

addition, a baseline time-effect curve was constructed using data from all subjects at 3 sites ("12-hour PFTs"). These subjects had additional PFTs checked at 20 min., 40 min., 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours after study drug administration. Baseline samples for FP and cortisol levels were drawn at 20 min. and 40 min. after study drug from all subjects. A subset of subjects (again, all subjects at 3 sites) had additional time points checked: 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours after study drug.

*Reviewer's Comment: The three sites for PK/PD determination and the 3 sites for time-effect curve assessment were the same.*

Eligible subjects needed to meet additional criteria at each clinic visit to continue in the study. "Stability limits" were therefore defined at Visit 2 for PEF<sub>R</sub> and FEV<sub>1</sub>:

- FEV<sub>1</sub> stability limit: 20% decrease from the best FEV<sub>1</sub> at Visit 2
- PEF<sub>R</sub> stability limit: 20% decrease from mean diary AM PEF<sub>R</sub> from the past 7 days

Subjects not meeting the following "continuation criteria" at each clinic visit (Visit 3 and beyond) were discontinued for lack of efficacy:

- No more than 2 days in the last 7 in which  $\geq 12$  puffs of Ventolin MDI were used
- No more than 3 days in the last 7 where the AM or PM PEF<sub>R</sub> was below the PEF<sub>R</sub> stability limit
- No more than 2 nights in the last 7 with awakenings requiring Ventolin.
- A clinic FEV<sub>1</sub>  $\geq$  the FEV<sub>1</sub> stability limit

Visits were scheduled weekly for the first 4 weeks, then every other week until study endpoint at 12-weeks. At Visits 3-9 the following procedures were performed:

- Assess subject's compliance including withholding medication (required for PFTs and other procedures to be performed)
- Assess subject's "continuation criteria" (must be met or patient was terminated for lack of efficacy)
- Review previous diary cards and dispense new cards
- Adverse event assessment especially acute asthma exacerbation
- PFTs. For additional subjects at 3 sites, "12-hr PFTs" were also performed. Readings were made predose and at 20 min., 40 min., 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours after study drug administration (Visits 3, 6, and 10)
- Collect/dispense study medication (Visits 4, 6, 7, 8, and 9: every 2 weeks)
- Oropharyngeal exam (Visits 6, 8, and 10)
- Physician global assessment (Visits 7 and 10)
- Clinical laboratory tests/plasma cortisol: (Visits 3, 6 and 10)

- Plasma samples for FP: Samples for FP and cortisol levels were drawn pre-dose and again at 20 min. and 40 min. after study drug from all subjects. A subset of subjects had additional time points checked: 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours after study drug (Visits 3, 6, and 10)

At study endpoint (Visit 10) or early termination, the usual scheduled clinic assessments were made, in addition to the same as performed at baseline (physical exam, etc.), and the special assessments summarized in the bullet points above. Study devices were collected, and overall compliance with study procedures was assessed by blister counts, completion of diary cards, and whether subject followed instructions to withhold medication on the morning of the clinic visit.

#### 4.3.1.6.5 Efficacy Assessments

The primary efficacy variable was AM pre-dose FEV<sub>1</sub>. FEV<sub>1</sub> was performed in triplicate using approved spirometric equipment according to ATS recommendations. The subject could be sitting or standing during the maneuver, but was required to be consistent throughout the study. If two FEV<sub>1</sub> readings were identical, the once with the highest FVC was utilized.

Secondary efficacy variables included all of the following:

- "Survival" in the study
- Physician-rated global assessment of efficacy (0=ineffective, 1=satisfactory, 2=effective, and 3=very effective)
- Pre-dose FVC and FEF<sub>25-75</sub>
- "12-hour PFTs"  
(Subjects who fell below their FEV<sub>1</sub> stability limit during this 12-hour period were treated with Ventolin, but not discontinued from the study. Further PFT assessments were not performed, but PK/PD samples continued to be drawn at the designated time points)
- Diary AM and PM PEFR  
(Using a \_\_\_\_\_ peak flow meter, AM before study medication and PM after study medication. The highest of three values was recorded. The AM/PM PEFR difference was also assessed as a secondary endpoint)
- Subject-rated daily symptom scores on a scale of 0 (none), 1 (mild), 2 (moderate), or 3 (continuous or disabling)
- Number of nighttime awakenings requiring Ventolin
- Rescue Ventolin use

#### 4.3.1.6.6 Safety Assessments

- Clinical Adverse Events (AE)
- Clinically significant changes in clinical laboratory values

- Clinically significant changes in physical examination, oropharyngeal exam, vital signs, or 12-lead ECG
- HPA-axis effects via basal AM cortisol
- Timed plasma concentrations of FP and cortisol

#### 4.3.1.6.7 Statistical Methods

*General Statements:* All statistical testing was two-sided. Treatment differences at or below the 0.05 level were considered significant. Pair-wise comparisons were performed without adjusting p-values for the number of comparisons made and pair-wise p-values were interpreted only when the overall test among treatment groups was statistically significant.

*Power Calculations:* Mean and standard deviation of the primary endpoint was estimated based on prior studies conducted by the sponsor. Enrollment was planned to obtain 240 evaluable (75-80 per arm) subjects to provide >80% power of detecting a difference in FEV<sub>1</sub> of 0.25L between any two treatment groups, using a t-test with a significance level of 0.05. The proposed sample size would also provide >80% power to detect a difference in AEs of 16% between any two treatment arms.

*Populations:* The Intent-to-Treat (ITT) Population was used for most calculations, unless otherwise stated. The ITT Population included any subject who had received at least one dose of study medication. The Efficacy Population was a subgroup that included only those subjects who had no major protocol violations during the study. The decision to exclude a subject from the Efficacy Population was to have been made prior to breaking the blind.

There was one center excluded from the efficacy analysis because the data was believed to be unreliable. A total of sixteen subjects were involved, and their data were analyzed for safety only, separately from the remainder of the ITT.

*Background Characteristics:* Comparisons between treatment groups were based on ANOVA F-test controlling for investigator for age, height, and weight, and on the Cochran-Mantel-Haenszel test controlling for investigator for gender, smoking history, method of contraception and ethnic origin.

*Efficacy:* The primary efficacy parameter was AM pre-dose FEV<sub>1</sub> in the ITT population. Testing for the primary and for most (continuous) secondary efficacy parameters was first performed on data from all investigators combined, assessing investigator and treatment-by-investigator interactions at a significance level of 0.10. An ANOVA F-test was used to compare change-from-baseline for each of the time-dependent variables at endpoint (or at other selected time points).

Endpoint was the last recorded value for the ITT population and the last evaluable value for the efficacy population.

Withdrawals from the study due to lack of efficacy were evaluated using Kaplan-Meier estimates of survival, and overall and pairwise treatment comparisons were based on the Log-rank test.

As stated above, continuous parameters such as PEFR measurements were tested with an ANOVA F-test controlling for investigator. Tests were performed on mean values over days within individual weeks. Parameters having discrete values such as symptom scores were analyzed using the non-parametric van Elteren test based on 7-day subject averages.

*Reviewer's Comment: The mean number of diary entries required in a given week before the data were considered sufficient for analysis was not stated.*

The 12-hour PFTs performed at 3 sites lacked sufficient numbers for valid statistical testing. Data were provided as descriptive only in supporting tables.

*Safety:* All safety assessments were based on the ITT population, including the dropped site, which was analyzed separately. Adverse events were tabulated by organ system, treatment group, severity, and relation to study drug. Laboratory variables, ECG, VS, and physical exam were reported by presence and/or direction of change and whether or not abnormal. AM plasma cortisol results were tabulated by treatment group based on an abnormality, defined as any basal (un-stimulated) reading <5 mcg/dL. No statistical tests were specified.

*Pharmacokinetics:* The ITT population provided FP plasma levels pre-dose and 20 min. and 40 min. after study medication for three time-points, baseline (1<sup>st</sup> dose), week 4, and week 12. Data consisted of FP plasma levels vs. time (when detectable) and  $C_{max}$  (calculated). The PK population from the 3 selected sites provided plasma FP and cortisol levels over 12 hours. Descriptive statistics were calculated for each of the derived PK parameters for each active treatment. The geometric  $LS_{mean}$  and 95% CI were regarded as the primary summary statistics for log transformed parameters. Median and range were the primary summary statistics for  $t_{max}$ ,  $C_{tau}$ , and other non-log transformed parameters. The derived PK parameters were analyzed for influence by gender, device and device-by-visit, FP accumulation with repeated dosing, and interaction between FP and terfenadine.

#### 4.3.1.7 Results

##### 4.3.1.7.1 Disposition

A total of 313 subjects were screened at 15 sites and entered into the preliminary 2-week baseline period. There were 100 withdrawals, most

due to failure to meet randomization criteria (52%), for a total of 213 eligible subjects. Other reasons for ineligibility included lack of reproducible lung function (28%), FEV<sub>1</sub> <50% or >80% predicted on Visit 1 or 2 (25%), and adverse event (3%). Subject distribution by site ranged from four (— 2%) to 31 (— 15%), with a mean of 14 patients/center and a median of 16 patients/center.

The 213 subjects who completed the screening period were randomized and entered into the double-blind treatment phase of the trial, 70 into placebo, 64 into FP 500 mcg BID DK, and 70 into FP 500 mcg BID DH. Fifty-eight (27%) of these 213 subjects discontinued prior to study endpoint, 53% in the placebo group, 16% in the DK group, and 14% in the DH group. The reason(s) for discontinuation are given by the table below, the most common being lack of efficacy by pre-defined criteria (15% overall). Adverse events accounted for only three (1%) of the total study discontinuations. The category "other" included failure to return, noncompliance, and prohibited medication.

#### SUBJECT DISPOSITION\*

	Placebo	DK FP500 BID	DH FP500 BID	Total
<b>Enrolled</b>	70	64	79	213
<b>Completed</b>	33 (47%)	54 (84%)	68 (86%)	155 (73%)
<b>Withdrawn</b>	37 (53%)	10 (16%)	11 (14%)	58 (27%)
<b>Lack of Efficacy</b>	25 (36%)	3 (5%)	5 (6%)	33 (15%)
<b>Adverse Event</b>	2 (3%)	1 (2%)	0 (0%)	3 (1%)
<b>Other</b>	10 (14%)	6 (10%)	6 (7%)	22 (10%)

\* From Volume 33, Table 2, p.99

#### 4.3.1.7.2 Demographics and Other Baseline Characteristics:

Treatment groups were demographically similar. About 55% were male, although adult asthmatics in this country are more likely to be female. The mean age was just under 33 years with a range from 12 to 76 years. As a group, subjects were predominantly Caucasian (81%) with African American and Latino comprising 6% and 10% overall, respectively. Most had never smoked (81%).

Asthma histories were also similar. About half of each group's reported durations of asthma were in excess of 15 years. Newly diagnosed asthmatics (duration ≤ 1 year) comprised <1% of the total enrollees. Surprisingly, only 42% of the group used inhaled CS at baseline. Eighty-five percent (85%) reported no ER visits and 99% reported no hospitalizations in the prior 12 months. Mean FEV<sub>1</sub> values were about 66% of predicted at baseline and comparable across treatment groups.

The comparability of orally inhaled corticosteroid (ICT) use at baseline across groups reflects stratification by this variable. Approximately 60%

of these subjects used TAA while 40% BDP. Prednisone or methylprednisolone were used by four subjects in the placebo group, one in the DK group, and two in the DH group (Table 7; Vol.33; p.105). Concurrent non-asthma medications and related medical conditions were not appreciably different between the three groups (Tables 8-10; Vol.33; p.106-110), with allergic or atopic disorders heading the list.

