. Further analysis on subjects with screening HIV RNA ≤100k copies/mL confirmed this finding. No significant race difference was observed in the FTC arm.

- **Randomization Stratum:** In both treatment arms, subjects with screening HIV RNA ≤100k copies/mL had significant better virologic response than those with screening HIV RNA >100k copies/mL. No significant differences in K-M were observed between the two treatment arms for higher (>100k copies/mL) or lower (≤100k copies/mL) screening HIV RNA stratum respectively.

- **Tolerability:** Among subjects who had at least eight weeks on study, LOVR results were consistent to those results of 'Randomization Stratum' with one exception: for subjects with screening HIV RNA ≤ 100 k copies/mL, subjects in the FTC arm had significant better virologic response than those in the d4T arm.

2. INTRODUCTION

2.1 Overview

This is a statistical review of the New Drug Application, NDA 21-500, for Emtriva® Capsules (FTC, emtricitabine) which is a nucleoside analogue, shown to be potent and selective inhibitor of Human Immunodeficiency Virus. The applicant Gilead Sciences, Inc. who acquired Triangle Pharmaceuticals, Inc. in 2003, is now seeking traditional approval of FTC upon review of this submission.

This submission consists of many studies. Among them Study FTC-303 and Study FTC-301A are relevant for the evaluation of efficacy endpoints. Study FTC302 was excluded because violation of regulatory rules.

2.2 Data Sources

2.2.1 Protocols and Amendments

2.2.1.1 Study FTC303

Title: A randomized, Open-label Non-inferiority Study of FTC Versus Lamivudine in Patients On a Stable Triple Antiretroviral Therapy Arm Containing Lamivudine, Stavudine, or Zidovudine, and a Protease Inhibitor Or Non-Nucleoside Reverse Transcriptase Inhibitor.

Study FTC303 was a randomized, active controlled, parallel group, open label, multi-center, non-inferiority study in antiretroviral treatment experienced HIV-infected individuals. At screening, patients had plasma RNA levels ≤ 400 copies/mL, and must be stable on an antiretroviral therapy arm for at least 12 weeks, containing lamivudine, NRTI (stavudine or zidovudine), PI (nelfinavir, ritonavir, soft gel saquinavir and indinavir), or NNRTI (delavirdine, efavirenz, or nevirapine).

Patients were randomized to the following two arms with in a 2:1 ratio:

Arm 1: Switch lamivudine to FTC (200 mg q.d.) in current arm (N=300)

Arm 2: Continue on current lamivudine containing (150 mg bid) arm (N=150).

Patients were randomization to one of four strata:

Stratum 1- screening HIV viral load of <50 copies/mL; PI in treatment arm;

Stratum 2- screening HIV viral load of <50 copies/mL; NNRTI in treatment arm;

Stratum 3- screening HIV viral load of 50-400 copies/mL; PI in treatment arm;

Stratum 4- screening HIV viral load of 50-400 copies/mL; NNRTI in treatment arm.

The study visits were scheduled at screening and at baseline, Week 2, Week 4, and every 4 weeks thereafter to Week 48. Plasma HIV viral load levels were measured for each study visit, and were determined by one central laboratory using the — Ultrasensitive assays. Results were sent to the individual investigators within 2 weeks of submission of the sample to the central laboratory. Note the results sent to the sites were blinded below 400 copies/mL. Lymphocyte subsets were measured at Weeks 12, 24, 36 and 48. The duration of treatment was 48 weeks and antiviral responses were assessed at week 24 and week 48.

2.2.1.2 Study FTC301A

Title: A Randomized, Double-Blind, Non-inferiority Trial Comparing Emtricitabine to Stavudine within a triple Drug Combination Containing Didanosine Plus Efavirenz in Antiretroviral-Drug Naïve HIV-1 infected Patients.

Study FTC301A enrolled HIV-1 infected and antiretroviral treatment naïve patients and who have plasma HIV RNA ≥ 5000 copies/mL. Twenty study subjects who were treated in Study FTC301 prior to Amendment 2 were allowed to enroll after a minimum 4-week washout period. Study subjects were randomly assigned in a 1:1 ratio to FTC (200 mg QD) and d4T (40 mg BID if ≥ 60 kg and 30 mg BID otherwise) in combination with ddI (400 mg QD if ≥ 60 kg and 250 mg QD otherwise) and efavirenz (600 mg QD). Randomization was stratified by screening HIV viral load of <100,000 copies/mL or ≥100,000 copies/mL or rolled over from trial FTC301 crossed by geographic region (North America, Europe, or Latin America). A total of nine strata were found.

1='North America:<=100k copies/mL'

2='Europe :<=100k copies/mL'

3='Latin America:<=100k copies/mL'

4='North America:>100k copies/mL'

5='Europe :>100k copies/mL'

6='Latin America:>100k copies/mL'

7='North America:FTC301 Rollover'

9='Latin America:FTC301 Rollover'

- HIV RNA levels were measured at baseline and every 4 weeks thereafter to at least 48 weeks, and antiviral responses were assessed at week 24 and week 48. Treatment was continued until the last enrolled subject completes 48 weeks of therapy.

2.2.2 Sample Size and Study Population

In Study FTC 303, the sample size calculation was based on a two-group large sample normal approximation test of proportions with a one-sided type I error of 0.025, would have a power of 85% to reject the null hypothesis that the study drug (FTC) and the standard (lamivudine) would not be equivalent, i.e., the difference in the virologic response would be 15% greater in favor of lamivudine, versus the alternative hypothesis that the virologic responses would be equivalent. The primary endpoint was defined as the proportion of participants with HIV RNA ≤ 400 copies/mL at week 48. The sample size of 300 for the FTC treatment arm and 150 for the lamivudine arm were obtained assuming that the expected mean difference in the virologic response (FTC-lamivudine) to be 5%, and the proportion of the virologic response in the lamivudine treated patients was 70%.

In Study FTC 301A, the sample size of 250 per arm was based on a two-group large sample normal approximation test of proportions with a power of 80%, one-sided type I error of 0.025, to demonstrate similarity of antiviral activity within 10% between FTC/ddI/EFV (FTC) and d4T/ddI/EFV (d4T) for the primary endpoint of proportion of participants with HIV RNA ≤ 50 copies/mL at week 48. In the sample size calculation, it was also assumed that the expected proportions of subjects with viral load ≤ 50 copies/mL at week 48 are 52.5% and 50% respectively, in the FTC and d4T treatment arms.

2.2.3 Efficacy Endpoints

The primary objective in Study FTC303 is the comparison between treatment groups the proportion of subjects with levels of plasma HIV RNA levels ≤ 400 copies/mL at Weeks 24 and 48. HIV RNA data from the ultrasensitive assays were used for efficacy analysis. The primary analysis of virologic success was accomplished using a stratum-adjusted two-sided 95% confidence interval for the difference between treatment arms in the proportion of subjects whose plasma HIV RNA ≤400 copies/mL at Week 48. Similar analysis was conduced for virologic failure, tolerability failure and effectiveness failure, as well as the proportion of subjects that had HIV-1 RNA ≤50 copies/mL. Time to virologic failure, the Average Area Under the Curve Minus Baseline (AAUCMB) for subjects with screening HIV RNA between 50-400 copies/mL were planned to be analyzed.

