CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-279

STATISTICAL REVIEW(S)

STATISTICAL REVIEW AND EVALUATION
CLINICAL STUDIES

Date	SEP 2 5 2001
NDA #	21-279
Applicant	Novartis
Name of Drug	Foradil™ (Formoterol fumarate dry powder capsules) oral inhalation via Aerolizer™
Indication	Treatment of patients with chronic obstructive pulmonary disease (COPD)
Document Reviewed	❑ Vol. 1, Sponsor's cover letter dated 9/22/2000 ❑ Vols. 14-38 (Clinical Trial Report) ❑ Data submission dated 9/22/2000
Statistical Reviewer	Ted J. Guo, Ph.D., Div II/OEB, HFD-715
Medical Input	Eugene Sullivan, MD, Division of Pulmonary Drug Products (ODE II, HFD-570)
Key Words	NDA, Clinical Studies, Foreign clinical data

Summary

The review of this NDA was based on evaluation of two multinational studies, Studies 056 and 058. In Study 056, US patients, included in 4 (out of 44, or 9.1%) trial centers, comprised 6.4% (50 out of 780) of all patients. In Study 058, US patients, included in 13 (out of 57, or 22.8%) trial centers, comprised 23.3% (199 out of 854) of all patients. This reviewer reached the following conclusions:

- In Studies 056 and 058, Foradil at 12 and 24 μg b.i.d. was superior to the placebo control.
- In both studies the differences in AUC of FEV_1 between 12 and 24 μg b.i.d. of Foradil were not statistically significant.
- In Study 056, Foradil at 12 and 24 μg b.i.d. was statistically superior to Ipratropium MDI 40 μg q.i.d., which was also shown to be statistically superior to the placebo.
- In Study 058, Foradil at 24 μg b.i.d., but not at 12 μg b.i.d., proved to be statistically superior to theophylline, which, as an open-labeled treatment, failed to demonstrate its superiority to the placebo.
- In Study 056, subjects receiving Foradil at 12 μg b.i.d. appeared to have a greater average AUC value of FEV_1 than did those patients receiving Foradil 24 μg b.i.d. However, in US patients, there was a greater FEV_1 response to Foradil 24.
- The lack of statistical significance favoring one dose to another suggests no particular preference for the dosage of Foradil.
- Descriptive comparison of AUC (of FEV_1) data appears to indicate that the treatment effects were somewhat different for the U.S. and non-U.S. patients in the studies.

In conclusion, the effectiveness of Foradil at 12 and 24 μg b.i.d. was demonstrated in Studies 056 and 058 as whole. However, a visual comparison of AUC (of FEV_1) between U.S. and non-U.S. patients suggested that Foradil is not as effective for the U.S. patients as for the non-U.S. patients.

Table of Contents

List of Tables

List of Figures

Introduction

Foradil™ is indicated for patients with chronic obstructive pulmonary disease (COPD). The proposed — : 12 — b.i.d., delivered via Aerolizer™. To support the efficacy claim for Foradil™, the sponsor, Novartis, submitted two placebo-controlled clinical studies (Table 1). Foreign clinical data comprised about 85% of the data in the two studies combined.

Table 1. Placebo-Controlled Studies Submitted

Clinical Study	Location	Type	Study Duration
Study 056	Multinational	Phase III	Nov. 1997-Apr. 1999
Study 058	Multinational	Phase III	Feb. 1997-Jun. 1999

The studies, Studies 056 and 058 had a similar design. In each study, the patient started with a 10- to 21-day run-in period prior to randomization. During the run-in period, patients received placebo with albuterol as rescue medication. In Study 056, a randomized patient was assigned to one of the following treatment groups: Foradil 12 μg, Foradil 24 μg, Ipratropium bromide 40 μg, or placebo; in Study 058, the positive control was open-label theophylline. Serial FEV_1 was measured hourly at the visit days. For the purpose of regulatory approval, the evaluation of Foradil is chiefly based on Studies 056 and 058. These studies were multinational studies.

Table 2 and Table 3 show the number of centers in the participating countries for Studies 056 and 058, respectively. Based on these tables, U.S. accounted for 17% (17/101) of all centers in two studies combined. Further below, Tables 4 and 5 show the number of patients in participating countries for Studies 056 and 058, respectively. Based on these tables, U.S. accounted for 15% (249/1634) of all patients in the two studies combined.

Table 2. Number of centers: Study 056

Country	N	%
-All-	44	100.00%
AUSTRALIA	6	13.64%
BELGIUM	1	2.27%
CANADA	5	11.36%
DENMARK	5	11.36%
FINLAND	3	6.82%
GREAT BRITAIN	2	4.55%
HOLLAND	10	22.73%
NORWAY	3	6.82%
POLAND	2	4.55%
RUSSIA	3	6.82%
USA	4	9.09%

Source: Reviewer's analysis data set spi56tab.sd2

Table 3. Number of centers: Study 058

Country	N	%
-All-	57	100.00%
AUSTRIA	1	1.75%
BELGIUM	1	1.75%
CZECH REPUBLIC	4	7.02%
FRANCE	3	5.26%
GERMANY	2	3.51%
GREECE	4	7.02%
HUNGARY	4	7.02%
ITALY	12	21.05%
S. AFRICA	6	10.53%
SLOVAKIA	3	5.26%
SPAIN	4	7.02%
USA	13	22.81%

Source: Reviewer's analysis data set spi58tab.sd2

Table 4. Number and percentages of patients by country: Study 056

Country	Count	Pct.
-ALL-	780	100.00
AUSTRALIA	83	10.64
BELGIUM	18	2.31
CANADA	78	10.00
DENMARK	122	15.64
FINLAND	42	5.38
GREAT BRITAIN	30	3.85
HOLLAND	181	23.21
NORWAY	46	5.90
POLAND	56	7.18
RUSSIA	74	9.49
USA	50	6.41