#### BACKGROUND CHARACTERISTICS\*

	Placebo	FP 500 BID DK	FP 500 BID DH	Total
<b>Number</b>	70	64	79	213
<b>Gender:</b>				
<b>Female</b>	46%	44%	46%	96 (45%)
<b>Male</b>	54%	56%	44%	117 (55%)
<b>Ethnicity: number</b>				
<b>Black</b>	3	4	6	13 (6%)
<b>Latino</b>	7	10	5	22 (10%)
<b>Caucasian</b>	59	50	64	173 (81%)
<b>Other</b>	1	0	4	5 (2%)
<b>Age (years):</b>				
<b>Mean (range)</b>	31.7 (13-73)	32.4 (13-62)	33.6 (12-76)	32.6 (12-76)
<b>Smoking history</b>				
<b>Never smoked</b>	81%	81%	81%	81%
<b>Former smoker</b>	19%	19%	19%	19%
<b>Inhaled CS use</b>				
<b>Yes</b>	43%	38%	44%	42%
<b>No</b>	57%	63%	56%	58%
<b>≥3 ER visits in prior 12 mos. (%)</b>	3%	2%	1%	2%
<b>Percent Predicted FEV<sub>1</sub> at Baseline (SE)</b>	66.83% (0.98)	65.18% (1.06)	67.34% (0.95)	

\* From Tables 4, 5 and 6; vol.33, pp.101-104

#### 4.3.1.7.3 Efficacy Analysis

##### *Populations and Compliance*

The population analyzed included all 203 subjects who received at least one dose of study medication (the ITT population). A subset analysis was performed using the 206-subject "efficacy population," comprised of the ITT subjects minus 7 subjects totally excluded because of major protocol violations. Ten additional subjects were partially excluded from the 206-patient efficacy subset because of protocol violations that occurred during the double-blind phase of the trial. This review will only consider the ITT population in the efficacy analysis.

The study drug compliance rate for both devices was determined for Visits 2, 4, 6, 7, 8, and 9 based on blister count, where compliance was defined as use of ≥70% of all doses. Based on these criteria, the mean compliance rate for both devices in each treatment group was approximately 94% or greater.

4.3.1.7.3.1 Primary Efficacy Variable: FEV<sub>1</sub>

Mean AM pre-dose FEV<sub>1</sub> was calculated for each treatment group at baseline and compared to mean AM pre-dose FEV<sub>1</sub> for each at end-point. Comparisons were made as mean FEV<sub>1</sub>, mean absolute change in FEV<sub>1</sub>, percent change in FEV<sub>1</sub>, and change in percent predicted FEV<sub>1</sub>. An F-test for overall treatment effect was performed prior to any pair-wise statistical comparisons. The last-value-carried-forward principle was used to calculate endpoint FEV<sub>1</sub> for each treatment group, to avoid bias introduced by the dropout of "sicker" patients, especially among the placebo subjects.

The results of this analysis are shown in the table below and in the attached Figures 2 and 3 (p.119; Vol.33). There was no significant difference in FEV<sub>1</sub> at baseline across treatment groups, which was 2.41L for placebo, 2.41L for the Diskus (FP 500 mcg BID DK), and 2.49L for the Diskhaler (FP 500 mcg BID DH). At endpoint, there was a statistically significant improvement in FEV<sub>1</sub> in each FP treatment group, 0.52L for the Diskus and 0.41L for the Diskhaler, compared to placebo, 0.03L (p<0.001). The significant difference could be demonstrated whether the difference was calculated as "liters," as "change from baseline in % predicted", or as "% change from baseline." There was no difference in the pair-wise comparison between the two FP groups at endpoint, a finding that was also independent of the way the difference in FEV<sub>1</sub> was calculated. Not included in the table below is the standard error for the change from baseline to endpoint in FEV<sub>1</sub>, which was 0.07 for placebo and 0.05 for both FP groups at endpoint.

MEAN CHANGE FROM BASELINE IN FEV<sub>1</sub> (L): ITT\*

	Placebo	FP 500 BID DK	FP 500 BID DH	p-value vs. placebo DK DH
N	69	63	79	
Baseline FEV <sub>1</sub> (L)	2.41	2.41	2.49	
% Predicted	66.8%	65.2%	67.3%	0.922**
FEV <sub>1</sub> : Mean change at Endpoint (L)	0.03	0.52	0.41	<0.001 <0.001
FEV <sub>1</sub> : % change at Endpoint	2.09	22.73	16.98	<0.001 <0.001
FEV <sub>1</sub> : Mean change in % Predicted	1.1	14.6	11.2	<0.001 <0.001

\* Intent-to-Treat Population; From Tables 11-15; vol.33

\*\*Overall (F-test)

The visit-by-visit change from baseline in FEV<sub>1</sub> is also shown in Tables 12-15. The overall treatment effect was statistically significant by Visit 3 (p<0.001), the first visit after study medication was begun, and was sustained throughout the 12-week trial. The pair-wise comparison

between each of the FP arms and placebo (placebo vs. DK and placebo vs. DH) was also significant by Visit 3 ( $p=0.001$  for both). There was no difference between DK and DH at any time-point, however.

#### 4.3.1.7.3.2 Secondary Endpoint: 12-hour FEV<sub>1</sub>

The 12-hour FEV<sub>1</sub> data is reported in Tables ST-11 through ST-16 (Vol.33). A subset of the ITT population, 40 patients from 3 centers, were tested during three non-consecutive day-long clinic visits, at baseline, after one week of study medication, and after 4 weeks of study medication. The 12-hour, serial post-dose FEV<sub>1</sub> determinations were performed in parallel with PK/PD sampling at the following time-points: pre-dose, 20 min., 40 min., and 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours. Other than study medication administered at time=0, patients were not eligible to receive any other concomitant anti-asthma medication during this interval, including Ventolin or theophylline. For this reason, discontinuations for failure to meet the pre-specified FEV<sub>1</sub> stability criteria was a common occurrence, and the spirometry time-effect curves were calculated using data from the relatively few participants who were able to complete all three of the 12-hour testing sessions, 8 to 14 of the original 40 subjects.

Results are not particularly surprising or illuminating. After the first dose of FP, there was some numeric improvement in FEV<sub>1</sub> during the 12-hour testing span, however, it did not differ between the placebo and the two FP arms. After one week of study drug treatment, baseline FEV<sub>1</sub> was greater for both FP groups compared to placebo, but during the 12 hours of testing, the two FP FEV<sub>1</sub>-time curves remained relatively flat. Similar results were reported for week 4. For both week 1 and week 4, the placebo FEV<sub>1</sub>-time curve started at about the same point as at baseline, but was down-sloping. A comparison between the two FP arms did not disclose any remarkable difference between the DK and DH devices, whether at baseline, week 1, or at week 4. No statistical testing was performed, however, presumably because of the small numbers of participants.

#### 4.3.1.7.3.3 Secondary Endpoint: Physician's global assessment

Physicians were asked to quantify their impression of the study treatments' efficacy for each of their subjects using the 4-point scale described earlier. The "physician's global assessment" was obtained during the 2-week baseline period, at the midpoint in the study (6 weeks), and at the final clinic visit (12 weeks). An endpoint value for this parameter was also reported, using LVCF data.

Data are reported in Table 17 and Table ST-17 (Vol.33). Baseline period efficacy did not differ between the three groups ( $p=0.732$ ), with only 8 to 13% of subjects rated as having "very effective" treatment. At the midpoint of the study, 13% of placebo patients were rated as receiving

“very effective” therapy, compared to 53% of the Diskus patients and 44% of the Diskhaler patients ( $p < 0.001$ ;  $p = \text{NS}$  for DK vs. DH). Data appeared nearly identical for the 12-week and endpoint estimates.

#### 4.3.1.7.3.4 Secondary Endpoint: FEF<sub>25-75%</sub> and FVC

FEF<sub>25-75%</sub> and FVC results were similar to FEV<sub>1</sub> and have been displayed in Tables 18-21 (Vol.33; pp.117-120). Baseline FEF<sub>25-75%</sub> did not differ between treatment groups, 1.77-1.90 L/sec. Improvement was significant by the end of the first week of treatment compared to placebo ( $p = 0.001$  overall treatment effect) and remained significant at all time points measured, including endpoint. At endpoint, there was a greater numerical improvement for the Diskus compared to the Diskhaler (0.83 L/sec vs. 0.62 L/sec), but this difference was not statistically significant ( $p = 0.090$ ). Also, the pair-wise comparison between the DK and placebo was significant at all time points, but this was not true for the DH at the Visit 9 and Visit 10 time points. The pair-wise comparison to placebo was significant at endpoint for each device, however ( $p < 0.001$  for both).

The data for FVC was not substantially different from the FEF<sub>25-75%</sub>.

#### 4.3.1.7.3.5 Secondary Endpoint: Survival in Study

There was a significant overall treatment effect on duration of study participation using the Log-rank test on Kaplan-Meier estimates of survival ( $p = 0.001$ ; see attached Figure 4; p.95; Vol.33). By the end of the study, 25 subjects (36%) in the placebo group had discontinued for lack of efficacy compared to 3 (5%) in the Diskus group and 5 (6%) in the Diskhaler group. Pair-wise comparisons of survival-in-study between placebo and each of the two FP arms were statistically significant ( $p = 0.001$  for each comparison). There was no significant difference in survival between the two FP arms.

#### 4.3.1.7.3.6 Secondary Endpoint: Diary PEFR

Mean AM PEFR, PM PEFR, and AM/PM PEFR differential were averaged weekly from diary card records of PEFR measured by subjects twice daily: before the AM dose of study medication and again after the PM dose. These data are shown in Tables ST-18- ST-20 (Vol.33; pp.253-5). The change from baseline was calculated for each of these three variables at all post-randomization clinic visits and at endpoint (Tables 22-24; Vol.33).

*Reviewer's Comment: It is unclear how many of seven possible AM PEFR diary entries needed to be recorded during a given week for the data to be considered "evaluable." Likewise for PM PEFR.*

Baseline AM PEFRs were similar across treatment groups at baseline, 421-436 L/min (see table, below). There was a statistically significant

treatment effect for FP compared to placebo starting at the first post-baseline clinic visit (week 1;  $p=0.003$ ) that was observed at all subsequent clinic visits, as well as at study endpoint ( $p<0.001$ ). Pair-wise treatment comparisons between placebo and each of the two FP groups were also significant at Week 1 ( $p=0.001$  for DK vs. placebo;  $p=0.005$  for DH vs. placebo), a finding that was also sustained for both devices at all subsequent measurements, as well as at study endpoint ( $p<0.001$  for each comparison). The improvement from baseline was numerically greater for the Diskus group (48 L/min) than for the Diskhaler group (27 L/min), and the pair-wise comparison between the Diskus and the Diskhaler statistically favored the Diskus ( $p=0.016$ ). The absolute change for the placebo group for this parameter was -14 L/min.

The mean change from baseline in diary PM PEFR followed a pattern similar to diary AM PEFR (see table below). Baseline values were comparable between treatment groups and slightly higher than AM PEFR values. Net improvement over time was more modest than for AM PEFR, and no overall treatment effect was seen until the Week-2 clinic visit. In addition to being numerically smaller, the difference in change from baseline to endpoint between the two devices was smaller and not statistically significant (DK, 26 L/min; DH, 21 L/min;  $p=0.170$ ). The pair-wise comparisons of the two FP devices with placebo each showed significance at Week 2 (DK,  $p=0.005$ ; DH,  $p=0.018$ ). Significance was sustained for the subsequent 10 time points for the DK group, but only for 9/10 time points for the DH group.