The definitions for the primary study endpoints in Study FTC303 including the following:

- Virologic success was defined as a continued suppression in HIV-1 RNA below 400 copies/mL.

Tolerability failure was defined as any adverse event severe enough to warrant permanent discontinuation of FTC, lamivudine, or the protocol specified triple therapy arm.

Effectiveness failure was defined as HIV-1 RNA > 400 copies/mL, after Week 4 of the study, tolerability failure, and clinical disease progression or loss to follow-up.

The primary study endpoint in Study FTC-301A was defined as a two-sided 95% stratum-adjusted confidence interval for the difference between the treatment arms in the proportion of subjects whose plasma HIV RNA ≤50 copies/mL at Week 24/48. The stratum-adjusted difference in the proportion of patients ≤50 copies/mL between two study arms and the stratum-adjusted two-sided 95% confidence interval was used.

Secondary efficacy endpoints and definitions are summarized in Table 1.

Table 1: Definitions of Secondary Efficacy Endpoints at Week 48[§]

Week 48 Efficacy Endpoints	Definitions
FTC301A	
The K-M estimates and 95% CI: Prob of virologic failure	Failure time includes those day 1 for whose who had never reached HIV RNA ≤400 copies/mL, and Rebounder at first time of rebounding
The K-M estimates and 95% CI: Prob of efficacy failure	Failure time includes the time of new CDC event, death, last visit date for those who loss-to-follow up
The K-M estimates and 95% CI: Time to effectiveness failure	Failure time includes the last date of study drug for whose who terminated study drug
Change from baseline CD4+ & %	Descriptive statistics, LOCF,
FTC303	
The K-M estimates and 95% CI: Prob of virologic failure	Failure time includes those day 1 for whose who had never reached HIV RNA ≤400 copies/mL, and Rebounder at first time of rebounding
The K-M estimates and 95% CI: Prob of efficacy failure	Failure time includes the time of new CDC event, death, last visit date for those who loss-to-follow up
The K-M estimates and 95% CI: Time to effectiveness failure	Failure time includes the last date of study drug for whose who terminated study drug
Change from baseline CD4+ & %	Descriptive statistics, LOCF,

2.2.4 Analysis Population and Subgroup Analyses

In Study FTC303, the sponsor assessed the significance of other parameters in predicting of virological failures using logistic regression model. Parameter of interest included screening and baseline plasma HIV RNA, demographics and baseline characteristics, time on stable therapy at study entry, antiretroviral medical history, and adherence to therapy.

In Study FTC-301A, four study populations were considered for the evaluation of efficacy and safety endpoints at Week 48. Table 2 shows the details.

Table 2: Study FTC301A: Analysis Populations and Definitions[s]

Analysis Population	Definitions in Addition to Randomized	Endpoints Analyzed
Intend-to-treat	at least one dose, non-completer as failure	Efficacy and Safety
As Treated	at least one dose, data collected when on study drugs	Efficacy
Randomized		Primary Efficacy
Virologic Evaluable	at least one dose, at least one HIV-1 RNA after week 12	Not specified

An exploratory analysis on the primary efficacy endpoint was planned to evaluate the predictive of the primary efficacy endpoint from baseline HIV-1 RNA, CD4+, presence of genotypic mutation at time of virologic failure, baseline patient characteristics (age, gender, ethnic origin), and therapy non-adherence.

Subgroup analyses were considered comparing the primary endpoint and selected secondary endpoints between treatment arms within subgroups of gender, ethnic origin, age and randomization stratum.

2.2.5 Plasma HIV-1 RNA and CD4+ Cell Count

2.2.5.1 Study FTC303

For the evaluation of plasma HIV-1 RNA levels, individual investigators sent specimens to a central laboratory within two weeks. The central laboratory used [redacted] (the ultrasensitive assay) to measure viral HIV-1 RNA levels and the results were sent to sites. Those results below 400 copies/ml were blinded to sites. The limit of detection (LOQ) for the ultrasensitive assay is 50 copies/mL.

The follow-up visits for plasma HIV-1 RNA were scheduled at screening, day 1, Week 2, Week 4, every 4 weeks thereafter until Week 48, and every 28 days after Week 48. In the HIV-1 RNA database, 146 (2.5%) person-visits were identified to have 2-4 observations

per person-visit for the TREATED population. Among these, 29 had discordant status in minimum and maximum, meaning that the minimum HIV RNA in the multiple observation set is below 50 or 400 copies/mL and the maximum is ≥ 50 or 400 copies/mL. Since more than 80% data points were concordant in nature, we used the mean HIV RNA to replace multiple observations per person-visit data.

The follow-up visits for CD4+ cell count were scheduled at screening, day 1, Week 12, every 12 weeks thereafter and at the end of study for both studies.

2.2.5.2 Study FTC301A

For the evaluation of plasma HIV-1 RNA levels, [redacted] ultrasensitive assay) and [redacted] (the standard assay) were used. The limits of detection (LOD) are 400 copies/mL and 50 copies/mL for the standard assay and the ultrasensitive assay, respectively.

The follow-up visits for plasma HIV-1 RNA were scheduled at screening, day 1, Week 4, every 4 weeks thereafter until Week 48, every 12 weeks post Week 48 and at the end of study.

Modified in Amendment 5, patients in South African would be excluded if results obtained with the Amplicor version 1.5 (v1.5) were ≥ 0.7 $\log_{10}$ higher than the results with the Amplicor version 1.0 (v1.0).

Amendment 6 specified the use of the [redacted] in evaluation of HIV RNA, but no particular rules were specified. The sponsor added the rules to Amendment 6 in clinical study report as follows. Patients would be eligible if (1) both v1.5 and v1.0 assay indicated screening HIV RNA ≥ 5000 copies/mL; or (2) a screening HIV RNA ≥ 5000 copies/mL by v1.5 and a screening HIV RNA < 5000 copies/mL by v1.0 and the difference between v1.5 and v1.0 was ≥ 0.7 $\log_{10}$. Otherwise, the patients were not to be eligible. According to Amendment 6, if a subject satisfied the inclusion criteria for screening plasma HIV-1 RNA, then the patient was followed throughout the study using the [redacted] assay or v1.0 assay for HIV-1 RNA measurements if the screening HIV RNA difference between the v1.5 and v1.0 assay results was ≥ 0.7 $\log_{10}$ or <0.7 $\log_{10}$.

Plasma HIV-1 RNA was measured using both the [redacted] assay. In the HIV RNA database, majority of the HIV-1 RNA measurements were obtained by the Ultrasensitive assay v1.0 (81%), followed the Standard assay v1.0 (14%), and Ultrasensitive or Standard assay v1.5 (5%).

In the presence of multiple measurement per visit date, an HIV-1 RNA value was determined first by v1.5 of the Ultrasensitive assay; or by v1.5 of the Standard assay if value of v1.5 of the Ultrasensitive assay was missing. If HIV-1 RNA by both v1.5 of the Ultrasensitive and Standard assay were missing, HIV-1 RNA measurement by v1.0 of the Ultrasensitive assay was used. Otherwise, HIV-1 RNA measurement by v1.0 of the

Standard assay was used. After selecting plasma HIV-1 RNA per visit date and deleting duplicates, there were two person-visits with repeated measurements. The mean value was replaced the repeated measurements since they both above LOQ=400 copies/mL.