Source: Reviewer's analysis data set spi56rec.sd2

Table 5. Number and percentages of patients by country: Study 058

Country	Count	Pct.
-ALL-	854	100.00
AUSTRIA	15	1.76
BELGIUM	2	0.23
CZECH REPUBLIC	68	7.96
FRANCE	42	4.92
GERMANY	26	3.04
GREECE	66	7.73
HUNGARY	61	7.14
ITALY	163	19.09
S. AFRICA	71	8.31
SLOVAKIA	73	8.55
SPAIN	68	7.96
USA	199	23.30

Source: Reviewer's analysis data set spi58rec.sd2

Sponsor's Analysis

Overview of Study 056

Study 056 was a 12-week double blind, placebo-controlled, parallel-group, multi-center study. The objective of the study was to assess the AUC of FEV_1 after 12 weeks of treatment in patients with COPD (p.12, vol. 14). Following a 21-day run-in period, the qualified patient was randomized to one of the following 4 treatment groups: Formoterol 12 μg b.i.d., Formoterol 24 μg b.i.d., ipratropium MDI 40 μg q.i.d., or placebo. Clinic visits were scheduled at days 28, 56, and 84 after the randomization (p.17, vol. 14).

The primary efficacy variable was the AUC of FEV_1 over the 12-hour spirometry test after 12 weeks of the treatment (p. 13, vol. 14).

Description of Study Plan (Study 056)

Table 6 highlights the characteristics of this study.

Table 6. Characteristics of Study 056

Study	General Feature	Specific Characteristics
Protocol 056 (Treating COPD) (p. 1, vol. 14)	12-week study	The study began with a 21-day baseline period followed by a 12-week treatment period: 11/26/97–4/23/99.
	Randomized	
	Double-blind	
	Double-dummy	
	Parallel-group	Formoterol 12 μg b.i.d., Formoterol 24 μg b.i.d., ipratropium MDI 40 μg q.i.d., or placebo.
	Multi-center	44 centers
	Primary efficacy variable: AUC of FEV1	AUC of FEV_1 after 12 weeks of treatment in patients with COPD

To summarize and evaluate the sponsor's findings and conclusions, this reviewer reanalyzed the sponsor's data and made part of his results on CDER Intranet using SAS/IntrNet. This tool produced most of the tables and graphs in this report.

Table 7 accounts for patients based on their study disposition. According to the sponsor, a total of 849 patients (based on data submitted) were randomized[1]. Among the 849 patients, 780 were identified as Intent-To-Treat (ITT)[2]; 698, a subgroup of the ITT patients, were evaluable ("acceptable") patients[3]. The sponsor also classified the "completers," but, in fact, there were no differences in numbers between the evaluable and patients and those who completed the study. In this review, the words, "evaluable" and "acceptable" are used interchangeably.

Table 7. Patient Counts (Study 056)

	Evaluable				Total	
	No		Yes			
	N	%	N	%	N	%
Treatment						
CtrlPbo	29	14.50	171	85.50	200	100.00
F12	13	6.70	181	93.30	194	100.00
F24	23	11.98	169	88.02	192	100.00
IPR	17	8.76	177	91.24	194	100.00
Total	82	10.51	698	89.49	780	100.00

Source: Reviewer's analysis data set spi56tab.sd2

Reviewer's Comment:

There are a few discrepancies in center counts between the above table and the sponsor's report (p. 12, vol. 52). Table 8 shows the differences in the number of centers by country based on pre-, post- randomization data and the sponsor's NDA. Please note the differences in center counts between column 1 and column 3. However, in general, the data submitted appear to be acceptable for review.

Table 8. Number of centers by country (Study 056)

Country	Before randomization, No centers were pooled	After randomization, Some centers were pooled	Center counts reported in the NDA (Study 056)
Australia	9	6	10
Belgium	3	1	4
Canada	7	5	7
Denmark	5	5	5
Finland	3	3	4
Great Britain	6	2	9
Norway	4	3	5
Holland	13	10	14
Poland	3	2	5
Russia	3	3	3
USA	6	4	6
	Total: 62	Total: 44	Total: 72

Source: Reviewer's analysis data set spi56_d, spi56tab

[1] The number of patients before randomization was approximately 824 (in fact, 849 based on data submitted–Reviewer), based on NDA p. 18, vol. 14.
[2] The ITT patients comprised all randomized patients who received at least one dose of medication.
[3] The "acceptable" patients were those ITT patients who completed the treatments and had at least six hours of washout period of rescue medication before Visit 5.

Sponsor's Statistical Methods (Study 056)

The following points highlight the sponsor's statistical method and analysis setup:

- The sponsor's statistical analysis was based on the ITT patients. Those were the randomized patients "who received at least one dose of trial medication (p.36, vol. 14)."
- "FEV_1 AUC over 12 hours, standardized with respect to time, after 3 months of treatment (Visit 5) was analyzed for the ITT population (p. 45, vol. 14)."
- "An ordered test procedures in pre-specified order was applied in order to deal with multiple testing without an adjustment of the significance level (p.37, vol. 14)." The ordered tests, pre-specified by the sponsor, were:
    - Test the hypothesis that there is no difference between the treatment with Formoterol 24 µg and placebo.
    - Test the hypothesis that there is no difference between the treatment with Formoterol 12 µg and placebo.
    - Test the hypothesis that there is no difference between the treatment with Formoterol 24 µg and ipratropium bromide.
    - Test the hypothesis that there is no difference between the treatment with Formoterol 12 µg and ipratropium bromide.
    - Test the hypothesis that there is no difference between the treatment with Formoterol 24 µg and Formoterol 12.
    - Test the hypothesis that there is no difference between the treatment with ipratropium bromide and placebo (p. 254, vol. 17).

