CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 20-154/S-029, S-030
20-155/S-021
20-156/S-022

STATISTICAL REVIEW

Statistical Review and Evaluation

NDA#:	20154, 20155 and 20156 (labeling supplements)
APPLICANT:	Bristol-Myers Squibb Company
NAME OF DRUG:	Videx® (didanosine) Chewable/Dispersible Buffered Tablets
INDICATION:	Treatment of HIV infection
DOCUMENTS REVIEWED:	Submission dated 4/30/99: Vol. 1, 6-14, 26
Clinical Reviewer:	HFD-530: Russ Fleischer, PA.C, MPH

A: Background

Didanosine (ddI) is a nucleoside analogue available in tablets in strengths of 25, 50, 100 and 150 mg of didanosine. The recommended dosing interval is twice daily.

These SNDAs contain one pivotal, randomized, multicenter, controlled trial (AI454-148) to support a new 200 mg strength tablets and a changing of dosing to once-daily in combination therapy. Two Phase II studies, AI454-143 and AI454-146, were also submitted as supportive evidence. Study AI454-146 will not be reviewed here since it does not contain information on the new dosage form.

Protocols

AI454-148

Title: "A Randomized Study of the Long-Term Suppression of Plasma HIV RNA Levels by Triple Combination Regimens in Treatment Naïve Subjects"

This is a multinational open-label study in anti-retroviral naïve HIV-infected subjects 12 (18 in Europe and Canada) years of age or older whose screening plasma HIV RNA level is ≥2000 copies/mL and whose CD4 cell count is ≥100 cells/mm^3. Seven hundred subjects were to be randomized 2:1 to the following two treatment groups:

Group 1:	ddI (400mg QD) + d4T (40mg BID) + NLF (750mg TID) if weigh ≥60Kg
	ddI (250mg QD) + d4T (30mg BID) + NLF (750mg TID) if weigh <60Kg
Group 2:	ZDV (300mg BID) + 3TC (150mg BID) + NLF (750mg TID)

The randomization was to be balanced by plasma HIV RNA level (<30,000 vs. >30,000 copies/mL) and investigative site using the method proposed by [redacted] with a centralized randomization code.

Dosing was planned for a period of at least 48 weeks after enrollment of the last subject. An early analysis was planned to occur when approximately 200 subjects have been treated for 6 months after randomization.

Participation in this trial may be discontinued for the following reasons:

- Increase in viral load to detectable levels (confirmed in a repeat assay, at least one value ≥1000 copies/mL).
- Major toxicity or pregnancy or use of prohibited medication.

Plasma HIV RNA level and CD4 cell counts were to be determined at screening (twice), day 1, Week 4, Week 8 and every 8 weeks thereafter, and at the final or early termination. Subjects who discontinued the randomized treatment were to be followed in the same manner.

The primary efficacy endpoint is the proportion of patients with HIV RNA < 400 copies/mL at Week 48. For the planned early submission the Week 24 data will be used instead. The secondary analysis includes time to viral load failure. Failure is defined as

- Failure to reach viral load <400 copies/mL by Week 24, or
- Experiencing AIDS-defining events or death, or
- Discontinued the randomized treatment except for switching between ZDV and d4T, or
- Confirmed rebound to detectable levels with at least one of which is >1000 copies/mL).

For the analysis of proportions below 400 copies/mL, subjects who discontinued the randomized treatment or lost to follow-up are regarded as failures. Missing values are regarded as failures unless bracketed by two values <400 copies/mL. The analysis uses an observation window of 8 weeks. The estimates and 95% confidence interval for the difference of proportions is adjusted for the HIV RNA strata with weights inversely proportional to within strata variance. A non-inferiority limit of 12% is used.

For the analysis of time to viral load failure, plots based on Kaplan-Meier estimates and Cox proportional hazards models are used to assess the treatment differences. The time averaged difference between the two groups in change from baseline (using $\log_{10}$ scale for HIV RNA level) is analyzed using repeated measures model with a compound symmetry covariance structure for plasma HIV RNA levels and CD4 cell counts.

All treated subjects are included in the analysis.

The trial design provides at least 90% power to demonstrate the equivalence of the two arms when the common response rate is assumed to be 75% in both arms.