The AM/PM PEFR differential was calculated by subtracting each AM PEFR from the previous evening's PM PEFR. Large differences between the two readings are generally taken as indicative of poor asthma control. The AM/PM differential declined for both FP treatment groups during this trial, although the difference was greater for the DK group (as would be expected, because this parameter is not independent but is calculated from the prior two measurements). In the pair-wise comparison between the two devices for this parameter, the difference was statistically significant at endpoint. A significant overall treatment effect was detected at the end of the first week of study medication. The pair-wise comparisons between placebo and each of the two FP arms were also significant at endpoint.

**APPEARS THIS WAY  
ON ORIGINAL**

**CHANGE FROM BASELINE TO ENDPOINT IN AM/PM PEFR\***

	Placebo	FP 500 BID DK	FP 500 BID DH	DK vs. DH (p-value)
N	68	62	77	
Baseline AM PEFR (L/min)	428	431	436	
ΔAM PEFR p-value**	-14	48 <0.001	27 <0.001	0.016
N	66	63	77	
Baseline PM PEFR (L/min)	451	461	460	
ΔPM PEFR p-value**	-8	26 <0.001	21 <0.001	0.170
N	68	62	77	
Baseline AM/PM PEFR Differential	23	28	24	
ΔAM/PM PEFR p-value**	6	-20 <0.001	-6 0.012	0.025

\* Tables 21-24; Vol. 33. ITT population

\*\*At endpoint; comparison vs. placebo for DK and DH; final column comparison DK vs. DH

**4.3.1.7.3.7 Secondary Endpoints: Symptom Scores, Nighttime Awakenings, and Rescue Ventolin Use**

Subjects recorded their asthma-related symptoms daily on their diary cards using a 0-3 severity scale, as described earlier in this review. Using this scale, symptoms were similar and relatively mild at baseline across treatment groups, all being <1.00. At endpoint, there was a statistically significant treatment effect for FP compared to placebo. The pair-wise comparison of each device with placebo was also significant at endpoint, although the Diskus showed numerical superiority to the Diskhaler. There was no significant difference between the two devices (p=0.207).

Nighttime awakenings requiring Ventolin were also infrequent and similar across treatment groups at baseline, ranging from approximately one night in ten for the placebo group (0.09) to one night in twenty for the DK group (0.05). A statistically significant treatment effect could be found at endpoint compared to baseline, however, the pairwise comparison with placebo for each device was significant only for the Diskhaler, not for the Diskus (see table below). There was no difference between the two devices on this parameter at endpoint (p=0.619).

Use of rescue Ventolin was to be recorded daily in the diary as number of puffs of the MDI used. At baseline, daily use of Ventolin was similar between treatment groups, approximately 3 ½ puffs per day. There was a statistically significant treatment effect for this parameter beginning at Week 1 and continuing for all subsequent weeks until the completion of the trial, including endpoint. The pair-wise comparison between placebo and each of the two FP groups was also significant throughout the trial beginning at Week 1. At endpoint, both the Diskus and the Diskhaler

groups had reduced their Ventolin use by approximately 1 ½ puffs per day (DK, -1.59; DH -1.31;  $p < 0.001$  for both compared to placebo). The placebo group, on the other hand, increased its daily Ventolin use by slightly greater than ½ puff per day (+0.66). Although the improvement for the Diskhaler group was numerically smaller, there was no statistical difference between DK and DH at endpoint or at any prior week during the study (see table below and Table 29; Vol.33).

#### CHANGE FROM BASELINE IN DIARY VARIABLES (ITT)\*

	Placebo	FP 500 BID DK	FP 500 BID DH
N	67	63	78
Asthma symptom score:			
Baseline	0.59	0.53	0.50
Change	0.02	-0.20	-0.11
p-value**		0.002	0.028
Nighttime Awakenings:			
Baseline	0.90	0.05	0.07
Change	0.80	-0.03	0.00
p-value**		0.069	0.008
Ventolin use (puffs/day)			
Baseline	3.45	3.45	3.27
Change	0.66	-1.59	-1.31
p-value**		<0.001	<0.001

\* From Tables 28-30; Vol.33.

\*\* Compared to placebo

#### 4.3.1.7.3.8 Efficacy by Demographic Subgroups

There was no indication that a difference in response to FP existed by gender subgroup on the primary endpoint. This appeared to be true by ethnic subgroup, as well, however number of non-Caucasian subjects was very small (see "Results: Demographics; also Tables ST 27-34; Vol.33).

The subgroup analysis by age is summarized below. Although there were relatively few adolescents studied during this protocol, there is a suggestion that the Diskhaler may have greater efficacy among these younger patients than the Diskus, a finding corroborated by the pivotal pediatric trial FLTA2006.

#### Mean Change from Baseline in FEV<sub>1</sub> by Age

Treatment Group	Adolescent 12-17 yrs			Adult 18-64 yrs			"Elderly" >64 years		
	PL	DK	DH	PL	DK	DH	PL	DK	DH
N	13	10	7	56	54	70	1	2	
FEV <sub>1</sub> : Change from Baseline (L)	0.33	0.67	0.85	-0.03	0.50	0.37	-0.15		0.31

#### 4.3.1.7.3.9 Efficacy by Inhaled Corticosteroid/Cromolyn use at Baseline

The study population was stratified by use of these agents at baseline (the ICT group) or whether they were managed on bronchodilator therapy alone (the BDT group). The ICT group constituted 42% of the study population overall compared to 58% for the BDT group (see "Results: Demographics").

Mean change from baseline in FEV<sub>1</sub> has been summarized by ICT and BDT strata and displayed in the table below. Endpoint data for these two subgroups were generally consistent with the overall ITT population, that is, the two FP groups showed more improvement than the placebo group, and the subjects in the Diskus group showed somewhat more improvement than the subjects in the Diskhaler group. BDT subjects in all three treatment groups showed greater numerical improvement in FEV<sub>1</sub> than ICT subjects. Placebo-treated ICT subjects experienced a fall in FEV<sub>1</sub> when CS treatment was withdrawn at randomization, as might be expected. Higher percentages of subjects in the BDT group remained in the study at Visit 10 compared to the ICT stratum. A full survival analysis by inhaled CS strata was not included in this submission, however.

#### Mean Change from Baseline in FEV<sub>1</sub> by ICT/BDT Strata

	Placebo		Diskus		Diskhaler	
	ICT	BDT	ICT	BDT	ICT	BDT
N	30	40	24	40	35	44
Baseline FEV <sub>1</sub> (L)	2.29	2.50	2.34	2.45	2.42	2.54
Mean Change at Endpoint	-0.15	0.17	0.41	0.59	0.31	0.49

#### 4.3.1.7.4 Safety Results

##### 4.3.1.7.4.1 Extent of Exposure

A total of 213 patients received at least one dose of study medication and therefore have been included in the safety analysis. Their extent of exposure is shown in the table below. On average, the FP-treated patients were exposed for approximately 78 days out of an 84-day trial. The placebo patients received approximately 20 fewer days of exposure.

#### EXTENT OF EXPOSURE TO STUDY MEDICATION\*

	Placebo	FP 500 BID DK	FP 500 BID DH
Number Baseline Completed	70 33 (47%)	64 54 (84%)	79 68 (86%)
Exposure(days): Mean	58.5	78.4	-77.9
Median	81.0	84.0	84.0

\* Table 31 and p.70; Vol.33

#### 4.3.1.7.4.2 Adverse Events (AE)

The adverse events identified in this trial are not substantially different from those reported in the ADVERSE REACTIONS section of the approved product labeling for Flovent™ Rotadisk. These common adverse events will therefore not be discussed in great detail in this review.

Overall, 61% of the placebo group reported at least one adverse event during this trial, which was comparable to the FP-treated groups, 66% in the Diskus group and 71% in the Diskhaler group. By organ system, the most commonly reported AE's in all treatment groups were within the ENT system (33-48%) followed by Neurologic (10-16%), Lower Respiratory (10-16%), GI (8-13%) and non-site specific (7-14%). In descending order of frequency, the top ENT AE's were throat irritation (9-22%), URI (10-16%), and nasal congestion (3-6%). Among the AEs which were more common in the FP-treated subjects were throat irritation, 9% of the placebo group compared to 22% of the Diskus group and 15% of the Diskhaler group; headaches, occurring in 7%, 14%, and 13% of placebo, DK, and DH patients, respectively; bronchitis, occurring in 1%, 14%, and 16% of placebo, DK, and DH patients, respectively; and cough, reported by 4%, 8%, and 9% of the placebo, DK, and DH groups, respectively. Oropharyngeal candidiasis or candidiasis unspecified site was reported for 3 placebo subjects, 3 Diskus subjects, and 6 Diskhaler subjects.

When analyzed by demographic subgroups, there was no apparent difference in overall numbers of AEs based upon gender or ethnicity. The number of non-Caucasian subjects was very small, for example, only 13 African American subjects were enrolled. Female subjects experienced approximately the same number of events as males, however, they seemed to be disproportionately affected by probable FP-related adverse local reactions, for example, throat irritation (6% placebo, 39% DK, and 17% DH compared to 9%, 22%, and 15%, respectively, in the overall ITT population). The other "local" reactions where the frequency of occurrence in female subjects exceeded that of the overall ITT population included dysphonia (0% placebo, 7% DK, and 14% DH vs. 0%, 3%, and 8%, respectively for the ITT) and bronchitis (3% placebo, 14% DK, and 14% DH vs. 1%, 8%, and 9%, respectively for the ITT). An analysis based on age was also performed, differentiating the adolescent subgroup (12-17 years) from the 18-64 years group and the >64 years group. The latter was comprised of only 3 subjects, making meaningful analysis impossible. The adolescent group, comprised of 30 subjects, had a slightly greater overall incidence of AEs than the population as a whole and were over-represented in the "drug interaction, overdose and trauma" category. Again, the small number of subjects places limits on the analysis. The 18-64 year old subjects had an AE profile similar to the ITT population as a whole.

There were no deaths and no serious AEs in the double-blind phase of this study. Three patients were withdrawn due the AEs, two placebo subjects (both due to URI) and one Diskus subject (URI in combination with lower respiratory tract infection).

There were eight reports of thrush, two each in the placebo and DK, and four in the DH group. Not unexpectedly, there were no reports of cataracts, glaucoma, or osteopenia in this 12-week trial. No adverse event specifically coded as "HPA axis suppression" was reported.

#### 4.3.1.7.4.3 Laboratory Data (excluding HPA-axis)

Blood samples for serum chemistry, LFT's, and hematology were obtained at baseline and at study endpoint. No subject was withdrawn for abnormal laboratory values, and no abnormal laboratory value was reported as an adverse event. A few subjects (1-4% per group, maximum) had "clinically significant" laboratory values by pre-specified criteria reported at any time post-randomization (Table 50; Vol.33, p.186). Each of these patients was further discussed in the text of the study report. A few more patients had laboratory values outside of the normal range, many of which were probably chance variation expected among a large group of patients. Abnormalities of relevance to this review, either because of known side-effects of CS or because of post-marketing surveillance, would include glucose, bicarbonate, potassium, eosinophil count, and alkaline phosphatase. These have been separately noted in this review (below).