  The sponsor detailed in the protocol, "In order to control the Type I error, a hierarchy of testing was set up with the contrasts tested in the order described... No contrast was to be considered statistically significant unless each preceding contrast examined within that family of contrasts was also statistically significant (p. 254, vol. 17)." The p-value cutoff point was set to 0.05.
- "Analysis of covariance was carried out to estimate all treatment differences for normalized AUC in FEV_1." "The statistical fixed-effects model considered country, center within country, sex, reversibility, smoking status at Visit 2 and treatment as main effects. In addition, baseline FEV1 (last measured FEV_1 value before treatment) was used as a covariate.

Under the above consideration, the sponsor produced the following statistical results.

Sponsor's Statistical Results (Study 056)

The sponsor's efficacy results are summarized in the following Table 9, in which the tests were performed in order from top to bottom, according to the protocol. The first two rows were tests for the primary objective. The analysis results shown in Table 9 were verified and confirmed by this reviewer.

Table 9. Mean differences in FEV_1 AUC over 12 hours after 3 months of treatment (ITT patients, Visit 5, Study 056)

Comparisons	Estimated mean difference in standardized[4] AUC of FEV_1	Two-sided 95% confidence intervals	Test p-values
Foradil 24 μg vs. placebo	0.194	0.145-0.243	<0.001
Foradil 12 μg vs. placebo	0.223	0.174-0.273	<0.001
Foradil 24 μg vs. ipratropium	0.057	0.007-0.106	0.024
Foradil 12 μg vs. ipratropium	0.086	0.037-0.136	0.001
Foradil 24 μg vs. Foradil 12 μg	-0.029	-0.079-0.020	0.245
Ipratropium vs. placebo	0.137	0.088-0.186	<0.001

Source: NDA Table 9-1, p. 46, vol. 14

Because of the concern from the reviewing medical officer, this reviewer paid special attention to the following analysis.

Analysis on patient-diary score as a secondary efficacy variable was based on rank statistics. The Cochran-Mantel-Haenszel (CMH) test was applied, though only the former was specified in the protocol. The summary of this analysis can be found on page 54, vol. 14, and details can be found in Appendix 5 Output 7 (p. 250, vol. 18). This reviewer noted the difference between the diary-score analysis based on entire treatment period and the analysis based on Visit 5, alone. The following Table 10 summarizes the sponsor's analysis base on rank statistics.

Table 10. Test based on rank statistics on patient-diary score (Study 056)

	Visit 5		Visit 3-5	
Comparison	P-value	Significant	P-value	Significant
Foradil 24 μg vs. placebo	0.159	No	0.007	Yes
Foradil 12 μg vs. placebo	0.004	Yes	0.000	Yes
Foradil 24 μg vs. Ipratropium	0.194	No	0.060	No
Foradil 12 μg vs. Ipratropium	0.010	Yes	0.009	Yes
Foradil 24 μg vs. Foradil 12 μg	0.300	No	0.419	No
Ipratropium vs. placebo	0.719	No	0.439	No

Based on the analysis of patient-diary scores (as a secondary efficacy variable), Foradil at 12 μg scored more often than Foradil at 24 μg comparing with their comparators. It appears that Foradil at 12 μg provided better improvement in patient-diary scores than Foradil at 24 μg.

The sponsor's analysis indicated that numerically Foradil at 12 μg performed better than did Foradil at 24 μg, although the difference is not statistically significant.

Sponsor's Conclusions (Study 056)

The sponsor's conclusion for the efficacy claim is quoted as follows.

[4] Based on model: FEV_1AUC=baselineFEV_1+country+center(country)+sex+smoke+reversibility+treatment

"Formoterol powder capsules for inhalation delivered via the Aerolizer™ device (Foradil® Aerolizer) was superior to placebo at two dose levels (12 and 24 μg) in patients with COPD with regard to FEV_1 AUC, primary outcome variable of this trial, with no reduction of effect over time (p.74, vol. 14)"

Reviewer's comment:

The effect of Foradil at 12 μg actually differs from Foradil at 24 μg over time, according to this reviewer's analysis. At the beginning of the treatment period (Visit 2), Foradil at both dosages showed little differences in terms of AUC of FEV_1. However, at Visit 5, Foradil at 24 μg demonstrated a greater effect than its lower dosage (12 μg). This finding is depicted in Figure 2 and Figure 3 below. The sponsor's above claim of "no reduction of effect over time" is overstated.

Figure 2. FEV_1 vs. time at Visit 2 (Study 056)

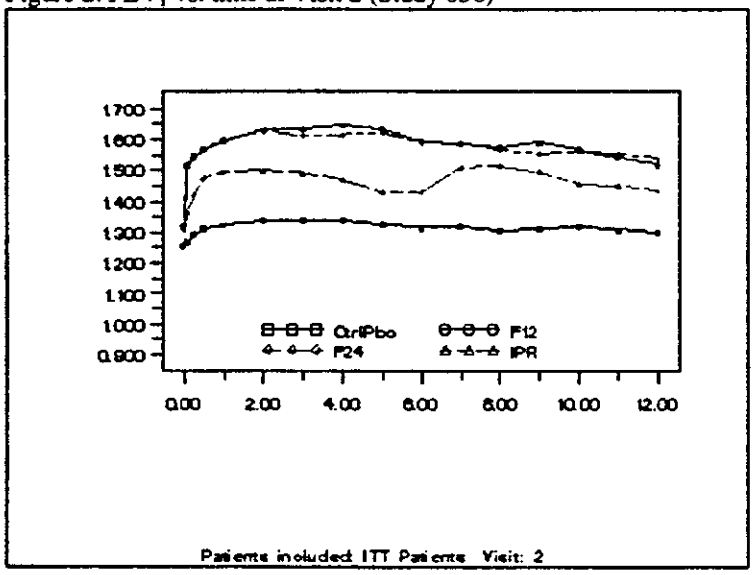
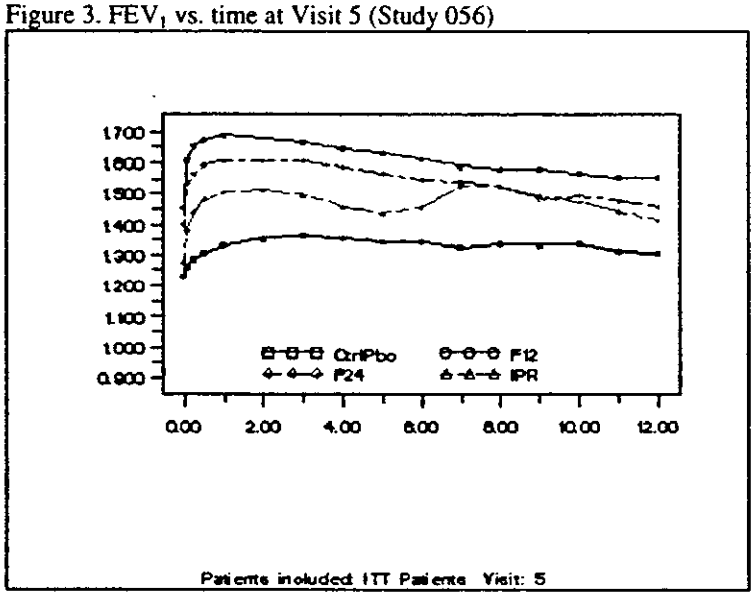