The follow-up visits for CD4+ cell count were scheduled at screening, day 1, Week 12, every 12 weeks thereafter and at the end of study for both studies.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy (Study FTC303)

3.1.1 Patient Disposition

This reviewer verified subjects who completed 48 weeks treatment in the treated population. using drug administrative database in FTC303.drugad to obtain a baseline date. The treated population included 440 study subjects, among them 433 subjects had drug administrative information in the database, including 140 in the 3TC treatment arm and 293 in the FTC treatment arm; and seven subjects (6 in the 3TC arm and one in the FTC arm) had no actual drug information and were considered as premature discontinuation. It was seen that 77% in the FTC arm and 81% in the 3TC am had completed 48-weeks of study, after imputing missing dates in the database.

The applicant created the baseline date or first date of taking study drug as the last date of HIV RNA or CD4 + cell count measurement in visit '00', not first date of taking study drug in FTC303.drugad. A missing in last date of study drug was imputed using the date of study termination, the last date of HIV RNA, or CD4+ cell count measurement in order. The discrepancies between the first date of study drug dosing and date of baseline (> 7 days) were observed in 5 study subjects. Therefore, this reviewer felt it appropriate to use the applicant's definition of baseline date and last date of study in evaluation of subject disposition.

Using the applicant's method, 227 (77%) in the FTC treatment arm and 119 (82%) in the 3TC treatment arm had completed a 48-week study. 67 (23%) in the FTC treatment arm and 27 (18%) in the 3TC treatment arm discontinued from study prior to Week 48, due to adverse event, study drug non-compliance, virological failure, protocol violation, death, and other reasons. Note there were 13 subjects (4%) in the FTC treatment arm discontinued from study due to adverse events. On the contrary, no subject discontinued from study in the 3TC treatment arm due to adverse event.

3.1.2 Demographics and Baseline Characteristics

Table 10 lists demographics and baseline characteristics of study subjects for Study FTC303. The treated population was 86% male, 64% were Caucasian, 21% Black and 13% Hispanic. The mean age was 42 years with the median age (range) 41 (22-80) years. The mean baseline CD4+ cell counts was 527, median was 488 and range was 37-1909 cells/mm^3. The mean baseline HIV RNA was 1.8 $\log_{10}$ copies/mL with a median 1.7 and the range was 1.7-4.0 $\log_{10}$ copies/mL. Overall, the median duration of prior antiretroviral therapy (ART) was 27.6 months.

Table 3: Demographics and Baseline Characteristics (FTC303)[§]

		FTC	3TC
Randomized		307	152
Treated		**294**	**146**
Age (years)	Mean (std)	42 (9)	43 (9)
	Median (Range)	41 (22-70)	42 (23-80)
Gender	Male	250 (85)	127 (87)
Race	Caucasian	193 (66)	87 (60)
	Black	59 (20)	34 (23)
	Hispanic	34 (12)	24 (16)
	Other	8 (3)	1 (2)
Strata[a]	HIV RNA <50 & PI	196 (67)	100 (68)
	HIV RNA <50 & NNRTI	57 (19)	27 (19)
	HIV RNA ≥50 & PI	35 (12)	17 (12)
	HIV RNA ≥50 & NNRTI	6 (2)	2 (1)
Region	North America	294 (100)	146 (100)
	Europe	0 (0)	0 (0)
	South America	0 (0)	0 (0)
HIV RNA in $\log_{10}$	Mean (std)	1.8 (0.3)	1.8 (0.3)
	Median (Range)	1.7	1.7
CD4+ (cells/mm^3)	Mean (std)	524 (269)	533 (278)
	Median (Range)	484	503
Weight (kg)	Mean (std)	77 (16)	78 (15)
	Median (Range)	77 (22-129)	78 (34-128)
AIDS[b]	# (%)	151 (51)	65 (45)

a Unit: copies/mL.
b Any Class B and C Events prior to entry, one less in the FTC treatment arm than the applicant's.

The demographic and baseline characteristics were well balanced except for the duration of prior ART. Prior to entry, subjects in the FTC arm had a median duration of ART 29.5 months, significantly greater than that (23.9 months) for the 3TC arm, p=0.0228 by the Wilcoxon test. Additionally, subjects in the 3TC arm had a median CD4+ of 503 cells/mm^3 significantly greater than 484 cells/mm^3 in the FTC arm, p=0.079 by the Wilcoxon test.

3.1.3 Efficacy Endpoints

The plasma HIV RNA levels at screening for all the subjects were all below 400 copies/mL. Therefore, based on LOQ=400 copies/mL, virologic success for a subject implies that this subject had a continued suppression through Week 48.

3.1.3.1 Outcomes Through Week 48

Study outcomes through Week 48 were obtained using longitudinal HIV-1 RNA data since

Day 1, subjects' discontinuation status and date of study termination, and occurrence and date of adverse event. Time to Loss-of-Virologic Response (TLOVR) data were obtained according to the definition of LOVR by the DAVDP. For example, a responder through Week 48 is the one who had achieved viral load < LOQ that was confirmed later prior to or at Week 48 visit window, and was maintained viral load < LOQ without discontinuation. Likewise, a virologic failure is the one who had either viral load rebound or who had never achieved viral load < LOQ. Table 4 shows virologic outcomes through Week 48 by LOQ=50 and 400 copies/mL. Table 5 shows treatment difference in LOVR and 95% CI at Week 48 by randomization stratum and LOQ.

Table 4: Study FTC303: Outcomes of Randomized Treatment Through Week 48

Outcome	Tradename + AZT/d4T + NNRTI/PI N=294	Lamivudine + AZT/d4T + NNRTI/PI n=146
LOQ = 400 copies/mL		
Completed study through Week 48	**227 (77)**	**120 (82)**
Responder[1]	222 (76)	118 (81)
Virologic Failure[2]	5 (<2)	2 (1)
Not completed study	**67 (23)**	**26 (18)**
Discontinuation while suppressed	35 (12)	11 (8)
Death	0 (0)	1 (<1)
Adverse Event	11 (4)	0 (0)
Other Reasons	24 (8)	10 (7)
Discontinuation after Rebound	20 (7)	11 (8)
Death	0 (0)	0(0)
Adverse Event	2 (<1)	0 (0)
Other Reasons	18 (6)	11 (8)
Discontinuation & Never Suppressed	12 (4)	4 (3)
Death	0 (0)	0 (0)
Adverse Event	0 (0)	0 (0)
Other Reasons	12 (4)	4 (3)
LOQ = 50 copies/mL		
Completed study through Week 48	**227 (77)**	**120 (82)**
Responder[1]	195 (66)	102 (70)
Rebound	28 (10)	15 (10)
Never Suppressed	4 (1)	3 (2)
Not completed study	**67 (23)**	**26 (18)**
Discontinuation while suppressed	29 (10)	10 (7)
Death	0 (0)	1 (<1)
Adverse Event	10 (3)	0 (0)
Other Reasons	19 (6)	9 (6)
Discontinuation after Rebound	14 (4)	6 (4)
Death	0 (0)	0(0)
Adverse Event	3 (1)	0 (0)
Other Reasons	11 (3)	6 (4)
Discontinuation & Never Suppressed	24 (8)	10 (7)
Death	0 (0)	0 (0)
Adverse Event	0 (0)	0 (0)
Other Reasons	24 (8)	10 (7)

1 Patients maintained confirmed HIV RNA ≤ LOQ through Week 48.
2 Includes confirmed viral rebound and one responder who had protocol violation.
3 Includes lost to follow-up, patient's withdrawal, non-compliance, protocol violation and other reasons.