*Reviewer's Comment: It is worth emphasizing at this point that the sponsor's definition of "clinically significant" was not synonymous with "outside of the normal range." A "clinically significant" abnormality was generally well above or below the accepted normal range for a given value. For example, a subject whose serum glucose was 90 mg/dL at baseline but which rose to 150 mg/dL post-randomization would be said to have hyperglycemia by most standard criteria (normal range 75-115 mg/dL). However, this patient would not be reported as having a "clinically significant" abnormality because the glucose was not outside of the pre-specified "clinically significant" range of >175 or <55 mg/dl. Although such a patient would have been reported numerically in one of the sponsor's "shift tables" (Table 52; Vol.33 for this study), there would be no simple way to retrieve his/her precise laboratory values or CRF, because patient identifiers have not been included in the table.*

There were no reported clinically significant elevations in bicarbonate. For glucose, there was one placebo and three Diskhaler subjects who had clinically significant low plasma glucose values (in the 50 mg/dL range). There were two Diskus and three Diskhaler patients whose glucose went from normal to high post-randomization, compared to two placebo patients with a normal to high shift.

No subject had clinically significant hypokalemia, or underwent a shift from normal to low potassium during the study, although there was one Diskhaler

patient (———1373) who had multiple unexplained high potassium readings, ranging from 6.1 meq/L to 6.5 meq/L during the study. No adverse clinical events were reported for this patient, ECG was normal (Vol.39; p.20), and no concomitant abnormalities in other laboratory values, such as platelets, were reported. Her bicarbonate level was normal.

A single Diskus patient (———1504) had an eosinophil count that was reported as clinically significantly elevated, at  $1.57 \times 10^3/\text{mcL}$ . The count was elevated at baseline, but had increased further post-randomization. The patient had other abnormalities, including a low neutrophil count of  $1.48 \times 10^3/\text{mcL}$ , and complaints of nasal congestion, knee pain, and sinusitis during the study. Although his CXR was "abnormal," the reading was "parenchymal scarring with adjacent pleural thickening" (Vol.39; p.30), without mention of infiltrates. Additional information available from the data listings did not suggest that this patient had a vasculitis. The shift tables gave a total of 14 patients whose eosinophil counts either started high and remained high during treatment, or increased from low or normal to high during treatment. There were 6 in the placebo group, 4 in the DK group, and 4 in the DH group. No further information is available about these individuals.

There were no subjects with "clinically significant" elevations in alkaline phosphatase (AP). A total of 6 patients had elevated AP at some point post-randomization, however, 5 of these had elevations at baseline. Two of these patients were in the placebo group, and the remainder were receiving FP. No additional information is available.

#### 4.3.1.7.4.4 HPA Axis Assessment

The HPA axis was assessed at baseline, 4 weeks, and at study endpoint (Visits 2, 6, and 10) by means of unstimulated (basal) AM plasma cortisol levels. Any value  $<5 \text{ mcg/dl}$  was considered abnormal. All samples were collected pre-dose. A subset of patients underwent extended serum cortisol testing over 12-hours concurrent with plasma FP determinations (see below).

Eleven (11;5%) patients had one or more post-randomization plasma cortisol values which were abnormally low, 1 (1%) in the placebo group, 4 (6%) in the Diskus group, and 6 (8%) in the Diskhaler group. For 6 of these subjects, only the final determination was abnormal, one in placebo, 2 in DK, and 3 in DH. For the other 5, one or more intermediate plasma cortisol values had been abnormal (see supporting table ST-40, Vol.33; p.310). For the 6 patients with abnormal readings at endpoint, follow-up was available for one, whose plasma cortisol had normalized.

HPA axis assessment was also performed by means of multiple timed, post-dosing plasma cortisol determinations drawn over a 12-hour period. Testing occurred at clinic Visits 2, 3, and 6 (baseline, after one week of dosing, and

at 4 weeks). A subset of patients (41 subjects at 3 sites) underwent extended testing over a 12-hour period at 20 min., 40 min., 1h, 2h, 3h, 4h, 5h, 6h, 8h, 10h, and 12h. No baseline plasma cortisol was drawn. This adrenal axis assessment was included as a secondary endpoint in a protocol primarily conducted to understand the pharmacokinetics of FP (see Appendix 10, Vol.35, and below). The reader is referred to the *Biopharmacology Review* of this submission for a full analysis and discussion of this experiment.

The study population was comprised of 13 patients in the placebo group (4F, 9M), 13 in the DK group (6F, 7M), and 15 in the DH group (6F, 9M). Only the plasma cortisol AUC values have been provided for each clinic visit (Table 19; Vol.35; pp.137-8). Three AUC values corresponding to each of the three clinic visits (V2, V3, and V6) are available for 10 of 13 placebo subjects, 9 of 15 DH subjects, and 11 of 13 DK subjects. For the 11 total subjects with missing data, most had Visit 2 values only and no Visit 3 or Visit 6 readings. The mean AUC was determined for each clinic visit (V2, V3, and V6; see Table 19). "In order to reduce the number of analyses, data from V2 was excluded" (see Vol.35, p.26), and comparisons were made between the 3 treatment groups for the V3 and V6 visits only. No statistically significant differences among treatment groups were found for the selected comparisons made using 12-hr plasma cortisol AUC data.

*Reviewer's Comment:* It is important to emphasize that this experiment was designed to understand FP pharmacokinetics, including possible drug interaction with terfenadine, gender effects, time to steady state,  $C_{max}$ , and so forth, not to assess the impact of FP on the adrenal axis. In the latter case, the final testing should have been performed at study endpoint rather than at the 1 and 4-week time-points. In addition, overnight testing would have been a far better choice since it is more sensitive to subtle changes in the adrenal axis than is daytime testing. An overnight urinary cortisol collection (or even a 24-hour collection) would have been superior to a 12-hour daytime plasma cortisol AUC for this purpose.

In addition to choice of test, there were problems in the design and conduct of this test. A baseline plasma cortisol should have been drawn, and the V2 (pre-dose) AUC values should have been included in the analysis. Finally, a comparison should also have been made between pre-treatment (i.e. V2) plasma cortisol AUC and endpoint cortisol AUC within each treatment group. In this case, there was no endpoint cortisol AUC, only a Week 4 (V6) cortisol, but nevertheless the data may have provided some insight on the effect of short term FP dosing on the HPA axis. Data from individuals could also be assessed before and after treatment. In the latter case, a decision about what to do about incomplete data (V2 but no V3 or V6 cortisol AUC) would need to be made prospectively.

In summary, this 12-hr plasma cortisol test is seriously flawed in design, conduct, and analysis for the purpose of drawing any reliable conclusions about the effects of FP on adrenal cortisol output.

---

#### 4.3.1.7.4.5 Other Safety Evaluations

These assessments included oropharyngeal examinations, vital signs, physical examinations, and ECG's. There were no clinically significant differences between placebo and treatment groups or between the two FP treatment groups relevant to this application.

#### 4.3.1.7.5 Comparative Pharmacokinetics: Diskus vs. Diskhaler

For a full discussion of the clinical pharmacology of the Diskus vs. the Diskhaler in the adult and adolescent population, the reader is referred to the *Biopharmacology Review* of this application.

A subset of subjects enrolled in this trial had FP plasma levels assayed over a 12-hr period post-dosing during clinic Visits 2, 3, and 6 (baseline, Week 1, and Week 4). Samples were taken as described in the previous (HPA axis) section. All subjects in the ITT population had samples drawn 20 and 40 minutes after dosing during these same clinic visits. The impact of gender, concomitant terfenadine, and device-by-visit was also investigated.

The following table shows the PK data for the 41 patients studied for 12-hours (Appendix 10; Vol.35; p.6). The point estimate of the ratio Diskus/Diskhaler for  $C_{MAX}$  was 0.77 (90%CI, 0.51, 1.16). The point estimate of the ratio Diskus/Diskhaler for AUC was 1.15 (90% CI, 0.69, 1.94).

#### FP PK VALUES FOLLOWING CLINIC VISIT 6 (pg/mL)

	DK 500 BID	DH 500 BID	p-value
$C_{MAX}$	92.1	119.5	0.183
$C_{MAX}$ 95%CI	63.7, 133.2	86.2, 165.6	
AUC	474.3	411.6	0.270
AUC 95%CI	297.8, 755.6	272.4, 621.9	

Based on these and other data, the systemic exposure of FP delivered from the Diskus to adult and adolescent patients with mild to moderate asthma appears to be similar to that from the Diskhaler. The AUC was slightly higher for the Diskus while the  $C_{MAX}$  was somewhat greater for the Diskhaler, however, the difference between the two values was not statistically significant for either comparison. This finding differs from that in the pediatric population, where the opposite appears to be true, that is, systemic exposure may be slightly higher with the Diskhaler.

*Reviewer's Comment: Given the apparent numerically higher AUC with the Diskus in this population, and the fact that almost all plasma FP results from lung deposition, the PK results are not out of keeping with the efficacy results, where there was a numerically, but not usually statistically, better response with the Diskus compared to the Diskhaler device.*

Because both drugs are metabolized by the same CYP3A4 microsomal enzymes, the possibility of a fluticasone/terfenadine drug interaction data was investigated. Terfenadine 60 mg BID had been co-administered with FP 500 mcg BID via DK or DH to selected patients throughout this trial. This was primarily because these individuals were receiving the drug at baseline for allergic rhinitis or other concomitant conditions, and they were allowed to continue during the double-blind phase of the trial. When data from these patients were analyzed, it was reported as demonstrating no statistically significant effect of concomitant terfenadine 60 mg BID on FP kinetics.

FP pharmacokinetics was analyzed separately for male and for female subjects. No significant gender effect was reported.

In summary, in adults and adolescents with mild to moderate asthma, dry powder fluticasone propionate administered via the Diskus multi-dose powder inhaler has systemic bioavailability which is similar to that of dry powder FP Rotadisk administered via the approved Diskhaler device. This finding corroborates the both the safety and efficacy data from this trial, that is, although there were minor numerical differences between the two devices on the various clinical endpoints, this difference did not achieve statistical significance. The PK results from this adult trial FLTA2001 stand in contrast to the PK data from the pediatric device comparison trial FLTA2006. Although more limited, these data show that pediatric asthmatics receive greater systemic exposure to FP with the Diskhaler device rather than the Diskus. The clinical data from FLTA2006 are silent or equivocal on this point, that is, although there was no significant difference between the two devices on any safety or efficacy endpoint, numerical differences in many secondary efficacy endpoints tended to favor the Diskhaler more often than the Diskus.<sup>11</sup>

It is not unreasonable to expect that differences in inspiratory effort and/or pulmonary mechanics (or some other unmeasured difference between children and adults), could affect drug delivery. If there is a difference, however, it is important that it be predictable and consistent in order for the drug product to be properly labeled, and to be considered safe and effective. More studies are needed to understand the scientific basis for

---

<sup>11</sup>

**Lingering doubts remain from clinical trial FLTA2002, which showed no difference in efficacy between FP 500 or 1000 mcg BID in oral CS-sparing effect. Recall from study FLI-210 (NDA #20-548 for Flovent MDI) that there was a clear dose response between 750 and 1000 mcg BID of FP delivered via the MDI, in spite of PK data arguing that the aerosol gave greater FP systemic bioavailability than the DPI. Those PK studies were not conducted with severe asthmatics, however.**

the observed differences in FP systemic bioavailability between the two devices in the two populations.<sup>12</sup>

#### 4.3.1.8 Conclusions

##### 4.3.1.8.1 Efficacy Conclusions:

Dry powder FP delivered from the Diskus multi-dose powder inhaler (MDPI) device at a dose of 500 mcg BID has been shown to be efficacious in the treatment of mild-to-moderate asthma in adult and adolescent patients, and to have efficacy similar to that of the approved Diskhaler. Efficacy was demonstrated for both devices for the primary endpoint, FEV<sub>1</sub>. Efficacy was supported for this device by the results of all but one of the secondary endpoints. Numerical superiority of the Diskus device over the Diskhaler was seen on the primary endpoint and on most of the secondary endpoints, although the difference achieved statistical significance on only two of the secondary endpoints. Both devices were superior to placebo on the endpoint survival-in-study.