AI454-143

Title: "A Randomized, Double-Blind, Study of the Antiviral Activity of Once-Daily and Twice-Daily Dosing of Didanosine in Combination with Twice-Daily Dosing of Stavudine in HIV-Infected Subjects."

This is a randomized, double-blind, 12-week study designed to determine the antiviral activity and tolerability of ddI dosed once or twice daily in combination with d4T in HIV-infected patients with a CD4 cell count of $\geq$100 cells/mm^3, plasma HIV RNA $\geq$10,000 copies/mL, and naïve to prior antiretroviral therapy. At the baseline subjects are stratified based on HIV RNA distribution (<30,000 copies/mL and $\geq$30,000 copies/mL).

One-hundred subjects are equally randomized to receive ddI 400 mg (2 x 200 mg) once daily+d4T (40 mg) twice-daily or ddI 200 mg twice-daily+d4T twice daily. The dosing was planned for at least 12 weeks after enrollment of the last patient, but continued until the last patient had completed 16 weeks of treatment.

The primary objective is to demonstrate that ddI dosed once daily is as effective as a standard twice-daily regimen. The primary endpoint is a comparison of the time-averaged difference (TAD) in $\log_{10}$(HIV RNA) between the two regimens over the first 12 weeks of therapy. This is computed by using a repeated measure model assuming compound symmetry for the covariates structure. If the upper limit of the 95% confidence interval for the difference is less than 0.5 $\log_{10}$, then the conclusion is that the regimens are equivalent. Secondary endpoints includ the proportion of patients with HIV RNA <400 copies/mL (Roche Amplicor® HIV-1 Monitor assay), and changes in CD4 cell counts at week 12.

B. Results of the Applicant's Analyses

The results of the applicant's analysis are included in the reviewer's analysis section.

C. Statistical Reviewer's Comments

AI454-148

The protocol-specified analysis population is all treated subjects. This being an open-label trial, the post-randomization dropouts may not be completely at random. Therefore it may create biases that can not be eliminated statistically. In this review all subjects randomized will be included in the analyses. The applicant has provided analyses based on this population.

Aside from the inherent difficulties associated with the interpretations of an open-label equivalence trial, which include biases caused by subject's and physician's knowledge of the treatment received, and the fact that any sloppiness in conducting the trial may make the results more alike between treatment arms, the design of this trial used "double substitution" where the two treatment arms differed in two of the three drugs used. Such a design makes it impossible to evaluate the contribution of ddI in the combination treatment.

The analyses used in the submission may not be appropriate for the adaptive randomization procedure used. The adaptive randomization procedure allows probabilities of treatment assignment to change according to how well certain pre-specified baseline covariates such as clinical sites and baseline HIV RNA levels are "balanced" between treatment arms. The possible outcomes and the probabilities associated with each possible outcome from this adaptive randomization procedure are different from traditionally used randomization procedures such as a simple randomization, with or without blocking and stratification, and analyses appropriate for the simple randomization may not be correct for the adaptive randomization. Upon request, the applicant has performed permutation-based analyses for the primary endpoint and they are reported in the next section. For all other endpoints or for secondary analyses, applicant's analysis methods is used in the review for the following reasons: (1) The applicant's simulation study showed that under a specific statistical model, the analysis results from the permutation-based methods were similar to the applicant's proposed methods for the primary analysis, (2) Permutation-based methods are computationally intensive and is not practical to be performed on all analysis due to time constraints, and (3) There is no permutation-based procedure for equivalence test that is agreed upon in the scientific community.

One subject was randomized twice. It was originally randomized to ZDV/3TC/NLF as subject 00003 00690 but never started the treatment and was later re-randomized to ZDV/3TC/NLF again as subject 00003 00731. This subject was not among the first 387 subjects randomized and was not included in the efficacy analysis.

According to the protocol and its amendments, the planned interim analysis was to be conducted when roughly 200 subjects complete 24 weeks of the trial. However, in the study report "200" became "at least one-half of the subjects". Upon further communication with the applicant, it appears that this change was made in response to FDA request (Fax dated 10/8/98) and was documented in a statistical analysis plan separate from the protocol amendments. Therefore the submitted interim analysis is accepted as valid.