3.1.3.2 Primary Efficacy Endpoint: LOQ=400 copies/mL

Using the TLOVR algorithm 222 (76%) of the subjects in the FTC treatment arm and 118

(81%) of the subjects in the 3TC treatment arm were responders, i.e., those that had maintained plasma HIV RNA levels ≤ 400 copies/mL through Week 48 without a rebound or discontinuation. The stratum-adjusted difference between FTC and 3TC treatment arm was –4.8% with 95% confidence interval of –12.8% to 3.3%. The Breslow-Day test for homogeneity across strata indicated no significant stratum differences (p=0.24). Note this result is slightly different from the applicant's result due to different definitions in dealing with multiple observations in HIV RNA per person-visit and rules for modification of a missing visit. Other situation was also observed. For example, in the 3TC arm, one subject #597 had two consecutive plasma HIV RNA above 400 copies/mL, and other values were all below 50 copies/mL. According to the TLOVR definition, this subject should be defined as a failure but was considered as a responder by the applicant.

3.1.3.3 Secondary Efficacy Endpoint: LOQ=50 copies/mL

Using the TLOVR algorithm 195 (66%) of the subjects in the FTC treatment arm and 102 (70%) of the subjects in the 3TC treatment arm were responders, i.e., those that had maintained plasma HIV RNA levels ≤ 50 copies/mL through Week 48 without a rebound. The stratum-adjusted difference between FTC and 3TC treatment arms was –3.0% with 95% confidence interval of –11.7% to 5.8%. The Breslow-Day test for homogeneity across strata indicated no significant stratum differences (p=0.39). These results are analogue to the applicants': p_{FTC} =67% and p_{3TC} =72.9%, $p_{3FTC\text{-}3TC}$ =-2.7%, 95% C.I. = (-11.0%, 5.7%).

Table 5: Summary of Treatment Difference in Virologic Response by Randomization Stratum[a]

Endpoint					95% CI	
	P_{FTC}(%)	P_{3TC}(%)	$P_{FTC\text{-}3TC}$(%)	P_{ajd}(%)	lower	upper
LOQ=400 copies/mL	75.5	80.8	-5.3	-4.8	-12.8	3.3
HIV RNA <50 & PI	80.6	86.0	-5.4			
HIV RNA <50 & NNRTI	77.2	77.8	-0.6			
HIV RNA ≥50 & PI	45.7	64.7	-19.0			
HIV RNA ≥50 & NNRTI	66.7	0.0	66.7			
LOQ=50 copies/mL	66.3	69.9	-3.6	-3.0	-11.7	5.8
HIV RNA <50 & PI	73.5	74.0	-5.0			
HIV RNA <50 & NNRTI	73.7	74.1	-4.0			
HIV RNA ≥50 & PI	22.9	47.1	-24.2			
HIV RNA ≥50 & NNRTI	16.7	0.0	16.7			

a. Based on TLOVR algorithm.

3.1.3.4 Time to Loss-of-Virologic-Response (TLOVR)

Using HIV-1 RNA data, subjects' discontinuation status and other information, this

reviewer obtained time to Loss-of-Virologic-Response (TLOVR) Kaplan-Meier curves by LOQ and treatment arm as shown in Figure 1. The top two TLOVR curves in solid lines are those for FTC (FTC_50) and 3TC (3TC_50) with respect to LOQ=50 copies/mL, and the bottom two TLOVR curves in dash lines are those for LOQ=400 copies/mL.

Overall, more subjects in the FTC treatment arm had loss-of virologic response, compared with those in the 3TC treatment arm. However, there was no evidence of statistical significant difference between the TLOVR-KM curves, p>0.05, by the log-rank test.

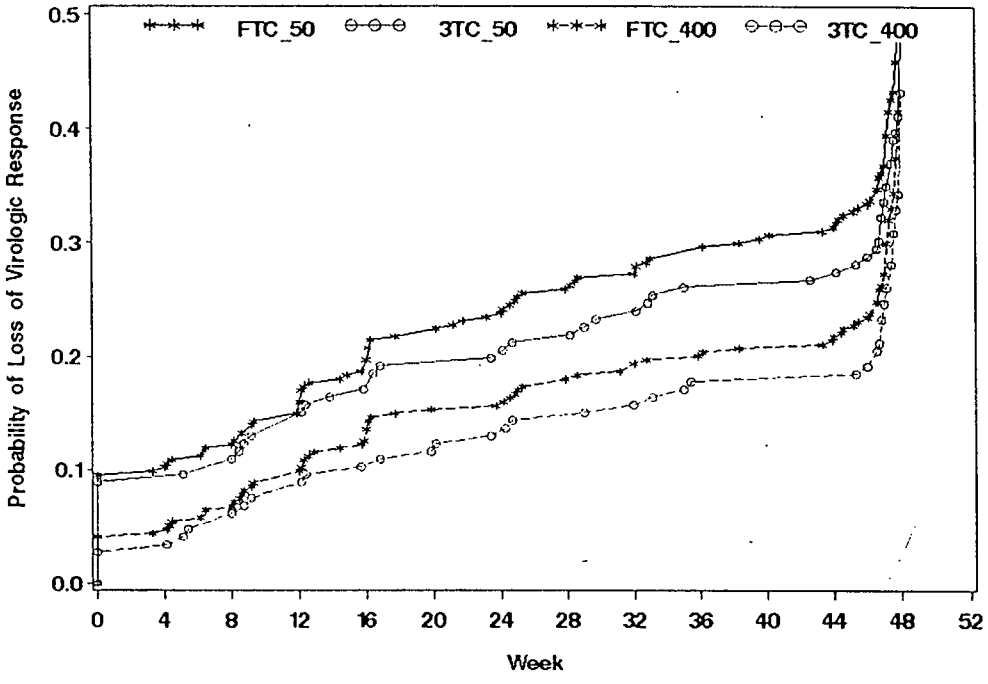

Figure 1: FTC303: Probability of Loss of Virologic Response by LOQ and Treatment Arm

Table 6: Summary of Treatment Difference in TLOVR at Week 48

Endpoint	P_{FTC}(%)	P_{3TC}(%)	$P_{FTC-3TC}$(%)	Se(%)	95% CI lower	95% CI upper
KM Probability of Failure[a]						
LOQ=50 copies/mL	33.3	28.8	4.5	4.7	-4.7	13.7
LOQ=400 copies/mL	23.5	19.2	4.3	4.1	-3.7	12.3

a. Point estimators at mid point of Week 48 window.

Table 6 shows the estimated LOVR at Week 48. The treatment differences are 4.3% to 4.5% which favors the 3TC arm. However, the 95% CIs contain zeros, indicating that the treatment differences in LOVR were not significantly different at type I error =0.05 level.

3.1.3.5 Temporal Trend in Plasma HIV RNA (log_{10} copies/mL)

- The temporal trend in plasma HIV RNA was investigated for treated population. Data were grouped into six subgroups, three per treatment arm, according to screening HIV RNA randomization stratum.