Subgroup analysis showed no significant difference in efficacy based on the subject's gender or ethnicity, although there were relatively few non-Caucasian subjects. With regard to inhaled corticosteroid use at baseline compared to inhaled bronchodilator use only, (ICT vs. BDT), the BDT subgroup showed a greater improvement on most efficacy endpoints, not unexpectedly. Also as expected, the ICT placebo group did generally show a deterioration in lung function, as expected upon withdrawal of inhaled corticosteroids.

This trial included a single dose of FP via the Diskus device, therefore no statement regarding dose-response can be made. The dose which was selected, FP 500 mcg BID via Diskhaler, is twice the highest recommended starting dose for patients with mild to moderate asthma and, in fact, corresponds to the lower of the two recommended starting doses for the oral CS-sparing indication (see earlier review of clinical trial FLTA2002). Although lower doses have been studied in this submission, no head-to-head comparisons of different BID doses in a single trial were submitted. The absence of such comparisons is a weakness in this NDA submission, since it could provide important information relevant to the NAEPP 2 guidelines calling for the titration of inhaled CS to the lowest effective dose.

##### 4.3.1.8.2 Safety Conclusions:

Based upon Study FLTA2001, dry powder FP 500 mcg BID administered via the Diskus appeared to be safe when used to treat adults and adolescents with mild-to-moderate asthma, and there appeared to be no safety difference between the two devices.

---

<sup>12</sup> : Perhaps starting with a hypothesis testable *in vitro*, that is, that flow rate affects drug delivery.

The most frequently occurring adverse events were in the ENT system, the most common being throat irritation, followed by neurological, the most common of which was headache. The overall profile was not different from that described in the approved labeling for Flovent Rötadisk Diskhaler.

There were no deaths in the study, no serious adverse events, and a total of three withdrawals due to adverse events.

Routine clinical laboratory assessments, physical examinations, ECG's, and vital signs did not disclose any unique or unexpected safety issue relevant to this product.

The assessments of HPA axis function included basal AM plasma cortisol and 12-hour plasma cortisol AUC. The latter was included as a secondary endpoint during a device comparison clinical trial designed to assess fluticasone propionate PK. The trial was not properly designed to assess this important safety parameter, and therefore samples were not obtained at key times, such as pre-dose or at study endpoint. With regard to basal AM cortisol testing, there were 11 patients who had abnormally low plasma cortisol levels post-randomization, 10 in the two FP groups and one in the placebo group.

Pharmacokinetic studies performed on a subset of subjects suggest that the Diskus and the Diskhaler are comparable in delivering dry powder FP to the lungs of adults and adolescents with mild to moderate asthma, a finding different from that observed for pediatric asthmatics.

A single BID dose of FP via Diskus was used in this trial, therefore, whether the safety profile would have improved had a lower BID dose of FP been used is not known.

#### **4.3.1.9 Labeling Considerations:**

Comments relevant to labeling this product for use in adults and adolescents will be deferred until the end of this section of the review, following the four supportive trials FLTA2003, FLTA2004, FLTA2005 and FLTA2016.

#### **4.3.2 FLTA2003:**

“A randomized, double-blind, double-dummy, parallel-group, comparative trial of inhaled fluticasone propionate 100 mcg BID and 200 mcg QD via multi-dose powder inhaler, beclomethasone dipropionate 168 mcg via metered-dose inhaler, and placebo in adolescent and adult patients with mild to moderate asthma.”

##### **4.3.2.1 Background Information**

This clinical trial, FLTA2003, ran concurrently with FLTA2004, which

had an identical design. The two trials differed in a single key inclusion criterion: baseline inhaled corticosteroid (CS) usage. FLTA2003 recruited inhaled CS-naïve subjects only (referred to by the acronym "BDT" or bronchodilator therapy) and FLTA2004 included only subjects already maintained on a stable dose of inhaled CS (ICS).

Taken together, FLTA2003 and FLTA2004 were similar in concept to pediatric clinical trial FLTA2008, which compared twice daily FP via Diskus to the same total dose administered once daily in the morning. The trials compared the same two doses, FP 100 mcg BID via Diskus and FP 200 mcg QD via Diskus. The two major differences were the addition to the adult studies of a positive control comparator arm (BDP 168 mcg BID via MDI) and the inclusion of both inhaled CS users and non-users at baseline by means of a balanced stratification in the pediatric study.

#### 4.3.2.2 Objectives

The objectives of this study were to compare the efficacy and safety of FP 100 mcg BID and FP 200 mcg QD administered to BDT asthmatics via multi-dose powder inhaler (MDPI or Diskus), BDP 168 mcg BID via metered-dose inhaler (MDI) and placebo BID terms of the following:

- **Efficacy:** Primary efficacy variable: FEV<sub>1</sub>; Secondary efficacy variables: survival in study, physician global assessment, patient-determined PEFR, symptom scores, rescue beta-agonist use, and nighttime awakenings requiring beta-agonist
- **Safety:** Physical examination, clinical laboratory, HPA-axis assessment via AM plasma cortisol, 12-lead ECGs, and adverse events

#### 4.3.2.3 Setting

Conducted at 25 outpatient sites in the US between 6 April 1995 and 2 March 1996. Enrollment per center ranged from 1 (<1%) to 21 (7%), with a mean of 12 patients/center and a median of 12 patients/center.

#### 4.3.2.4 Endpoints

##### 4.3.2.4.1 Efficacy Endpoints:

- The primary efficacy variable was change from baseline in AM pre-dose FEV<sub>1</sub> determined at each clinic visit.
- Secondary efficacy variables:
  - Survival in study
  - Diary AM and PM PEFR
  - Patient-rated Symptom Scores (scale of 0-3 where 0=ineffective and 3=very effective)
  - Rescue  $\beta$ -agonist use
  - Nighttime awakenings requiring  $\beta$ -agonist

##### 4.3.2.4.2 Safety Endpoints

- Adverse events
- Clinical laboratory tests

- Basal AM plasma cortisol
- Physical examination
- Vital Signs
- 12-lead ECG

#### 4.3.2.5 Design

FLTA2003 was a 12-week, randomized, double-blind, double dummy, placebo-controlled, multi-center clinical trial in adult and adolescent patients with mild to moderate chronic asthma not managed on inhaled CS. After an initial screening visit, subjects entered a 2-week, single blind, double-dummy run-in period with placebo dispensed from two different devices, the Diskus (DK) and a conventional metered dose inhaler (MDI). In addition to becoming familiar with these two devices, all subjects were switched from their usual  $\beta$ -agonist bronchodilator to Ventolin and were instructed to discontinue all other anti-asthma medications with the following exceptions: salmeterol and/or theophylline, as long as they were maintained on a stable dose throughout the trial. At the end of the two-week run-in period, eligible subjects entered the 12-week double-blind phase of the study. Subjects were assigned randomly to one of 4 treatment groups, placebo, FP 100 mcg BID via DK, FP 200 mcg QD via DK, or BDP 168 mcg BID via MDI. Assessments occurred weekly during the first 4 weeks of the 12-week dosing period, then biweekly until the end of the study (Weeks, 0, 1, 2, 3, 4, 6, 8, 10, and 12).

#### 4.3.2.6 Summary of Protocol (includes all amendments)

##### 4.3.2.6.1 Study Population

###### *Inclusion Criteria*

- Male or female
  - If female, surgically sterilized, post-menopausal or practicing acceptable contraception
- Age 12 years or older
- Diagnosis of asthma by ATS criteria for at least 6 months
- Best FEV<sub>1</sub> 50-80% predicted (Crapo; or Polgar if age 12-17 years; multiplied by 0.88 if subject was African American)
- Variability of FEV<sub>1</sub> of 15% or increase in FEV<sub>1</sub> within 30' of 2-4 puffs albuterol

###### *Exclusion Criteria*

- Current use of inhaled CS (volunteers must not have used inhaled CS for at least one month prior to Visit 1)
- Life-threatening asthma
- Use of nonsteroidal immunosuppressive therapy for asthma, such as cyclosporine, methotrexate, or gold
- Cromolyn or nedocromil use in prior 4 weeks
- URI or lower resp. tract infection in prior 2 weeks

- Influenza vaccination in prior 2 weeks
- >10 pack-year hx/o cigarettes and/or smoking any tobacco products in prior year
- Other significant concomitant disease or medical condition
- Mentally challenged
- Concomitant psychiatric disorder
- History of alcohol or substance abuse
- Allergy to corticosteroids (CS) or  $\beta$ -agonists
- Clinically significant abnormality on screening laboratory or 12-lead ECG
- AM plasma cortisol < 5 mcg/dL
- Glaucoma or posterior subcapsular cataracts (PSC)
- Clinically significant abnormality on CXR
- Prior participation in MDPI study (Diskhaler participation OK)

*Disallowed Medications*

- At time of enrollment:
  - Any antibiotic in prior 2 weeks
  - Any investigational drug in prior 90 days
  - Oral, intranasal, or parenteral CS in prior month
  - Inhaled CS use in prior month
- Specifically prohibited during the trial:
  - Anticholinergics
  - Anticonvulsants
  - Antidepressants, including polycyclics and MAOI's
  - Long acting antihistamines or antihistamine/decongestant combinations (astemizole must have been continued 6 weeks prior to Visit 1)
  - Long acting oral decongestants — nasal spray was allowed for a 5-day period as needed)
  - All antihistamines except loratidine, if started prior to the study and continued throughout its duration
  - Phenothiazines
  - Orally inhaled nedocromil or cromolyn (intranasal cromolyn OK if discontinued 12 hours prior to clinic visit)
  - Macrolide antibiotics
  - Quinolone antibiotics
  - $\beta$ -blockers
  - digitalis
  - ketoconazole, fluconazole
- All anti-asthma medications except Ventelin MDI (substituted for any other  $\beta$ -agonist), theophylline (if on a stable dose prior to start of study), or salmeterol (if on a stable dose prior to start of study)

#### 4.3.2.6.2 Treatment Arms and Dosing

Subjects were randomized to one of four treatment groups (see table below). Each subject received two DKs, Device A and Device B, and an MDI, Device C. A dose consisted of two blisters from DK Device A and four puffs from MDI Device C administered at 8:00 AM and 2 blisters from DK Device B and four puffs from MDI Device C administered at 8:00 PM. MDI Device C was exchanged every two weeks and DK Devices A and B were exchanged every four weeks until the end of the study.