Figure 3. FEV_1 vs. time at Visit 5 (Study 056)



Overview of Study 058

Study 058 was a 12-month double blind, placebo-controlled, parallel-group, multi-center study. Following a 21-day run-in period, the qualified patient was randomized to one of the following 4 treatment groups: Formoterol 12 µg b.i.d., Formoterol 24 µg b.i.d., theophylline 200-400 mg, or placebo. Clinic visits were scheduled at weeks 3, 6, 9, and 12 months after the randomization (p.17, vol. 20). Although the trial period was 12 months, the efficacy assessment was based on the AUC of FEV_1 after 12 weeks of treatment in patients with COPD (p.17, vol. 20).

The primary efficacy variable was the AUC of FEV_1 over the 12-hour spirometry test, after 12 weeks of the treatment (p. 13, vol. 20).

Description of Study Plan (Study 058)

Table 11 highlights the characteristics of this study.

Table 11. Characteristics of Study 058

Study	General Feature	Specific Characteristics
Protocol 058 (Treating COPD) (p. 1, vol. 20)	12-week study	The study began with a 21-day baseline period followed by a 12-week treatment period: 11/26/97–4/23/99.
	Randomized	
	Double-blind	Note that theophylline was open-labeled.
	Double-dummy	
	Parallel-group	Formoterol 12 µg b.i.d., Formoterol 24 µg b.i.d., theophylline, or placebo.
	Multi-center	57 centers
	Primary efficacy variable: AUC of FEV1	AUC of FEV_1 after 12 weeks of treatment in patients with COPD

Table 12 accounts for the number of patients based on their disposition. This table describes the accountability of the patients. According to the sponsor, a total of 854 patients (based on data submitted) were randomized. All the 854 patients were identified as Intent-To-Treat (ITT)[5]; 723 were evaluable (acceptable) patients[6], comprising a subgroup of the ITT patients.

[5] The ITT patients comprised all randomized patients who received at least one dose of medication.
[6] The Acceptable patients were those ITT patients who completed the treatments and had at least six hours of washout period of rescue medication before Visit 3.

Table 12. Patient Counts (Study 058)

	Evaluable				Total	
	No		Yes			
	N	%	N	%	N	%
Treatment						
CtrlPbo	35	26.72	185	25.59	220	25.76
F12	21	16.03	190	26.28	211	24.71
F24	18	13.74	196	27.11	214	25.06
THE	57	43.51	152	21.02	209	24.47
Total	131	100.00	723	100.00	854	100.00
	Complete				Total	
	No		Yes			
	N	%	N	%	N	%
Treatment						
CtrlPbo	59	25.43	161	25.88	220	25.76
F12	52	22.41	159	25.56	211	24.71
F24	40	17.24	174	27.97	214	25.06
THE	81	34.91	128	20.58	209	24.47
Total	232	100.00	622	100.00	854	100.00

Source: Reviewer's analysis data set spi58tab.sd2

Table 13 shows the differences in the number of centers by country based on pre-, post- randomization data, and the sponsor's NDA. Please note the differences in center counts between column 1 and column 3. US centers comprised about 23% (13/57) of the total centers (Table 3). In addition, U.S. patients comprised 23.3% (199/854) of the total patients (Table 5).

Table 13. Number of centers by country (Study 058)

	Before randomization, No centers were pooled	After randomization, Some centers were pooled	Center counts reported in the NDA (Study 058)
Austria	2	1	2
Belgium	1	1	1
Czech Republic	5	4	5
Germany	5	2	5
Spain	6	4	6
France	9	3	9
Greece	5	4	5
Hungary	4	4	4
Italy	16	12	16
Slovakia	3	3	3
USA	19	13	19
S. Africa	6	6	6
Total	81	57	81

Source: Reviewer's analysis data set spi58_d, spi58tab

Sponsor's Statistical Methods (Study 058)

The sponsor applied a pre-specified ordered test procedure, the same procedure for Study 056. The purpose of this procedure was to control the Type I error rate (p. 36, vol. 20).

Sponsor's Statistical Results (Study 058)

The sponsor's efficacy results are summarized in the following Table 14.

Table 14. Mean differences in FEV_1 AUC over 12 hours after 3 months of treatment (ITT patients, Visit 3, Study 058)

Comparisons	Estimated mean AUC of FEV_1, standardized[7]	Two-sided 95% confidence intervals	Test p-values
Foradil 24 μg vs. placebo	0.208	0.152-0.264	<0.001
Foradil 12 μg vs. placebo	0.200	0.144-0.257	<0.001
Foradil 24 μg vs. theophylline	0.092	0.034-0.151	0.002
Foradil 12 μg vs. theophylline	0.085	0.026-0.144	0.005
Foradil 24 μg vs. Foradil 12 μg	0.008	-0.048-0.063	0.787
Theophylline vs. placebo	0.116	0.056-0.176	<0.001

Source: NDA Table 9-1, p. 50, vol. 20

The analysis results shown in Table 14 were verified and confirmed by this reviewer.