  - FTC-1 or 3TC-1 – subjects with screening HIV RNA ≤ 50 copies/mL and treated with PI;
  - FTC-2 or 3TC-2– subjects with screening HIV RNA ≤ 50 copies/mL and treated with NNRTI;
  - FTC-3 or 3TC-3 – subjects with screening HIV RNA 50 –400 copies/mL.

Figure 2 shows the mean plasma HIV RNA curves for six subgroups. For FTC-1, FTC-2, 3TC-1 and 3TC-2 groups, the mean curves are flat lines around the nadir 1.7 log_{10} copies/mL (50 copies/mL) through Week 48. The mean curve for FTC-3 lies in between 1.8 and 2.2 log_{10}, less than 2.6 log_{10} (400 copies/mL) and the mean curve for 3TC-3 mostly above the nadir but under the mean curve of the FTC-3 group. It appears that the HIV RNA suppression for subjects in the treated population were well maintained through 48 weeks of therapy.

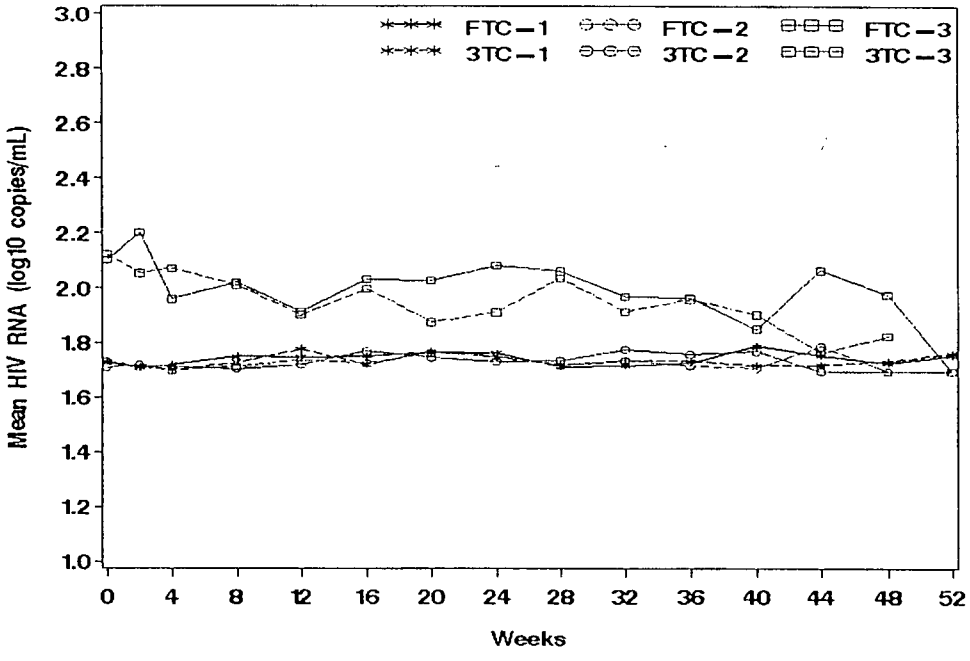

Figure 2: FTC303: Mean Plasma HIV RNA (log_{10} copies/mL) by Treatment Arm and Randomization Stratum

3.1.3.6 Change from Baseline CD4+Cell Count and CD4%

- The upper plot in Figure 3 shows the mean change from baseline in CD4+ Cell Count cells/mm^3 by treatment arm and randomization stratum. At Week 48, the mean increase was 61 cells/mm^3 in the 3TC arm, significantly greater than a mean increase of 29 cells/mm^3 in the FTC arm, p=0.047 by the Wilcoxon test. The mean change from baseline CD4+ cell count was increased for strata 1 and 2 where the increase was favorable for subjects in lamivudine groups (dashed lines). For stratum 3, it appears both groups had decline in CD4+ cell count, and subjects in the 3TC arm had more reductions in mean CD4+ cell counts.

  A significant qualitative interaction in CD4+ cell count between treatment arms and randomization strata was observed. This reviewer fitted mixed models

  $$\Delta CD4+ = (CD4+)_0 * Treatment * Stratum + Treatment * Stratum * Time \qquad Model1$$

  with compound symmetry (CS).

  Results show (1) there is an evidence of a relationship between the baseline CD4+ cell count and change from baseline CD4+. The p-value is 0.0110; and (2) there is a strong evidence of a Treatment*Stratum*Time interaction (p=0.0062), meaning that the change from baseline CD4+ were not the same for treatment arm and randomization strata.

  Further analyses were performed using data stratified by treatment arm and randomization stratum. The results show that the change from baseline CD4+ cell count for the subjects in the 3TC arm were not the same for three randomization strata, p=0.0388 by the F-test. On the contrary, for subjects in the FTC arm, the change from baseline CD4+ cell count were not significantly different for three randomization strata, p=0.5418 by the F-test.

  The bottom plot of Figure 3 shows the mean change from baseline in CD4% by treatment arm and randomization stratum. At Week 48, the mean increase from baseline in CD4 percent (ΔCD4%) was 2.5% in the FTC arm compared to 1.7% in the 3TC arm, p=0.045, by the Wilcoxon test. The mean change from baseline CD4% increased over time for most of the subgroups and the increase was favorable for the FTC subgroups. For stratum 3, however, it appears that both groups had flat or downturn up to Week 24 and then went up.

  The mean change from baseline in CD4% and CD4+ cell count had a revised treatment difference.

- It is known that CD4+ is the product of CD4% and total T-lymphocyte count, and the later includes CD4+ and CD8+ cell counts. Although the median/mean baseline CD4+ cell count was 503/533 cells/mm^3 in the 3TC arm, significantly *greater than* a median/mean of 484/524 cells/mm^3 for subjects in the FTC arm, p=0.0790, by the Wilcoxon test, the median/mean baseline CD8+ cell count was 839/932 cells/mm^3 in the 3TC arm, *less than* a median/mean of 889/950 cells/mm^3 for subjects in the FTC arm, p=0.30, by the Wilcoxon test.

- CD4% and CD4+ cell count is highly correlated at baseline, i.e., the Spearman correlation coefficient is $\rho = 0.73$ for baseline CD4% and CD4+ cell count. However, the degree of correlation between the change from baseline of the two variables is greatly reduced, i.e., $\rho = 0.17$ for ΔCD4% and ΔCD4+ cell count.