#### TREATMENT ARMS AND DOSING STRATEGY

Treatment	Twice Daily Dosing AM and PM
FP 100 mcg BID	2 blisters FP 50 mcg via DK Device A (AM) 2 blisters FP 50 mcg via DK Device B (PM) 4 puffs placebo MDI Device C (AM and PM)
FP 200 mcg QD	2 blisters FP 100 mcg via DK Device A (AM) 2 blisters placebo via DK Device B (PM) 4 puffs placebo MDI Device C (AM and PM)
BDP 168 mcg BID	2 blisters placebo via DK Device A (AM) 2 blisters placebo via DK Device B (PM) 4 puffs BDP 42 mcg MDI Device C (AM and PM)
Placebo	2 blisters placebo via DK Device A (AM) 2 blisters FP placebo via DK Device B (PM) 4 puffs placebo MDI Device C (AM and PM)

*Reviewer's Comment: Diskus devices were exchanged every 4 weeks, which ought to be life-of-device for the once daily dosing indication at 200 mcg daily administered from the 100 mcg/blister device. The 100 mcg BID dose was administered solely from the 50 mcg/blister device (for blinding purposes), not the 100 mcg/blister Diskus. No trial in adults and adolescents has been included in this submission utilizing the 100 mcg/blister device dosed at 100 mcg BID. There is also no trial in adults utilizing the 50 mcg/blister Diskus dosed at 200 mcg QD (see "Efficacy Conclusions, below).*

#### 4.3.2.6.3 Treatment Assignment:

Subjects were given a number at Visit 1. Eligible subjects who completed the screening period and met randomization criteria were randomly assigned to one of four treatment arms in accordance with a code. Eligible subjects were assigned the lowest available treatment number in the chronological order of presentation. Subject and treatment numbers were unique and could not be reassigned. No specific attempt to balance enrollment at individual centers was mentioned in the protocol.

#### 4.3.2.6.4 Study Sequence

*Screening Period (Visits 1- 2):* The screening period was used to confirm eligibility, assess asthma stability, obtain baseline data, assess compliance, and instruct the subjects in the use of all the devices and study procedures

to be used during this trial. (See the attached "Figure 1" for a summary schedule of events: Vol.95; p.77).

With the exception of the CXR, all screening and baseline tests indicated on Figure 1 were to be completed at Visit 1 in order to be available at Visit 2. A CXR was optional for patients who could present an acceptable CXR performed in the prior 12 months. Other routine assessments performed at Visit 1 included medical history, physical examination, vital signs, oropharyngeal exam, clinical laboratory tests, pregnancy test if applicable, AM plasma cortisol, FEV<sub>1</sub> with reversibility testing, if appropriate, and PEFr.

Subjects received instructions on daily routine assessments and procedures they were to perform for the subsequent two weeks. Diary PEFr was to be measured twice daily in triplicate using a  Peak Flow Meter, and the highest value recorded in the subject's diary. AM PEFr was to be measured at 8:00 AM before study medication but after other diary assessments. PM PEFr was to be measured at 8:00 PM after study medication had been given.

Subjects received diary cards at Visit 1, and were instructed to record their asthma symptoms, rescue  $\beta$ -agonist use, and nighttime awakenings daily throughout the study.

The screening period of this trial was single-blind. Each subject received two placebo Diskus's, Device A and Device B, and one MDI, Device C. Subjects were encouraged to take their medication at the same time every day, two blisters from Device A and four puffs from Device C at 8:00 AM and two blisters from Device B and four puffs from Device C at 8:00 PM.

Subjects could continue to take their baseline asthma medication at this time, except that Ventolin was substituted for their own particular  $\beta$ -agonist. The Ventolin was to be used only to treat symptoms, and not taken on a regular basis (even if that was how it was previously taken). Theophylline and salmeterol could both be continued during the baseline period as well as for the duration of the trial, if they had been used previously in the management of the patient's asthma. Doses must remain constant, however, throughout the study, and both medications were to be withheld prior to each clinic visit, salmeterol for at least 12 hours and theophylline for 12-36 hours. Subjects were also to withhold Ventolin for 6 hours and the AM dose of study medication on the morning of the scheduled clinic appointments.

*Treatment Period (Visits 2 - 10):* To be eligible for the study, in addition to meeting the Inclusion/Exclusion criteria above, subjects had to have met the following criteria:

- Their asthma had been relatively stable. "Stable" was defined as having no day in the last 7 in which  $\geq 12$  puffs of Ventolin MDI was used **and** no more than 4 mornings in the last 7 where the AM PEFR was decreased  $>20\%$  from the prior PM PEFR **and** no more than 2 nights in the last 7 with awakenings requiring Ventolin.
- Their clinic spirometry met the following criteria:
  - Best FEV<sub>1</sub> 50-80% predicted (Polgar for ages 12-17 years; Crapo for 18 years and older)
  - Best FEV<sub>1</sub> from Visit 2 within  $\pm 15\%$  of Best FEV<sub>1</sub> from Visit 1.
- Adequate compliance was demonstrated:
  - At least 70% of study medication had been used
  - Diary card had been completed
  - Anti-asthma medications had been withheld as required

At Visit 2, subjects exchanged their placebo devices for a 2- or 4-week supply of the appropriate Diskus (DK) and MDI devices, as determined by their randomization code. Instructions regarding withholding medications prior to clinic visits were repeated.

Other assessments that occurred at Visit 2 can be found summarized on the attached Figure 1 (Vol.95; p.77). These included adverse event assessment, oropharyngeal exam, baseline PFTs, and collect/dispense diary card.

Eligible subjects needed to meet additional criteria at each clinic visit to continue in the study. "Stability limits" were therefore defined at Visit 2 for PEFR and FEV<sub>1</sub>:

- FEV<sub>1</sub> stability limit: 20% decrease from the best FEV<sub>1</sub> at Visit 2
- PEFR stability limit: 20% decrease from mean diary AM PEFR from the past 7 days

Subjects not meeting the following "continuation criteria" at each clinic visit (Visit 3 and beyond) were discontinued for lack of efficacy:

- No more than 2 days in the last 7 in which  $\geq 12$  puffs of Ventolin MDI was used
- No more than 3 days in the last 7 where the AM or PM PEFR was below the PEFR stability limit
- No more than 2 nights in the last 7 with awakenings requiring Ventolin.
- A clinic FEV<sub>1</sub>  $\geq$  the FEV<sub>1</sub> stability limit

A subject could also be discontinued for lack of efficacy if they experienced a clinical asthma exacerbation requiring emergency intervention or treatment with a proscribed medication. All data from subjects discontinued for lack of efficacy prior to the time of their discontinuation was included in the analysis, carried forward (LVCF) to

endpoint as the last evaluable value. Termination procedures similar to Visit 10 (Week 12) study endpoint procedures were also conducted.

Visits were scheduled weekly for the first 4 weeks, then every other week until study endpoint at 12-weeks. At Visits 3-9 the following procedures were performed:

- Assess subject's compliance including withholding medication (required for PFTs and other procedures to be performed)
- Assess subject's "continuation criteria" (must be met or patient was terminated for lack of efficacy)
- Review previous diary cards and dispense new cards
- Adverse event assessment especially acute asthma exacerbation
- PFTs.
- Collect/dispense study medication (Diskus: Visit 2, Visit 6 or 4 weeks, and Visit 8 or 8 weeks; MDI: Visits 2, 4, 6, 7, 8, and 9: every 2 weeks)
- Oropharyngeal exam (Visits 6, 8, and 10)
- Clinical laboratory tests/plasma cortisol: (Visit 10 or endpoint)

At study endpoint (Visit 10) or early termination, the usual scheduled clinic assessments were made, in addition to the same as performed at baseline (physical exam, etc.), and the special assessments summarized in the bullet points above. Study devices were collected, and overall compliance with study procedures was assessed by blister counts, completion of diary cards, and whether subject followed instructions to withhold medication on the morning of the clinic visit.

#### 4.3.2.6.5 Efficacy Assessments

The primary efficacy variable was pre-dose FEV<sub>1</sub>. FEV<sub>1</sub> was performed in triplicate using approved spirometric equipment according to ATS recommendations. The subject could be sitting or standing during the maneuver, but was required to be consistent throughout the study. If two FEV<sub>1</sub> readings were identical, the once with the highest FVC was recorded.

Secondary efficacy variables included all of the following:

- Survival in the study
- Diary AM and PM PEFr  
(Using a peak flow meter, AM before study medication and PM after study medication. The highest of three values was recorded. The AM/PM PEFr difference was also assessed as a secondary endpoint)
- Subject-rated daily symptom scores on a scale of 0 (none), 1 (mild), 2 (moderate), or 3 (continuous or disabling)
- Number of nighttime awakenings requiring Ventolin
- Rescue Ventolin use

#### 4.3.2.6.6 Safety Assessments

- Clinical Adverse Events (AE)
- Clinically significant changes in clinical laboratory values
- Clinically significant changes in physical examination, oropharyngeal exam, vital signs, or 12-lead ECG
- HPA-axis effects via basal AM cortisol

#### 4.3.2.6.7 Statistical Methods

*General Statements:* All statistical tests were two-sided. Treatment differences at or below the 0.05 level were considered significant. Pair-wise comparisons were performed without adjusting p-values for the number of comparisons made and pair-wise p-values were interpreted only when the overall test among treatment groups was statistically significant.

*Power Calculations:* Mean and standard deviation of the primary endpoint was estimated based on prior studies conducted by the sponsor. Enrollment was planned to obtain 280 evaluable (70 per arm) subjects to provide >80% power of detecting a difference in FEV<sub>1</sub> of 0.25L between any two treatment groups, using a t-test with a significance level of 0.05. The proposed sample size would also provide >80% power to detect a difference in AE of 16% between any two treatment arms.

*Populations:* The Intent-to-Treat (ITT) Population was used for most calculations, unless otherwise stated. The ITT Population included any subject who had received at least one dose of study medication. The Efficacy Population was a subgroup that included only those subjects who had no major protocol violations during the study. The decision to exclude a subject from the Efficacy Population was to have been made prior to breaking the blind.

*Background Characteristics:* Comparisons between treatment groups were based on ANOVA F-test controlling for investigator for age, height, and weight, and on the Cochran-Mantel-Haenszel test controlling for investigator for gender, smoking history, method of contraception and ethnic origin.

*Efficacy:* The primary efficacy parameter was AM pre-dose FEV<sub>1</sub> in the ITT population. Testing for the primary and for most (continuous) secondary efficacy parameters was first performed on data from all investigators combined, assessing investigator effects and treatment-by-investigator interactions at a significance level of 0.10. An ANOVA F-test was used to compare change-from-baseline for each of the time-dependent variables at endpoint (or at other selected time points). Endpoint was the last recorded value for the ITT population and the last evaluable value for the efficacy population.

Withdrawals from the study due to lack of efficacy were evaluated using Kaplan-Meier estimates of survival, and overall and pairwise treatment comparisons were based on the Log-rank test.

As stated above, continuous parameters such as PEFr measurements were tested with an ANOVA F-test controlling for investigator. Tests were performed on mean values over days within individual weeks. Parameters having discrete values such as symptom scores were analyzed using the non-parametric van Elteren test based on 7-day subject averages.

*Reviewer's Comment: The minimum number of diary entries required in a given week before data could be analyzed was not stated.*

*Safety:* All safety assessments were based on the ITT population. Adverse events were tabulated by organ system, treatment group, severity, and relation to study drug. Laboratory variables, ECG, VS, and physical exam were reported by presence and/or direction of change and whether or not abnormal. AM plasma cortisol results were tabulated by treatment group based on an abnormal value, defined as any basal (un-stimulated) reading <5 mcg/dL. No statistical tests were specified.

#### 4.3.2.7 Results

##### 4.3.2.7.1 Disposition

A total of 390 subjects were screened at 25 sites during the preliminary 2-week baseline period. There were 91 withdrawals, most due to failure to meet randomization criteria (34, 37%) followed by lack of reproducible lung function (30, 33%) and "other," including use of a proscribed medication (17, 19%), for a total of 299 eligible subjects.