Analysis of a secondary efficacy variable is noted here to for reviewing medical reviewer's information and for future reference. The analysis on patient-diary score as a secondary efficacy variable was based on rank statistics. The Cochran-Mantel-Haenszel (CMH) test was applied though only the former was specified in the protocol. The summary of this analysis can be found on page 60, vol. 20, and details can be found in Appendix 5 Output 7 (p. 286, vol. 25). The following Table 15 summarizes the sponsor's analysis based on rank statistics. The sponsor concluded, "There were no statistically significant results for any of the paired treatment contrasts at any visit. (p. 61, vol. 20)"

Table 15. Test based on rank statistics on patient-diary score (Study 058)

	Visit 3		Visit 3-6	
Comparison	P-value	Significant	P-value	Significant
Foradil 24 μg vs. placebo	0.691	No	0.379	No
Foradil 12 μg vs. placebo	0.080	No	0.094	No
Foradil 24 μg vs. Theophylline	0.460	No	0.185	No
Foradil 12 μg vs. Theophylline	0.140	No	0.090	No
Foradil 24 μg vs. Foradil 12 μg	0.237	No	0.630	No
Theophylline vs. placebo	0.971	No	0.759	No

Sponsor's Conclusions (Study 058)

The sponsor's conclusion for the efficacy claim is quoted as follows.

"Formoterol at both 24 μg and 12 μg b.i.d. produced statistically and clinically significant improvement in lung function, as measured by FEV_1 area under the curve, when compared to placebo after 12 weeks of treatment. The estimated improvement was 208mL and 200mL for 24 μg and 12 μg b.i.d., respectively. Formoterol at both doses was also statistically significant when compared to theophylline. (p.13, vol. 20)"

[7] Based on model: FEV_1AUC=baselineFEV_1+country+center(country)+sex+smoke+reversibility+treatment

Reviewer's Evaluation of Study 056

As a first step, this reviewer examined the completeness of the patient data.

Table 16 shows the percentages of patients who stayed on study by treatment and by visit. A reasonably large proportion (ranging from approximately 89% to 94%) of the patients completed the study. The placebo-treated patient group had the lowest percentage of patients (88.5%) who completed the study. The percentage of completers among the U.S. patients was even lower (76.9%).

Table 16. Numbers and percentages of patients who stayed in the study (Study 056)

Treatment	Visit	#Patients	#Total Patients	Percent Completed
CtrlPbo	2	200	200	100.0%
CtrlPbo	5	177	200	88.5%
F12	2	194	194	100.0%
F12	5	183	194	94.3%
F24	2	192	192	100.0%
F24	5	174	192	90.6%
IPR	2	194	194	100.0%
IPR	5	177	194	91.2%

Source: Reviewer's analysis data set spi56rec.sd2

Table 17 describes the mean values of FEV_1 at baseline for the treatment groups. The baseline is defined as the pre-treatment FEV_1 values at Visit 2. The overall difference in FEV_1 appears to be small among the treatment groups (p=0.3493).

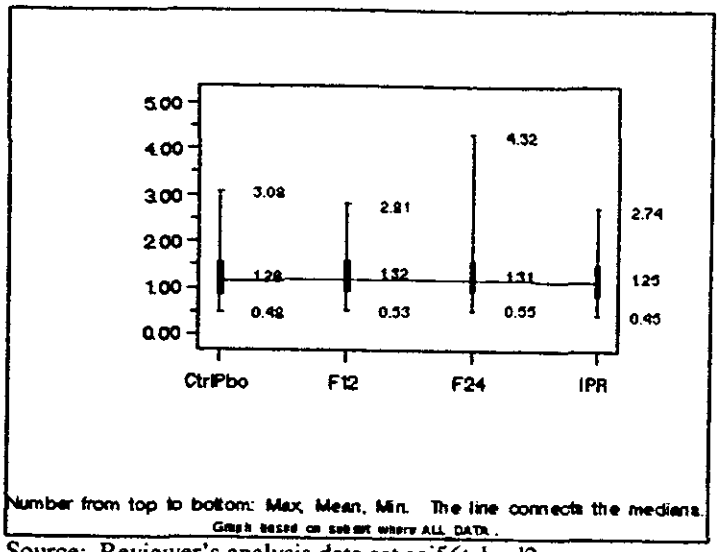
Table 17. Baseline FEV_1 (Study 056)

Treat	N	MEAN	STD	MIN	MAX
CtrlPbo	200	1.2552	0.4671	0.4800	3.0800
F12	194	1.3216	0.4992	0.5300	2.8100
F24	192	1.3111	0.5144	0.5500	4.3200
IPR	194	1.2517	0.4712	0.4500	2.7400
Test for Group Diff: F				DF	P
1.0976				3	0.3493

Source: Reviewer's analysis data set spi56tab.sd2

A graphic representation of baseline FEV_1 values is depicted in Figure 1, below. The bottom and top edges of the boxes mark the 25th and 75th percentiles of the sample. A line connects the medians. The maximums, minimums, and means are printed on the right hand side of the box plot.

Figure 1. Baseline FEV_1 (Study 056)



Source: Reviewer's analysis data set spi56tab.sd2

Table 18 shows selected statistics for the AUC values of FEV_1 at Visit 2 for the four treatment groups.

Table 18. AUC of FEV_1 at Visit 2 (Study 056)

Treatment	N	Mean	Std	Missing	Min	Max
-All-	773	1.4852	0.5386	7	0.5654	4.5962
CtrlPbo	199	1.3026	0.4818	1	0.5654	2.9275
F12	192	1.5902	0.5630	2	0.7005	3.6355
F24	191	1.5841	0.5540	1	0.7331	4.5962
IPR	191	1.4711	0.5049	3	0.5698	3.0437

Source: Reviewer's analysis data set spi56tab.sd2

The same statistics for Visit 5 are shown in Table 19.