The protocol-specified primary endpoint TAD is regarded as secondary here. Instead, to be consistent with current clinical goal of HIV therapy, which is to suppress plasma HIV RNA levels, proportions below 400 copies/mL is regarded as the primary endpoint. The statistical power for demonstrating equivalence in proportions below 400 copies/mL is fairly low due to the small sample size.

D. Statistical Reviewer's Analyses

D.1. Study 148

Baseline Characteristics

Seven hundred fifty six subjects were randomized and 725 of them initiated the treatment. The efficacy of this submission is based on subjects who were randomized through July 31, 1998 (387) and initiated treatment (375). Of all subjects randomized, 71% are male, 26% are black, 56% are white and 14% are Hispanic. The mean age is 34.7 years old (range 17 – 70). The median baseline HIV RNA level is 4.69 $\log_{10}$ copies/mL and the median baseline CD4 is 340 cells/mm^3. 36% of the subjects have baseline HIV RNA values <30,000 copies/mL while the remaining 64% have values ≥ 30,000 copies/mL.

Subject Accountability

The following table presents the disposition of subjects.

Subject Status and Reason Discontinued by Treatment Group and Study

Treatment	ddI/d4T/NLF	ZDV/3TC/NFL
Total Randomized	503	253
No. Never Started Treatment	25	6
No. Started Treatment	478	247
Total Randomized through 7/31/98	255	132
No. Never Started Treatment	10	2
No. Started Treatment	245	130
No. discontinued randomized treatment	61	30
Lost to Follow-up	18	9
Subject Withdrew	13	10
Adverse Event	13	5
Non-compliance	7	3
Disease Progression or Relapse	6	3
Death	3	–
Pregnancy	1	–

Based on tables on pages 114 of Vol.13 and dataset submitted.

Among all 756 subjects randomized, 31 subjects (4.1%) never started the randomized treatment. This rate is higher in the ddI/d4T/NFF arm (5.0%) than in the ZDV/3TC/NLF arm (2.4%). Among the 387 subjects included for the Week 24 analysis, these rates are comparable to the overall population. Additionally, 91 of these 387 subjects (23.5%) initiated and later discontinued the randomized treatment. This rate is nearly identical for the two treatment arms.

Efficacy Endpoints

The tables below display the results for HIV RNA viral load and CD4+ counts. In the analysis of viral load, a subject with viral load below 400 copies/mL is regarded as a "Success". Subjects who discontinued the randomized treatment earlier are regarded as failures while missing values are filled with the worst of the pre and post neighboring values.

Week 24 HIV RNA Status

All Randomized Subjects through July 31, 1998

Treatment	DdI/d4T/NLF	ZDV/3TC/NFL
Total Randomized through 7/31/98	255	132
No. never started treatment	10	2
Discontinued with last RNA value before Week 24	47	26
Week 24 HIV RNA <400 copies/mL	152	73
Week 24 HIV RNA ≥400 copies/mL	41	27
Week 24 HIV RNA missing	5	4
Week 16 and 32 RNA <400 copies/mL	2	1
Week 16 or 32 RNA >400 copies/mL or missing	3	3

Reviewer's Calculation Based on Data submitted.

Proportion of Subjects with HIV RNA <400 copies/mL by Treatment

With Missing Imputed and Failures Carried Forward

All Subjects Randomized through July 31, 1998

Treatment	DdI/d4T/NLF	ZDV/3TC/NFL
Total Randomized through 7/31/98	255	132
No. of Success	154	74
Success Rate	60.6%	56.1%
Difference*	4.3%	
p-value and 95% Confidence interval*	0.41 (-6.0%, 14.7%)	

*: Baseline HIV RNA stratum adjusted difference. Weight for a stratum with n_1 and n_2 subjects in the two arms is proportional to $n_1*n_2/(n_1+n_2)$.

Based on Table S.10.1.2E on page 124 of Vol. 13.

In the table above, the estimate of the treatment difference is based on the baseline HIV RNA stratum adjusted difference and the 95% confidence interval is based on normal approximation as presented in the applicant's submission. As discussed earlier, this analysis may not be suitable for the adaptive randomization procedure used for treatment assignment. Instead a permutation-based procedure is used here. The details of the procedure are attached in the Appendix.