The 299 subjects who completed the screening period were randomized and entered into the double-blind treatment phase of the trial, 73 into placebo, 73 into FP 100 mcg BID, 77 into FP 200 mcg QD, and 76 into BDP 168 BID. Eighty-four (28%) of these subjects discontinued prior to study endpoint, 48% in the placebo group, 22% in the FP 100 BID group, 17% in the FP 200 QD group, and 26% in the BDP 168 BID group. The reason(s) for discontinuation are given by the table below, the most common being lack of efficacy by pre-defined criteria (14% overall). Adverse events accounted for only three (1%) of the total study discontinuations. The category "other" included protocol violations, noncompliance, and prohibited medication use.

**SUBJECT DISPOSITION\***

	Placebo	FP100 BID	FP 200 QD	BDP 168 BID	Total
Enrolled	73	73	77	76	299
Completed	38	57	64	56	215 (72%)
Withdrawn	35	16	13	20	84 (28%)
Lack of Efficacy	19	5	7	10	41 (14%)
Adverse Event	2	0	1	0	3 (1%)
Other	14	11	5	10	40 (13%)

\* From Volume 95, Table 2, p.83 and p.46.

#### 4.3.2.7.2 Demographics and Other Baseline Characteristics:

Treatment groups were demographically similar. About 60% of enrollees were male, although adult asthmatics in this country are more likely to be female. The mean age was 33 years, with a range from 12 to 78 years. Most had never smoked (80%). As a group, they were overwhelmingly Caucasian (88%) with a vanishingly small representation of Black and Latino subjects, comprising 5% and 4% overall, respectively. One of the four treatment groups, FP 100 BID, was 97% Caucasian.

*Reviewer's Comment: It is frustrating to this reviewer to continue to see sparse representation by ethnic minorities, who bear a disproportionate burden of asthma in this country. The lack of diversity in this pivotal trial is particularly glaring, with one-treatment arm comprised of over 97% Caucasian subjects. While there are no data to suggest important differences in responses to asthma treatments such as inhaled CS between ethnic groups, a specific argument that there is no evidence that the safety or efficacy of Flovent Diskus differs between ethnic groups is circular, because none of the clinical trials submitted in this NDA have had sufficient power to detect such a difference.*

Asthma histories were similar. Over half of each group reported a duration of asthma in excess of 15 years. Newly diagnosed asthmatics (duration  $\leq 1$  year) comprised  $<1\%$  of the total enrollees. Ninety-three percent (93%) reported no ER visits and 98% reported no hospitalizations in the prior 12 months. FEV<sub>1</sub> values were about 68% of predicted at baseline and comparable across treatment groups.

Concurrent anti-asthma medication included the inhaled  $\beta$ -agonist albuterol (Ventolin), taken by 100% of subjects, as specified in the protocol. Theophylline was taken by 18%, 19%, 21%, and 18% of the placebo, twice daily FP, once daily FP, and BDP groups, respectively. Daily doses and/or serum levels were not provided. Salmeterol could also be continued during the study period per protocol, but its' usage was somewhat less balanced between treatment arms than was theophylline. Salmeterol was taken by 8%, 12%, 14%, and 14% of the placebo, twice daily FP, once daily FP, and BDP groups, respectively (Vol.95; Table 6). Concurrent non-asthma medications and related medical conditions were not appreciably different between the four groups (Tables 7-9; Vol.95), with allergic or atopic disorders heading the list.

**BACKGROUND CHARACTERISTICS\***

	Placebo	FP 100 BID	FP 200 QD	BDP 168 BID	Total
<b>Number</b>	73	73	77	76	299
<b>Gender:</b>					
<b>Female</b>	37%	42%	43%	43%	124 (41%)
<b>Male</b>	63%	58%	57%	57%	175 (59%)
<b>Ethnicity:</b>					
<b>Black</b>	4	1	5	5	15 (5%)
<b>Latino</b>	2	1	4	4	11 (4%)
<b>Caucasian</b>	65	71 (97%)	66	62	264 (88%)
<b>Other</b>	2	0	2	5	9 (3%)
<b>Age (yrs):</b>					
<b>Mean (range)</b>	33 (12-78)	33 (12-65)	35 (13-71)	30 (12-70)	33 (12-78)
<b>Smoking history:</b>					
<b>Never smoked</b>	84%	71%	77%	88%	239 (80%)
<b>Former smoker</b>	16%	29%	23%	12%	60 (20%)
<b>Asthma Duration:</b>					
<b>&lt; 15 years</b>	47%	42%	39%	39%	125 (42%)
<b>≥ 15 years</b>	53%	58%	61%	61%	174 (58%)
<b>ER visits (one yr) :</b>					
<b>0</b>	89%	97%	92%	95%	93%
<b>≥3</b>	1%	0	1%	0	<1%
<b>FEV<sub>1</sub> at Baseline:</b>					
<b>Liters (SE)</b>	2.62 (0.07)	2.60 (0.07)	2.58 (0.07)	2.53 (0.07)	
<b>% Predicted</b>	68.52%	68.79%	67.72%	68.43%	

\* From Tables 3, 4, and 5; vol.95, pp.85-89

## 4.3.2.7.3 Efficacy Analysis

## 4.3.2.7.3.1 Populations and Compliance

The population analyzed included all 299 subjects who received at least one dose of study medication (the ITT population). A subset analysis was performed using the 283 subject "efficacy population," comprised of the ITT subjects minus 16 subjects excluded because of a post hoc determination that they had not met inclusion/exclusion criteria. Data from 9 additional subjects were "partially excluded" because of protocol violations, also found post hoc. This review will only consider the ITT population in the efficacy analysis.

Compliance rates were defined as the percent of scheduled doses used from study drug dispensed at each visit. The study drug compliance rate for both devices was determined for Visits 2-10 based on blister count. MDI compliance could not be directly determined other than from diary data. Mean compliance rate exceeded 100% by these criteria for all four groups.

*Reviewer's Comment: Compliance rates in excess of 100% by blister count suggests some form of non-compliance (over-dosing) or more likely, a problem with the device resulting in a dose being actuated or a blister pocket being perforated, but the dose not being properly delivered, requiring re-administration by the subject.*

4.3.2.7.3.2 Primary Efficacy Variable: FEV<sub>1</sub>

Mean AM pre-dose FEV<sub>1</sub> was calculated for each treatment group at baseline and compared to mean AM pre-dose FEV<sub>1</sub> for each at end-point. Comparisons were made as mean FEV<sub>1</sub>, mean absolute change in FEV<sub>1</sub>, percent change in FEV<sub>1</sub>, and change in percent predicted FEV<sub>1</sub>. An F-test for overall treatment effect was performed prior to any pair-wise statistical comparisons. The last-value-carried-forward principle was used to calculate endpoint FEV<sub>1</sub> for each treatment group, to avoid bias introduced by the dropout of "sicker" patients, especially among the placebo subjects.

The results of this analysis are shown in the table below and in the attached Figure 3 (p.79; Vol.95). There was no significant difference in FEV<sub>1</sub> at baseline across treatment groups, which was 2.62L for placebo, 2.60L for FP 100 mcg BID, 2.58 L for FP 200 mcg QD, and 2.56L for BDP 168 mcg BID. At endpoint, there was a statistically significant treatment effect overall (p=0.001). Pair-wise comparisons between placebo and each of the three treatment groups showed statistical significance for both BID regimens, but not for FP once daily, although the p-value was close (see table below). Inspection of the mean change from baseline in FEV<sub>1</sub> showed a substantial numerical difference at endpoint between FP once daily and each of the two BID arms, 0.49L for FP 100 BID and 0.48L for BDP 168 BID, compared to 0.37 L for FP 200 QD and 0.21L for placebo. The pair-wise comparison between BDP and once daily showed the difference to be significant (p=0.048), although not between FP once daily and FP twice daily (p=0.112). The same analysis performed using the efficacy population was not substantially different. Both FP 100 mcg BID and BDP 168 mcg BID were shown to be efficacious using this primary endpoint, but once daily was not. The difference between FP once daily and BDP twice daily was again statistically significant.

MEAN CHANGE FROM BASELINE IN FEV<sub>1</sub> (L): ITT\*

	Placebo	FP 100 BID	FP 200 QD	BDP 168 BID	p vs. Placebo		
					FP100	FP200	BDP
N	73	73	77	76			
Baseline FEV <sub>1</sub> (L)	2.62	2.60	2.58	2.56	0.806**		
Mean change at Endpoint (L)	0.21	0.49	0.37	0.48	<0.001	0.054	0.002
% change at Endpoint	8.09%	19.39%	14.40%	18.93%	<0.001	0.060	0.002
Mean change in % Predicted	5.28	12.77	9.45	12.33	<0.001	0.048	0.001

\* Intent-to-Treat Population; From Tables 11-15; vol.95

\*\*Overall (F-test)

Figure 3 shows the mean change in FEV<sub>1</sub> over time, and Table 11 (Vol.95; p.99-100) shows the mean numerical value of FEV<sub>1</sub> at each clinic visit. All three treatment arms showed initial rapid improvement over the first week, however, a significant change in FEV<sub>1</sub> was not achieved again until Week 6, and then not again until Week 12, although the two BID arms were very close. The once daily arm never achieved a significant separation from placebo, an observation that can be seen graphically on Figure 3.

*Reviewer's Comment: Interpreting these mean week-by-week changes in FEV<sub>1</sub> clinically, an inhaled CS-naïve asthmatic started on FP 200 mcg once daily would reach the same improvement in FEV<sub>1</sub> at 12-weeks that a patient started on FP 100 mcg BID had reached at 6-weeks. Also, the improvement in lung function with once daily dosing is so "close to the line" in efficacy that the question of a patient's long-term stability at this dosing frequency arises. This reviewer questions whether it is valid or satisfactory to extrapolate data from a 12-week trial to answer this clinically relevant question.*

*There is one non-clinical question with regard to Flovent Diskus that could have clinical repercussions, and that is this issue will be more fully addressed in the CMC Review, there is a suggestion on Figure 3 of a "fall-off" in efficacy at the 4-week "life-of-device" time-point for both once daily and twice daily FP that is not seen with BDP. It may be argued whether this apparent tail-off in efficacy is real, and if so, whether it is related to stability issues.*

#### 4.3.2.7.3.3 Secondary Endpoint: Survival in Study

There was a significant overall treatment effect on duration of study participation using the Logrank test on Kaplan-Meier estimates of survival ( $p=0.005$ ; see attached Figure 4; p.80; Vol.95). By the end of the study, 19 subjects (26%) in the placebo group had discontinued for lack of efficacy compared to 5 (7%) in the FP 100 BID group, 7 (9%) in the FP 200 QD group, and 10 (13) in the BDP group. Pair-wise comparisons of survival-in-study between placebo and each of the two FP arms were statistically significant ( $p=0.007$  and  $p=0.008$  for twice daily and once daily, respectively). There was no significant difference in survival between the two FP arms. The pair-wise comparison between placebo and BDP approached but did not achieve statistical significance.

#### 4.3.2.7.3.4 Secondary Endpoint: Diary PEFr

Mean AM PEFr, PM PEFr, and AM/PM PEFr differential were averaged weekly from diary card records of PEFr measured by subjects twice daily: before the AM dose of study medication and again after the PM dose. The change from baseline was calculated for each of these three variables at all post-randomization clinic visits and at endpoint (Tables 16-20; Vol.95).