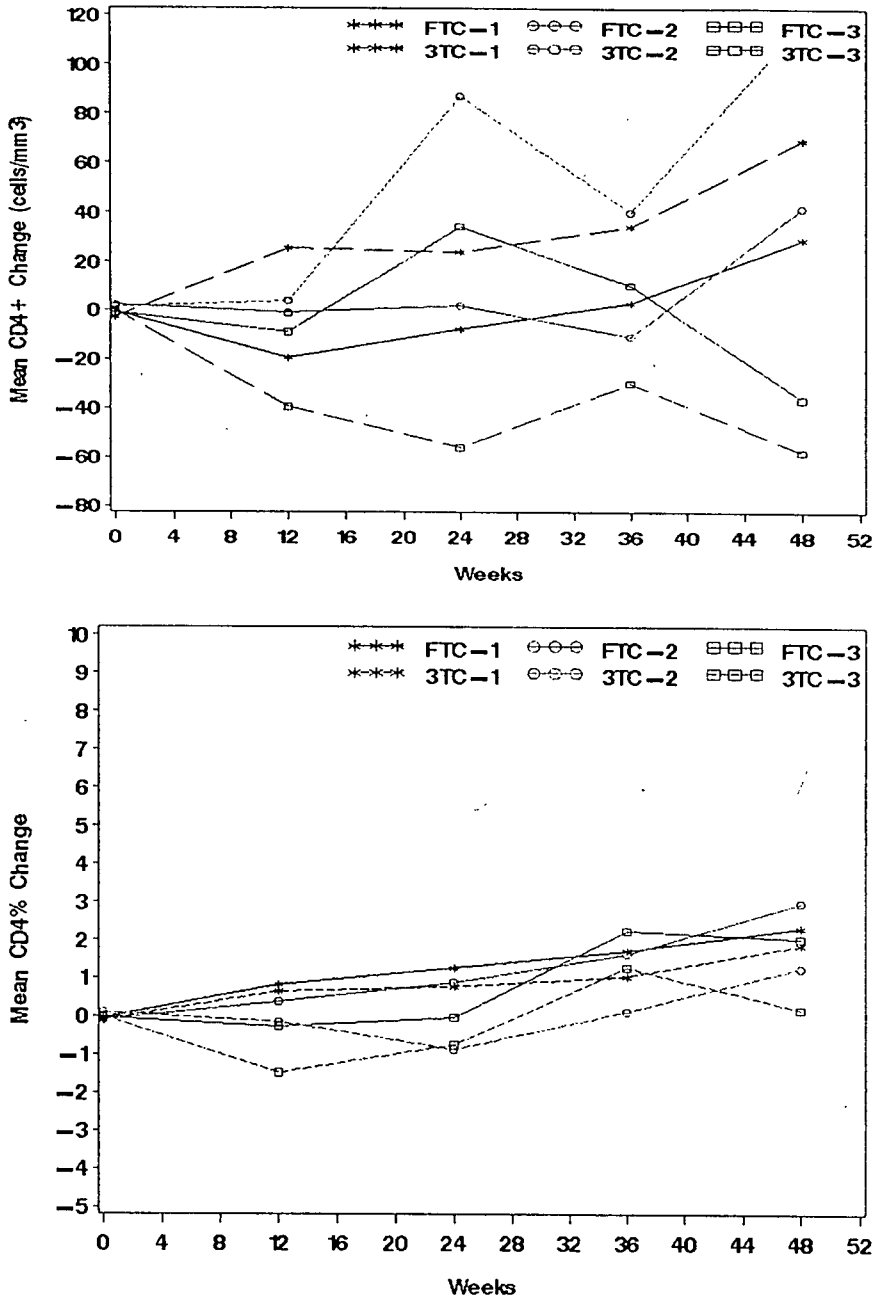

Figure 3: FTC303: Mean Change From Baseline CD4+ & CD4% by Treatment Arm and Randomization Stratum

3.1.4 Estimation of Efficacy for Emtricitabline: An Indirect Approach

The Study FTC303 did not have a control arm to estimate the efficacy of lamivudine. Therefore, the efficacy of emtricitabine can not be directly inferred by this study design. Additionally, the study subjects in Study FTC303 were exposed to lamivudine prior to entry. Resistance to lamivudine may play a major role in virologic response for the study regimens. Because of these factors, isolation of lamivudine effect from emtricitabline effect for the FTC arm would be intricate.

Statistical modeling was carried out to estimate efficacy of emtricitabline indirectly using information from a historical lamivudine study (035 in NDA 20-685) and the current Study FTC303. In NDA 20-685, the applicant was seeking the approval of indinavir sulfate. In Study 035, approximately 80% subjects in the indinavir +ZDV+3TC arm (n=40) and 39.2% subjects in indinavir arm (n=28) had plasma HIV-1 RNA below 500 copies/mL at Week 24. Note subjects in indinavir arm were ZDV-experienced (median exposure 30 months), protease-inhibitor- and lamivudine-naïve. Because of the extensive ZDV exposure, the reviewer assumes that the difference between the two arms in the 035 can be attributed mainly to 3TC.

3.1.4.1 Efficacy of emtricitabine relative to placebo at Week 24

Assuming drug efficacy to be ε, then the efficacy of emtricitabine relative to placebo is

$$\varepsilon\,(\text{FTC-Placebo}) = \varepsilon\,(\text{FTC-3TC}) + \varepsilon\,(\text{3TC-Placebo}) \qquad \text{(Eq.1)}$$

The estimate of ε (3TC-Placebo) obtained from Study 035 may need to be discounted by a factor $1-\lambda$, $(0 \le \lambda \le 1)$, which reflects the uncertainty associated with using data from a historical trial, possible contribution of the ZDV, and dissimilarity of the Study FTC303 and Study 035.

In the following, the FDA reviewer first obtained the efficacy of emtricitabine relative to placebo at Week 24.

- In Study FTC303, ε (FTC-3TC) can be estimated as difference in proportion of subjects with plasma HIV-1 RNA below LOQ=400 copies/mL at Week 24 between FTC and 3TC arms. The estimate of [ε(FTC-3TC)] was -3.3%, which favors the 3TC arm[1]. The standard error of [ε(FTC-3TC)] was 3.6%. Using HIV-1 RNA data in Study FTC303, at Week 24, the difference of ε(FTC-3TC) by LOQ=400 copies/mL and ε (FTC-3TC) by LOQ=500 copies/mL = -0.5%. If one used LOQ=500 copies/mL, the estimate of ε (FTC-3TC) would be –2.8% with the adjusted standard error of ε(FTC-3TC) by LOQ=500 copies/mL to be 3.6% ignoring the variance by difference in LOQ.

- The second term in the right side of Eq. 1 can be estimated from Study 035 in NDA 20-685. The mean difference (3TC-Placebo) in proportion of subjects with plasma HIV-1 RNA below 500 copies/mL at Week 24 was 40.8% and standard deviation of the

difference was 11% using an exact method and 12% using Normal approximation.

- The overall standard error for the right side of Eq. 1 would be 11.6% to 12.5%, if no inflation of the variance would be counted. Assuming one can replicate the lamivudine efficacy in a clinical trial like Study 035 with limited effect attributed by ZDV and a study population similar to FTC303, then with a discount between 5%-30%, the estimated efficacy of emtricitabine at Week 24 would be between 25%-35%, with the lower 95% CI limit ranging 1.3%-11.1% if the standard error was 11.6%. These indicate that emtricitabline is at least 1.3% better than placebo with 95% confidence even with a 30% discount of the results from Study 035.

Figure 4 shows the estimated efficacy of emtricitabine relative to placebo and 95% CI limits, based on λ= [5%,45%] and the overall standard error in right side of Eq. 1 of 11.6%. The middle line denoted by 'eff' is the estimated ε (FTC-Placebo), the line in the bottom denoted by 'lower' is the lower 95% CI limit. Likewise, the line in the top denoted by 'upper' is the upper 95% CI limit.

1. Source: Text table 4, Summary of treatment Differences in TLOVR (HIV RNA 400 copies/mL) at Week 24/48: Treated Population, Page 6, Section 3.2.2, Vol 57).