With the permutation-based procedure, the 95% confidence interval became (-7.5%, 16.7%), which is wider, but the lower bound is still adequate for equivalence claim.

There are 9 missing values and their impact on the analysis is small. In the worst case scenario, when all the 5 missing values in the ddI arm are regarded as failures while the 4 missing values in ZDV/3TC/NFL are regarded as successes, the success rates are 152/255=59.6% for the ddI/d4T/NFL arm and 77/132=58.3% for the ZDV/3TC/NFL arm. The estimated treatment difference becomes 1.3% and the 95% confidence interval becomes (-10.5%, 13.7%) using permutation-based procedure. Overall, the proportion below 400 copies/mL at Week 24 in

ddI/d4T/NFL arm is very likely to be no more than 11% worse than the ZDV/3TC/NFL arm when early discontinuations and lost to follow-ups are regarded as above 400 copies/mL.

At Week 24, the median CD4 cell count change for ddI/d4T/NFL is 164 and for ZDV/3TC/NFL it is 172. The estimated time-averaged difference of CD4 change over 24 weeks for the two arms is 11.5 in favor of ZDV/3TC/NFL and is not statistically significantly different (p=0.42, using baseline HIV RNA-stratified CMH test).

Overall, the trial demonstrated the equivalence of the two treatments in achieving HIV-1 RNA <400 copies/mL at Week 24. The CD4 changes were similar between the two treatments.

Subgroup Analysis

The reviewer conducted the subgroup analyses by age, gender, race and baseline HIV RNA level (<30,000 vs. ≥30,000 copies/mL) for the proportions below 400 copies/mL at Week 24. The homogeneity of treatment differences among the treatment groups across stratums defined by these factors is tested using the Breslow-Day test. Age (≤35 vs. >35 years old), gender and race appear to have no interaction with the treatment (p-values > 0.2). For the baseline HIV RNA this interaction is statistically significant (p-value=0.031). The table below summarizes the proportions in the subgroups.

Proportion <400 copies/mL at Week 24 by Baseline HIV RNA

	Baseline HIV RNA <30,000 copies/mL	Baseline HIV RNA ≥30,000 copies/mL
ddI/d4T/NFL	55/82=67.1%	99/173=57.2%
ZDV/3TC/NFL	21/45=46.7%	53/87=60.9%
Difference	20.4%	-3.7%

The statistical test of the difference between the two treatment groups in baseline HIV RNA <30,000 copies/mL stratum (67.1% vs. 46.7%) yielded a p-value of 0.03. This suggests that ddI/d4T/NFL may be more efficacious than ZDV/3TC/NFL for subjects with baseline HIV RNA <30,000 copies/mL. However, the response pattern of ZDV/3TC/NFL in the two stratums is atypical in that contrary to most clinical trials, subjects with baseline HIV RNA ≥30,000 copies/mL have markedly a higher response rate than those with baseline HIV RNA <30,000 copies/mL. In addition, the response rates between the two stratums are quite similar for both treatment arms at Week 16 and the p-value for the interaction between HIV RNA stratum and the treatment is 0.556. Further investigation at Week 48 and additional studies are required to clarify this issue.

D.2. Study 143

Baseline Characteristics

Eighty-eight subjects were randomized and one ddI BID subject received no treatment. Of all subjects randomized and have at least one dose of study medication, 77% are male, 48% are black, 44% are white and 8% are Hispanic. The mean age is 35.8 years old (range 18 – 54). The median baseline HIV RNA level is 4.65 $\log_{10}$ copies/mL and the median baseline CD4 is 410 cells/mm^3. Thirty-four percent of the subjects have baseline HIV RNA values <30,000 copies/mL while the remaining 66% have values ≥ 30,000 copies/mL.

Subject Accountability

The following table presents the disposition of subjects.

Subject Status and Reason Discontinued by Treatment Group and Study

Treatment	ddI(QD)/d4T			ddI(BID)/d4T		
Total Randomized	44			44		
No. Never Started Treatment	0			1		
No. Started Treatment	44			43		
No. discontinued randomized treatment	Before Week 12	After Week 12	Total	Before Week 12	After Week 12	Total
All discontinued	9	10	19	7	14	21
Lost to Follow-up	1	1	2	3	4	7
Subject Withdrew	3	3	6	0	3	3
Adverse Event	1	1	2	2	4	6
Non-compliance	2	0	2	1	1	2
Disease Progression or Relapse	0	1	1	0	0	0
Other	2	4	6	1	2	3

Based on tables on pages 57, 67 and 103 of Vol.6.