Table 19. AUC of FEV_1 at Visit 5 (Study 056)

Treatment	N	Mean	Std	Missing	Min	Max
-All-	686	1.4917	0.5397	30	0.5568	3.6596
CtrlPbo	166	1.3155	0.4734	13	0.5568	3.4115
F12	179	1.6200	0.5747	4	0.6934	3.6596
F24	168	1.5518	0.5383	6	0.5845	3.2543
IPR	173	1.4696	0.5204	7	0.5863	2.8872

Source: Reviewer's analysis data set spi56tab.sd2

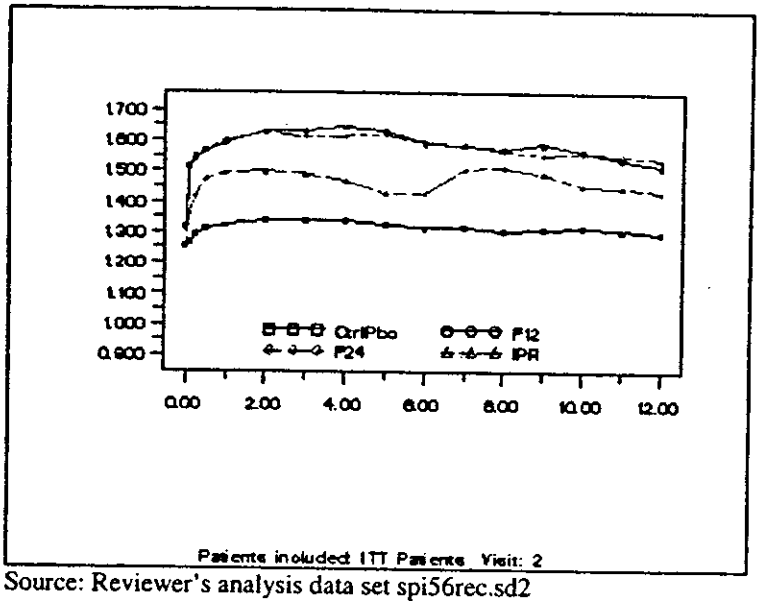
Table 18 and Table 19 show that, without adjustment for other factors in addition to the treatment effect, for Visits 2 and 5, the mean AUC of FEV_1 in the placebo groups were smaller than those in the Foradil groups.

The mean AUC of FEV_1 in the ipratropium groups were greater than those in the placebo group, but smaller than those in the Foradil groups.

Please note that the variable, **Missing** represents the number of patients with missing AUC of FEV_1 values. Please also note that the patient counts shown in the above table are slightly smaller than those shown in Table 16, where the few patients without "Time 0" FEV_1 values were not counted.

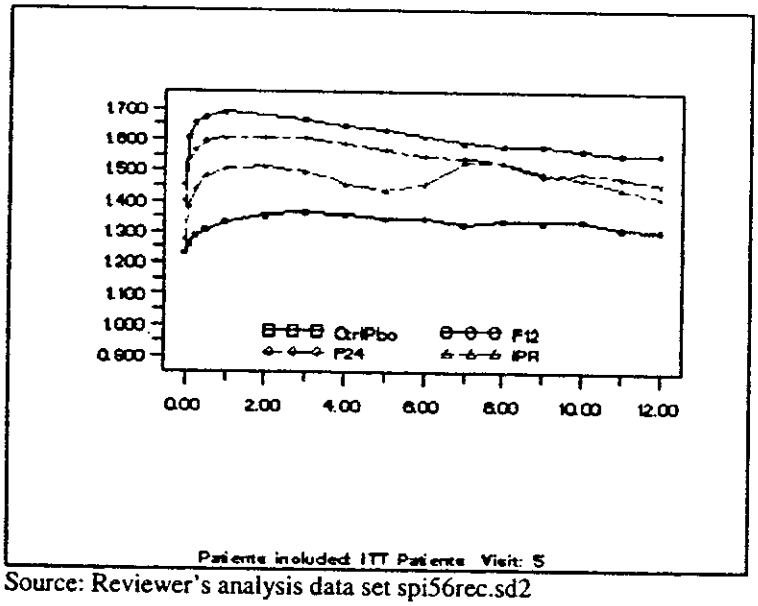
The values of FEV_1 over time by treatment for both visits are shown in the following Figure 2 and Figure 3. Figure 2 demonstrated that the FEV_1 values in Foradil groups (12 and 24 µg doses) were clearly greater than those in the ipratropium groups, which was greater than those in the placebo group. The difference between the two Foradil groups appeared to be small. Foradil's statistical superiority is confirmed in this review section.

Figure 2. FEV_1 vs. time at Visit 2 (Study 056)



Source: Reviewer's analysis data set spi56rec.sd2

Figure 3. FEV_1 vs. time at Visit 5 (Study 056)



Source: Reviewer's analysis data set spi56rec.sd2

Figure 3 above shows that Foradil at 12 and 24 μg had a greater AUC of FEV_1 than did the placebo at Visit 5. Interestingly, the Foradil at 12 μg dose level shows a greater AUC of FEV_1 than the 24 μg dose. The difference between Foradil at 24 μg and ipratropium appeared to diminish at about 8 hours post-dosing and onward.

This reviewer's statistical analysis was based on a statistical model including the factors: treatment, country, center within country, and baseline FEV_1 as the covariate. This model was simpler than that of the sponsor's. But the statistical conclusions based on the two models are the same.

Table 20 and Table 21 describe the statistical analysis of AUC (0-12hours) of FEV_1 at Visit 5. Foradil at 12 and 24 μg proves to be superior to two competitors, placebo and ipratropium. Ipratropium is also superior to the placebo. The difference between the two Foradil regimens appears to be small and insignificant.