3.1.4.2 Efficacy of emtricitabine relative to placebo at Week 48

Since Study 035 was a 24-week study, in order to use it for FTC303 Week 48 evaluation, additional assumptions would be needed to bridge the virologic response (1-LOVR) at Week 24 and the virologic response at Week 48. Denoting the probability of LOVR-free at time t is $P(\text{FTC-Placebo})_t$, the efficacy at Week 48 can be estimated:

$$P(\text{FTC-Placebo})_{48} = \kappa\, P(\text{FTC-Placebo})_{24} \qquad \text{(Eq.2)}$$

where κ is a ratio of the two probabilities. Using the FTC303 TLOVR data, the ratio of 1-LOVR at Week 48 vs. that at Week 24 for the FTC arm is 0.59. Assuming that this ratio is similar to FTC-Placebo, then the effective size ε(FTC-Placebo) at Week 48 would be 15% to 21%.

3.1.4.3 Discussion

- The superiority of FTC over placebo in maintaining HIV RNA below assay limit could not be established without strong assumptions on the 3TC's contribution in the control regimen. The reviewer used the indinavir study 035 to derive the 3TC's contribution in efficacy in a combination regimen. Because Study 035 is small, only when we assume that more than 70% of the 3TC effect observed in Study 035 was preserved we can indirectly prove that FTC is superior to placebo at significance level 0.05.

- Since patients in FTC303 were already 3TC-experienced, it is unlikely that 3TC contribution in this population will be as large as the 3TC contribution in the 3TC-naïve

population in the Study 035. The combination regimen used in the historical trial also differed from the combination in FTC303 study. Given these differences the 70% preservation rate may be too high. The 70% preservation rate is also higher than the more commonly used 50% preservation rate. Therefore, we can not firmly establish the superiority of FTC over placebo.

- The indirect method is based on several assumptions linking the historical study with the current study.

  - Prediction of the efficacy by emtricitabine can be implemented based on the associated between LOVR using LOQ=50 copies/mL and LOQ=400 copies/mL from FTC303 data. If the plasma HIV-1 RNA information for Study 035 is available, then we may obtain the probability of LOVR based on LOQ=50 copies/mL and go over the above steps.

  - Using the FTV303 data to estimate κ in Eq. 2 was based on an assumption that $P(\text{FTC-Placebo})_{48} / P(\text{FTC-Placebo})_{24} \approx P(\text{FTC regimen})_{48} / P(\text{FTC regimen})_{24}$. One alternative approach is to predict $P(\text{3TC-Placebo})_{48}$ in Study 035 and obtain such a ratio.

  - The variance estimate of efficacy by emtricitabine relative to placebo at Week 48 is intricate. To further investigate var(κ), one might use bootstrapping or Bayesian method. One may consider inflating the variance of ε (3TC-Placebo) from Study 035 because in practice the source of variations would be more than what one understands.

- In the future, a careful study design in a triple therapy regimen should be considered in order to obtain a scientific result.

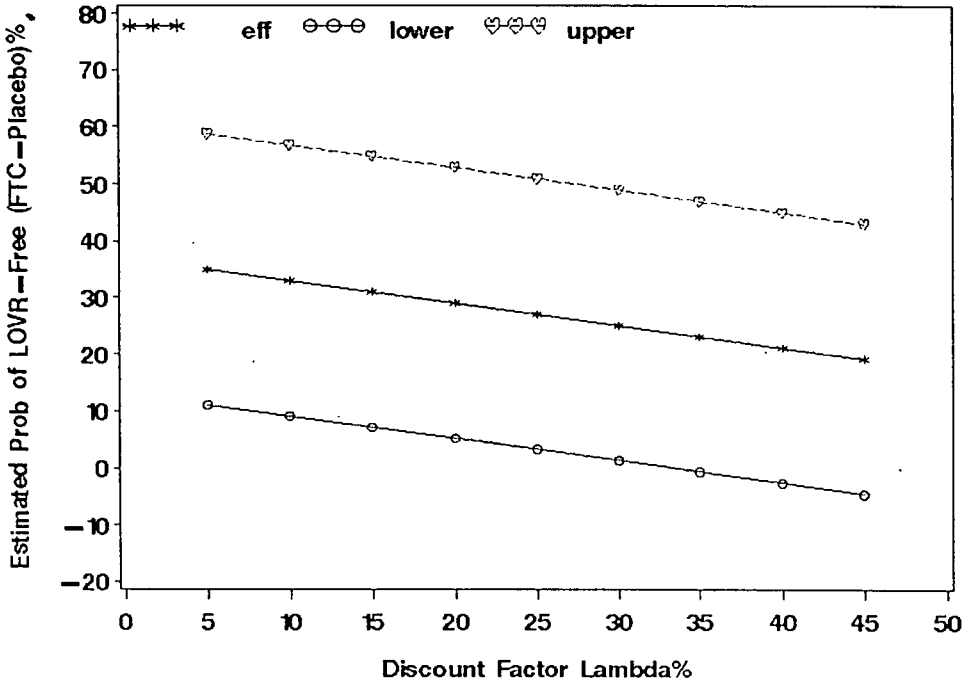

Figure 4: Study FTC-303: Estimated Probability of LOVR-free (FTC-Placebo) at Week 24 for Various Discount Factor λ

3.2 Evaluation of Efficacy (Study FTC301A)

3.2.1 Patient Disposition

Table 7 lists the disposition through Week 48. 580 subjects were randomized, 571 (97.8%) of which had received at least one dose of study medication, and comprised the Intent-to-treat (ITT) population. A total of 442 subjects (77%) were completed study therapy at Week 48. More subjects (83%) in the FTC treatment arm than those in the d4T (72%) had completed the study. More subjects in the d4T treatment arm were discontinued due to adverse events, virologic failure and loss-to-follow up than in the FTC group. In addition, 209 (73%) of the subjects in the d4T treatment arm and 234 (82%) in the FTC treatment arm had drug exposure exceeded 40 weeks.

Table 7: Study FTC301A: Subject Disposition Through Week 48

	d4T n (%)	FTC n (%)	Total n (%)
Randomized	290	294	584
Did not Return at Baseline	5 (1.7)	8 (2.7)	13 (2.2)
Treated	285 (100)	286 (100)	571 (100)
Completed study through Week 48[a]	**205 (72)**	**237 (83)**	**442 (77)**
Study Discontinuation	**80 (28)**	**49 (17)**	**129 (23)**
Adverse Events	33 (11.6)	16 (5.6)	49 (8.6)
Virologic Failure	22 (7.7)	8 (2.8)	30 (5.3)
Loss to follow-up	12 (4.2)	9 (3.1)	21 (3.7)
Investigator or subject request	4 (1.4)	5 (1.7)	9 (1.6)
Protocol violation	2 (<1)	7 (2.4)	9 (1.6)
Noncompliance	3 (1.1)	2 (<1)	5 (<1)
Other[b]	4 (1.4)	2 (<1)	6 (1.1)

a. Based on HIV RNA collection dates.
b. Including two subjects in the d4T treatment arm with no Week 48 HIV-1 RNA data.