*Reviewer's Comment: The sponsor did not specify the minimum number of AM or PM PEFR diary entries recorded over the course of a given week that would be required before the data could be considered "evaluable."*

Baseline AM PEFRs were similar across treatment groups at baseline, 425-434 L/min (see table, below). There was a statistically significant treatment effect at study endpoint ( $p=0.002$ ) as well as significant pair-wise treatment comparisons between placebo and each of the three treatment groups at endpoint ( $p=0.005$ ,  $<0.001$ ,  $<0.001$  for FP twice daily, FP once daily, and BDP, respectively). The improvement from baseline was numerically greater for the two twice daily dosing groups (31 L/min for both FP twice daily and BDP twice daily) than for the FP once daily group (27 L/min). There was no significant difference between any two treatment groups, however.

The mean change from baseline in diary PM PEFR followed a pattern similar to diary AM PEFR (see table below). Baseline values were comparable between treatment groups and slightly higher than AM PEFR values. Net improvement over time was more modest than for AM PEFR, with the final change from baseline to endpoint being 25 L/min for FP 100 BID, 26 L/min for FP 200 QD, and 27 L/min for BDP twice daily, compared to 5 L/min for placebo. The overall treatment effect was significant at endpoint ( $p=0.004$ ), as were each of the pair-wise comparisons between active treatment and placebo (see table below).

#### CHANGE FROM BASELINE TO ENDPOINT IN AM/PM PEFR\*

	Placebo	FP 100 BID	FP 200 QD	BDP 168 BID	p vs. placebo		
					FP100	FP200	BDP
N	73	73	77	76			
Baseline AM PEFR (L/min)	428	434	434	425			
$\Delta$ AM PEFR	1	31	27	31	0.005	<0.001	<0.001
N	70	72	77	75			
Baseline PM PEFR (L/min)	449	454	452	444			
$\Delta$ PM PEFR	5	25	26	27	0.009	0.001	<0.001
N	73	73	77	76			
Baseline AM/PM PEFR Differential (L/min)	20	20	17	19			
$\Delta$ AM/PM PEFR Differential	3	-7	-4	-4	0.278	0.287	0.313

\* Tables 21-24; Vol. 95. ITT population

The AM/PM differentials for each subject were calculated at the various time-points by subtracting each AM PEFR from the previous evening's PM PEFR. A high AM/PM differential is considered indicative of asthma instability. These data are shown in Tables 19 and 20 (Vol.95). Mean change from baseline to endpoint in AM/PM differential is shown in the

table above. There was a numerical decrease in AM/PM differential in all active treatment groups compared to placebo, greater for twice daily FP than for once daily or for BDP. However, there was no overall statistically significant treatment effect at endpoint, nor were any of the pair-wise comparisons with placebo significant.

#### 4.3.2.7.3.5 Secondary Endpoints: Symptom Scores, Nighttime Awakenings, and Rescue Ventolin Use

Subjects recorded their asthma-related symptoms daily on their diary cards using a 0-3 severity scale, as described earlier in this review. Using this scale, symptoms were similar and relatively mild at baseline across treatment groups, all being approximately 1.00. At endpoint, there was a statistically significant treatment effect overall ( $p=0.043$ ; see table below). The pair-wise comparison of each treatment arm with placebo was significant at endpoint for twice daily FP and BDP, but not for once daily FP. There was no significant difference between any two active treatment arms.

Nighttime awakenings requiring Ventolin were also infrequent and similar across treatment groups at baseline, ranging from approximately one night in ten for each group (placebo, 0.11; FP BID, 0.10; FP QD, 0.09; BDP, 0.11). At study endpoint, there was no statistically significant treatment effect overall, nor were any of the pair-wise comparisons with placebo significant (see table below).

Use of rescue Ventolin was to be recorded daily in the diary as number of puffs of the MDI used. At baseline, daily use of Ventolin was similar between treatment groups, approximately 2 ½ puffs per day. At study endpoint, all three active treatment arms had succeeded in reducing their daily Ventolin requirements by approximately one puff per day. In contrast, the placebo group had a numerical increase in Ventolin requirements (+0.22 puffs/day). The change was greatest for the twice daily FP group (-1.07 puffs/day), followed by the BDP group (-0.90 puffs/day) and the once daily FP group (-0.82 puffs/day; see table below). There was a statistically significant treatment effect for this parameter when measured at study endpoint ( $p=0.002$ ). The pair-wise comparisons between placebo and each of the three active treatment groups were also significant at study endpoint. There was no statistical difference between any two active treatment arms, however.

**APPEARS THIS WAY  
ON ORIGINAL**

**CHANGE FROM BASELINE IN DIARY VARIABLES (ITT)\***

	Placebo	FP 100 BID	FP 200 QD	BDP 168 BID
N	73	73	77	76
Asthma symptom score:				
Baseline	1.07	1.07	1.03	1.05
Change	-0.12	-0.40	-0.37	-0.38
p-value**		0.016	0.063	0.013
Nighttime Awakenings:				
Baseline	0.11	0.10	0.09	0.11
Change	-0.03	-0.06	-0.06	-0.06
p-value**		0.300	0.065	0.112
Ventolin use (puffs/day)				
Baseline	2.66	2.46	2.25	2.67
Change	0.22	-1.07	-0.82	-0.90
p-value**		0.004	0.019	<0.001

\* From Tables 22-24; Vol.95.

\*\* Compared to placebo

**4.3.2.7.3.6 Efficacy by Demographic Subgroups**

There was no indication that a difference in the primary endpoint existed by gender subgroup. Representation by non-Caucasian subjects was too low to make any determination for these ethnic subgroups. With regard to age, most subjects were between 18 and 64 years, with adolescents age 12-17 constituting only 25 out of 299 (8%) enrollees, and geriatric subjects > 64 years comprising only 7 of 299 (2%). Again, there did not appear to be any difference between these subgroups and the ITT population of the primary endpoint, but numbers were extremely small (see "Results: Demographics; also Tables ST 12-19; Vol.95).

**4.3.2.7.4 Safety Results****4.3.2.7.4.1 Extent of Exposure**

A total of 299 patients received at least one dose of study medication and therefore have been included in the safety analysis. Their extent of exposure is shown in the table below. On average, the patients who received active treatment were exposed for approximately 75-78 days out of an 84-day trial. The placebo patients received approximately 10 fewer days of exposure.

**EXTENT OF EXPOSURE TO STUDY MEDICATION\***

	Placebo	FP 100 BID	FP 200 QD	BDP 168 BID
Number				
Baseline	73	73	77	76
Completed	38 (52%)	57 (78%)	64 (83%)	56 (74%)
Exposure(days):				
Mean	64.8	74.3	78.7	74.3
Median	83.0	85.0	85.0	85.0

\* Table 25 and p.58; Vol.95

#### 4.3.2.7.4.2 Adverse Events (AE)

The adverse events identified in this trial are not substantially different from those reported in the ADVERSE REACTIONS section of the approved product labeling for Flovent™ Rotadisk. These common adverse events will therefore not be discussed in great detail in this review.

Overall, 64% of the placebo group reported at least one adverse event during this trial, which was comparable to the active-treatment groups, 63% of FP 100 BID, 62% of FP 200 QD, and 61% of the BDP BID group. By organ system, the most commonly reported AE's in all treatment groups were within the ENT system (38-47%) followed by GI (12-21%), Neurologic (8-19%), and Lower Respiratory (8-15%). In descending order of frequency, the top ENT AE's were URTI (13-25%), throat irritation (12-21%), upper respiratory inflammation (4-7%) and nasal congestion (3-7%).

Among the AEs which were more common in the FP-treated subjects were throat irritation, 12% of the placebo group compared to 16% of the FP 100 BID group, 16% of the FP 200 QD group, and 21% among the BDP group; headaches, occurring in 5%, 15%, 16%, and 14% of placebo, FP 100 BID, FP 200 QD, and BDP QD patients, respectively; dysphonia, reported by 0%, 1%, 1%, and 5% of placebo, FP 100 BID, FP 200 QD, and BDP QD patients, respectively; and cough, reported by 1%, 5%, 4%, and 0% of the placebo, FP 100 BID, FP 200 QD, and BDP QD patients, respectively.

Events more commonly reported among once daily compared to twice daily inhaled CS users, or vice versa, include URTI, more common among the once daily CS users (reported by 25% of FP 200 QD patients compared to 16% of the FP 200 BID group and 13% of the BDP group) and Musculoskeletal system events, primarily muscle and joint pains, more common among the twice daily users (occurring in 3% of the FP once daily group compared to 12% of the FP 100 BID group and 5% of the BDP group). Local CS effects such as dysphonia or thrush did not appear to differ between the once and twice daily groups.

Events of particular interest include oropharyngeal candidiasis or candidiasis unspecified site, which was reported for 0 placebo subjects, 2 FP 100 BID subjects, 2 FP 200 QD subjects, and 2 BDP subjects, respectively. Not unexpectedly, there were no reports of cataracts, glaucoma, or osteopenia in this 12-week trial. No adverse event specifically coded as "HPA axis suppression" was reported.

When analyzed by demographic subgroups, there were no apparent differences in the overall number or nature of AEs based upon gender, age or ethnicity. The number of non-Caucasian subjects was very small, however, <12% of the ITT. Likewise, there were few enrollees at each end of the age spectrum, 25 who were between 12-17 years and 8 who were >64 years.

There were no deaths during the double-blind phase of this study. Four patients experienced serious AEs and three patients were withdrawn due to AEs. One SAE occurred during screening (diabetes mellitus), and the subject was dropped from the study. The other three SAEs included one placebo patient with asthma exacerbation and bilateral pneumothoraces (this was also a discontinuation), one subject in the FP 100 BID group who experienced left salivolithiasis, and one patient in the BDP group who became pregnant at study Week 10 and miscarried at Week 12. The additional two dropouts due to AEs included one placebo subject with a rash and one FP 200 QD subject with palpitations.

#### 4.3.2.7.4.3 Laboratory Data (excluding HPA-axis)

Blood samples for serum chemistry, LFT's, and hematology were obtained at baseline and at study endpoint. One subject was withdrawn for an abnormal laboratory test, an elevated GGTP after 6 days on BDP. The event was not coded as an AE. No further details are available.

A few subjects (1-3% per group, maximum) had "clinically significant" laboratory values by pre-specified criteria reported at any time post-randomization (Table 29; Vol.95). These are summarized by test in Table 32 and specific values appear in Table ST-30 (Vol.95). A few more patients had laboratory values outside of the normal range, many of which were probably chance variation expected among a large group of patients. These are summarized in "shift" tables, and appear in Table 31 (Vol.95). Abnormalities of relevance to this review, either because of known side-effects of CS or because of post-marketing surveillance, would include bicarbonate, potassium, glucose, eosinophil count, and alkaline phosphatase. These have been separately noted below.

There were no reported "clinically significant" (>40 meq/L) elevations in bicarbonate. There were no reported "clinically significant" decreases in potassium (<3.0). The shift tables (Tables 30-31; Vol.95) similarly showed no high abnormal values for bicarbonate nor low abnormal values for potassium.

There were two reported cases of "clinically significant" elevations in plasma glucose (>175 mg/dL), one occurring in a placebo subject and the other in a patient in the FP 100 BID group. This latter patient (7302) had a normal glucose at baseline (113 mg/dL; nml range 65-115 mg/dL), and an elevated reading of 238 mg/dL at endpoint. The latter was reportedly drawn 6 days after discontinuation of study drug, however. Follow-up glucose drawn one month later remained abnormal (124 mg/dL), but was not in the clinically significant range. Shift tables showed 5 patients whose glucose went from normal to elevated during the trial, 3 in placebo and 2 in FP 200 QD. No further information is available about these patients.