The two arms are similar in number discontinued and reason for discontinuations.

Efficacy Endpoints

The tables below display the results for HIV RNA viral load and CD4+ counts. In the analysis of viral load, a subject with viral load below 400 copies/mL is regarded as a "Success". Subjects who discontinued the randomized treatment earlier are regarded as failures while missing values are filled with the worst of the pre and post neighboring values.

Proportion of Subjects with HIV RNA <400 copies/mL by Treatment
With Missing Imputed and Failures Carried Forward
All Subjects Randomized

Treatment	ddI(QD)/d4T	ddI(BID)/d4T
Total Randomized	44	44
No. of Success	18	14
Success Rate	40.9%	38.6%
Difference	2.3%	
95% Confidence interval	(-18%, 23%)	

Based on Table S.10.2.1 on page 122 of Vol. 6.

The lower bound of the 95% confidence interval is too low to reach equivalence conclusions.

F. Overall Assessment

Based on the available data up to Week 24 cutoff date, Study 148 demonstrated that ddI+d4T+NLF is likely to be no more than 11% worse than ZDV+3TC+NLF in achieving viral load < 400 copies/mL at Week 24. The CD4 changes over time are similar for the two groups. However, the double substitution in the design does not allow valid inference on the contribution of ddI.

In Study 143, proportions below 400 copies/mL at Week 12 were numerically similar between ddI(QD)/d4T and ddI(BID)/d4T. However, the lower bound of the 95% confidence interval for the treatment difference is too low and the treatment period is too short for any definitive conclusions on the comparison between the two treatment arms.

Greg Soon, Ph.D.
Mathematical Statistician

Concur: Girish Aras, Ph.D.

cc:
Archival IND 21007
HFD-530
HFD-530/HJolson (via TeamLinks)
HFD-530/TCvetkovich
HFD-530/RFleischer
HFD-530/DSillivan
HFD-725/GAras
HFD-725/GSoon
HFD-725/MHuque

GS/October 25, 1999/MSWord7.0/D:\GREGSOON\NDA\Bristol-Myers Squibb\20154\Review.doc
This review contains 12 pages.

Appendix: Calculating confidence interval base on a re-randomization test

An ad hoc 95% confidence interval for the difference of the proportions with HIV RNA <400 copies/mL in Study 408 can be generated with the following procedure. This procedure is based on an inversion of the re-randomization test.

1. For each arm, classify subjects into three groups: Group I consisting of subjects who were on the original randomized therapy at Week 24 with Week 24 HIV RNA <400 copies/mL, Group II consisting of subjects who were on the original randomized therapy at Week 24 with Week 24 HIV RNA >400 copies/mL, and Group III consisting of subjects who discontinued the randomized therapy or lost to follow-ups by Week 24. For subjects who were on the original randomized therapy at Week 24 but with missing Week 24 HIV RNA values, the Week 24 HIV RNA values can be imputed using the worst of pre and post HIV RNA values.
2. Find the largest l in the interval (0,1) such that for any number p in the interval (0, l) the two-sided p-value based on the following procedure is <0.05.
   2.1. Let m= p x {# of subjects in the experimental arm}. Re-classify m subjects from Group II with the smallest HIV RNA values in the experimental arm as "successes". For all other subjects only those in Group I will be regarded as "successes".
   2.2. Calculate the two-sided p-value based on this new classification using the re-randomization test with the original randomization procedure.
3. Similarly, find the largest u in the interval (0, 1) such that for any number p in the interval (0, u) the p-value based on the following procedure is <0.05.
   3.1. Let m= p x {# of subjects in the control arm}. Re-classify m subjects from Group II with the smallest HIV RNA values in the control arm as "successes". For all other subjects only those in Group I will be regarded as "successes".
   3.2. Calculate the p-value based on this new classification using the re-randomization test with the original randomization procedure.
4. Report the 95% confidence interval as ($-l$,u).