Upon this reviewer's request, the applicant reported plasma HIV-1 RNA data status and censoring at Week 48 for the ITT population. 195 subjects (68.4%) in the d4T arm and 230 (80.4%) in the FTC arm had Week 48 HIV-1 RNA data. 19 subjects (12 in the d4T arm and 7 in the FTC arm) did not have Week 48 data. Therefore, using the HIV-1 RNA data, the number of subjects completed a 48 week study should be ≤ 207 in the d4T arm, and ≤ 237 in the FTC arm. Therefore, study completion status obtained by this reviewer and the applicant are very close.

3.2.2 Demographics and Baseline Characteristics

- Table 8 lists demographics and baseline characteristics of study subjects in Study FTC301A.

Table 8: Study FTC301A: Demographics and Baseline Characteristics

		d4T	FTC
Randomized		290	294
Treated		285	286
Age (years)	Mean (std)	37(10)	36 (9)
	Median (Range)	35 (18-67)	35 (18-69)
Gender	Male	246 (86)	239 (84)
Race	Caucasian	159 (56)	136 (48)
	Black	40 (14)	52 (18)
	Hispanic	70 (25)	77 (27)
	Other	16 (5)	21 (7)
Strata[a]	N. America-lower HIV RNA	71 (25)	72 (25)
	Europe-lower HIV RNA	31 (11)	31 (11)
	S. America-lower HIV RNA	60 (21)	56 (20)
	N. America-higher HIV RNA	51 (18)	53 (19)
	Europe- higher HIV RNA	30 (11)	31 (11)
	S. America- higher HIV RNA	32 (11)	33 (12)
	N. America-FTC301	6 (2)	6 (2)
	S. America-FTC301	4 (1)	4(1)
Region	North America	128 (45)	131 (46)
	Europe	61 (21)	62 (22)
	South America	96 (34)	93 (33)
HIV RNA in $\log_{10}$	Mean (std)	4.8 (0.7)	4.8 (0.7)
	Median (Range)	4.9 [illegible]	4.9 [illegible]
CD4+ (cells/mm^3)	Mean (std)	324 (192)	312 (192)
	Median (Range)	301 [illegible]	282 [illegible]
Weight (kg)	Mean (std)	74 (15)	73 (15)
	Median (Range)	72 (42-157)	72 (42-155)
AIDS	# (%)	54 (19)	65 (23)

[a] lower HIV RNA- between 5000 and 100,000 copies/mL; higher HIV RNA- > 100,000 copies/mL; FTC301- previously treated under FTC301-protocol.

The treated population had a mean age 36 years. 85% was male, 52% Caucasian, 26% Hispanic and 16% Black, 45% from North America, 33% from South American, and 22% from European counties. The mean and median baseline CD4+ cell counts were 318 and 288 cells/mm^3, respectively. The mean baseline HIV RNA was 4.84 in $\log_{10}$ copies/mL. The treated population included 57% subjects with screening HIV RNA between 5000 and 100,000 copies/mL, 40% with screening HIV RNA above 100,000 copies/mL and 3% previously treated under the FTC301-protocol. Prior to entry, 21% of the subjects in the treated population had history of HIV-1 related events.

The two treatment groups were well balanced with respect to demographic variables and baseline characteristics.

3.2.3 Efficacy Endpoints

- The statistical reviewer verified sponsor's fundamental findings that: (1) the virologic response was similar for the FTC study arm and equivalent d4T reference arm in Study FTC-301A; (2) the virologic response was sustained over 48 weeks; and (3) the virologic response was associated with a significant rise in CD4+ cell counts and CD4%.

3.2.3.1 Outcome

Using HIV-1 RNA data, subjects' discontinuation status, this reviewer obtained Time to Loss-of-Virologic-Response (TLOVR) outcomes. Tables 9 & 10 show outcomes by treatment arm for LOQ=400 and 50 copies/mL, respectively.

Table 9. Study 301A: Outcomes of Randomized Treatment Through Week 48 (LOQ=400)

Outcome	Tradename + ddI +EFV N=286		Stavudine + ddI + EFV n=285	
LOQ = 400 copies/mL	**n**	**%**	**n**	**%**
Completed study through Week 48	**235**	**82**	**203**	**71**
Responder[1]	**231**	**81**	**193**	**68**
Never Suppressed	2	1	2	1
Rebound	2	1	8	3
Not completed week 48 study	**51**	**18**	**82**	**29**
Discontinuation while suppressed	**16**	**6**	**31**	**11**
Death	0	0	0	0
Adverse Event	9	3	22	8
Other Reasons[3]	7	2	9	3
Virologic Failure[2]	0	0	0	0
Discontinuation after Rebound	**10**	**3**	**37**	**13**
Death	0	0	1	0
Adverse Event	3	1	12	4
Other Reasons	3	1	2	1
Virologic Failure	4	1	12	4
Discontinuation & Never Suppressed	**25**	**9**	**24**	**8**
Death	0	0	0	0
Adverse Event	7	2	2	1
Other Reasons	13	5	12	4
Virologic Failure	5	2	10	4

4 Patients maintained confirmed HIV RNA ≤ LOQ through Week 48.
5 Includes confirmed viral rebound, protocol violation.
6 Includes lost to follow-up, patient's withdrawal, non-compliance, protocol violation and other reasons.

Table 10. Study 301A: Outcomes of Randomized Treatment Through Week 48 (LOQ=50)

Outcome	Tradename + ddI +EFV N=286		Stavudine + ddI+EFV N=285	
LOQ = 50 copies/mL	**n**	**%**	**n**	**%**
Completed study through Week 48	**234**	**81.8**	**199**	**69.8**
Responder[1]	223	78.0	167	58.6
Never Suppressed	3	1.0	7	2.5
Rebound	8	2.8	25	8.8
Not completed week 48 study	**52**	**18.2**	**86**	**30.2**
Discontinuation while suppressed	**16**	**5.6**	**35**	**12.3**
Death	0	0.0	1	0.4
Adverse Event	7	2.4	26	9.1
Other Reasons[3]	9	3.1	8	2.8
Virologic Failure[2]	0	0.0	0	0.0
Discontinuation after Rebound	**2**	**0.7**	**5**	**1.8**
Death	0	0.0	0	0.0
Adverse Event	1	0.3	2	0.7
Other Reasons	0	0.0	0	0.0
Virologic Failure	1	0.3	3	1.1
Discontinuation & Never Suppressed	**34**	**11.9**	**46**	**16.1**
Death	0	0.0	0	0.0
Adverse Event	11	3.8	10	3.5
Other Reasons	16	5.6	17	6.0
Virologic Failure	7	2.4	19	6.7

1. Patients maintained confirmed HIV RNA ≤ LOQ through Week 48.
2. Includes confirmed viral rebound, protocol violation.
3. Includes lost to follow-up, patient's withdrawal, non-compliance, protocol violation and other reasons.

3.2.3.2 Time to Loss-of-Virologic-Response (TLOVR)

Using HIV-1 RNA data, subjects' discontinuation status and study drug information, this reviewer obtained time to Loss-of-Virologic-Response (TLOVR) curves by treatment arm and LOQ as shown in Figure 5. Comparing the two Kaplan-Meier curves between treatment arms, we obtained p=0.0016 and 0.0373 for LOQ=50 and 400 copies/mL, respectively, by the Log-rank test. It appears that subjects in the FTC treatment arm had significant lower LOVR or greater virologic responses than those in the d4T treatment arm.