Table 20. Estimated Mean of AUC of FEV_1 at Visit 5 (Study 056)

TREAT	95% Lower Confidence Limit	AUC of FEV_1 (LOCF) LS_MEAN	95% Upper Confidence Limit
CtrlPbo	1.286729	1.324992	1.363254
F12	1.507856	1.545147	1.582438
F24	1.484490	1.522313	1.560136
IPR	1.412623	1.450734	1.488846

Source: Reviewer's analysis data set spi56ta2.sd2

Table 21. Comparisons in AUC of FEV_1 at Visit 5 (Study 056)

General Linear Models Procedure
Tukey's Studentized Range (HSD) Test for variable: AUC of FEV_1 (LOCF)
NOTE: This test controls the type I experiment wise error rate.
Alpha= 0.05 Confidence= 0.95 df= 703 MSE= 0.061977
Critical Value of Studentized Range= 3.642
Comparisons significant at the 0.05 level are indicated by '***'.

TREAT Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
F12 - F24	-0.01545	0.05051	0.11647	
F12 - IPR	0.09853	0.16439	0.23026	***
F12 - CtrlPbo	0.23112	0.29707	0.36303	***
F24 - F12	-0.11647	-0.05051	0.01545	
F24 - IPR	0.04749	0.11388	0.18027	***
F24 - CtrlPbo	0.18008	0.24656	0.31304	***
IPR - F12	-0.23026	-0.16439	-0.09853	***
IPR - F24	-0.18027	-0.11388	-0.04749	***
IPR - CtrlPbo	0.06629	0.13268	0.19907	***
CtrlPbo - F12	-0.36303	-0.29707	-0.23112	***
CtrlPbo - F24	-0.31304	-0.24656	-0.18008	***
CtrlPbo - IPR	-0.19907	-0.13268	-0.06629	***

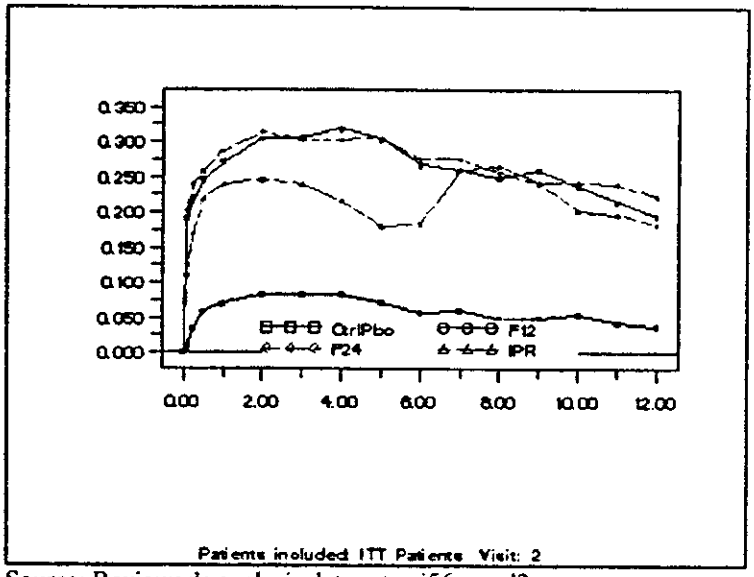
Source: Reviewer's analysis data set spi56ta2.sd2

For the data exploration purpose, and for the interest of the reviewing medical officer, Dr. Sullivan, it is useful to visualize the FEV_1 changes from test-day pre-dose values.

The following graphs, Figure 4 and Figure 5 show the changes in FEV_1 from pre-dosing values by visit.

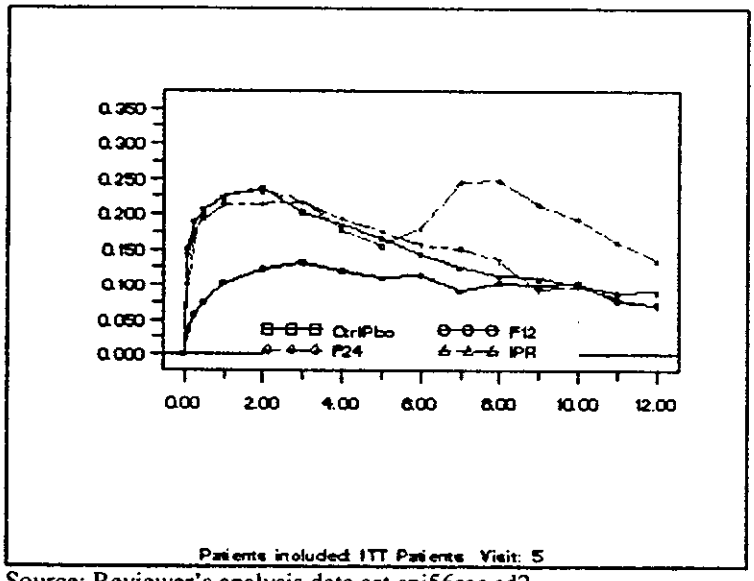
For Visit 2, the FEV_1 changes for Foradil at the two dosages are comparable with that for ipratropium, and clearly greater than that of the placebo. Particularly, the Foradil outperformed ipratropium for the first sever (7) hours post-dosing.

Figure 4. Changes in FEV_1 from test-day pre-dose values vs. time at Visit 2 (Study 056)



Source: Reviewer's analysis data set spi56rec.sd2

Figure 5. Changes in FEV_1 from test-day pre-dose values vs. time at Visit 5 (Study 056)



Source: Reviewer's analysis data set spi56rec.sd2

Unlike Visit 2, the changes in FEV_1 from test-day pre-dose values in the Foradil groups at Visit 5 appeared to be greater than that in placebo group. But the changes in FEV_1 did not appear to be much different from the ipratropium.

Additional Analysis (Study 056)

In Study 056, U.S. accounts for 9.1% (4/44) of the centers and 6.4% (50/780) of all patients. Regarding regulatory decisions based on international study as such, it is useful to compare the response of U.S. patients to the study drug, Foradil 12 and 24μg b.i.d., with that or non-U.S. patients.

In this effort, this reviewer will not perform hypothesis tests by U.S. and non-U.S. patient groups. They would be otherwise underpowered. Instead, visual examinations and descriptive analyses will be applied to facilitate the review team with information should any labeling and other recommendations might be considered.

Comparisons of Pre-Dosing FEV_1 Between U.S. and Non-U.S. Patient Populations

The following graphs (Figure 6 and Figure 7) and Table 22 show the changes in pre-dosing FEV_1 from Visit 2 to Visit 5, for U.S. and non-U.S. patients, respectively. This comparison may show how differently the two patient populations responded to the treatment at the endpoint visit (Visit 5). Note that among U.S. patients treated with placebo, pre-dosing FEV_1 from Visit 2 to Visit 5 decreased noticeably: from 1.13 to 0.99 by 12%, while FEV_1 increased slightly in the other treatment groups. Among non-U.S. patients, pre-dosing FEV_1 in the placebo group only decreased slightly: from 1.26 to 1.25 by 0.79%. FEV_1 in other treatment groups increased from Visit 2 to Visit 5 in U.S. and non-U.S. patient populations.

Figure 6. Pre-dosing FEV_1 (Study 056, U.S. patients)

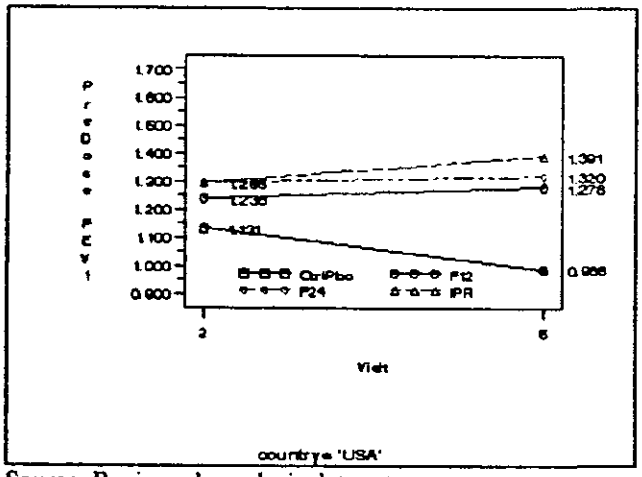
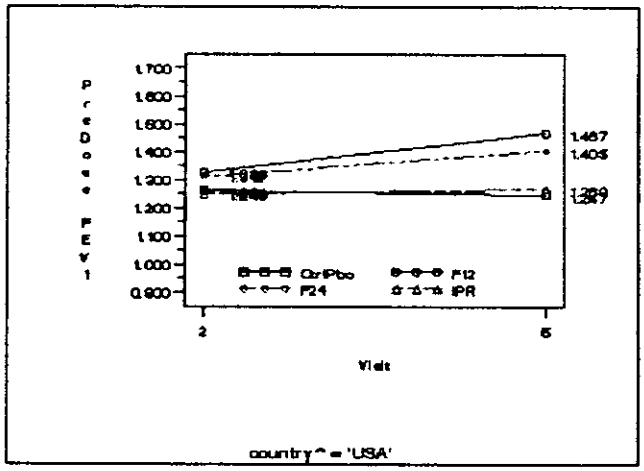


Figure 7 Pre-dosing FEV_1 (Study 056, non-U.S. patients)



Source: Reviewer's analysis data sets

The cause of noticeable drop in pre-dosing FEV_1 from Visit 2 to Visit 5 in the U.S. patients is not clear to this reviewer. It is worth further exploring the difference in response to the testing drug between U.S. and non-U.S. patients.

In addition to the observations on the placebo groups for U.S. and non-U.S. patients, Table 22 shows that the pre-dosing FEV_1 increased from Visit 2 to Visit 5 in active treatment groups in both patient populations.

Table 22. Pre-dosing FEV_1 (Study 056)

Pre-dosing FEV_1	Treatment							
	CtrlPbo		F12		F24		IPR	
	Visit		Visit		Visit		Visit	
	2	5	2	5	2	5	2	5
Country								
ALL COUNTRIES	1.26	1.23	1.32	1.45	1.31	1.40	1.25	1.28
NON-USA	1.26	1.25	1.33	1.47	1.31	1.41	1.25	1.27
USA	1.13	0.99	1.24	1.28	1.30	1.32	1.30	1.39

If the dropout rates in non-U.S. patients treated with placebo were higher than that in the U.S. patients due to the lack of improvement or the worsening of condition, one might think the remaining (non-U.S.) patients in the placebo group would do better than those in the U.S. placebo group. This is not the case here. Table 23, below, indicates that the U.S. patients in the placebo group had a greater dropout rate (23.1%) than the non-U.S. patients treated with placebo had (10.7%). Therefore, the dropout rate may not be closely related to the difference in pre-dosing FEV_1 at Visit 5 in the U.S. and non-U.S. placebo groups.

Table 23. Dropouts of patients at Visit 5 (Study 056)

	NON_US				USA			
	Total ITT	Completer	% Complete	% Dropout	Total ITT	Completer	% Complete	% Dropout
Treat								
CtrlPbo	187	167	89.3%	10.7%	13	10	76.9%	23.1%
F12	181	170	93.9%	6.1%	13	13	100.0%	0.0%
F24	181	164	90.6%	9.4%	11	10	90.9%	9.1%
IPR	181	164	90.6%	9.4%	13	13	100.0%	0.0%

Examination of Country-Treatment Interaction

Having seen the difference in pre-dosing FEV_1 between the U.S. and non-U.S. patient populations, it is appears to be useful to further explore the difference in country-treatment interaction. By interaction, it means whether the same treatment affects patients in different country more or less the same way. If the treatment effect (statistically) significantly differs from country to country, particularly in this trial based on from foreign clinical data, the decisions, including labeling recommendations might be more complicated than usual.

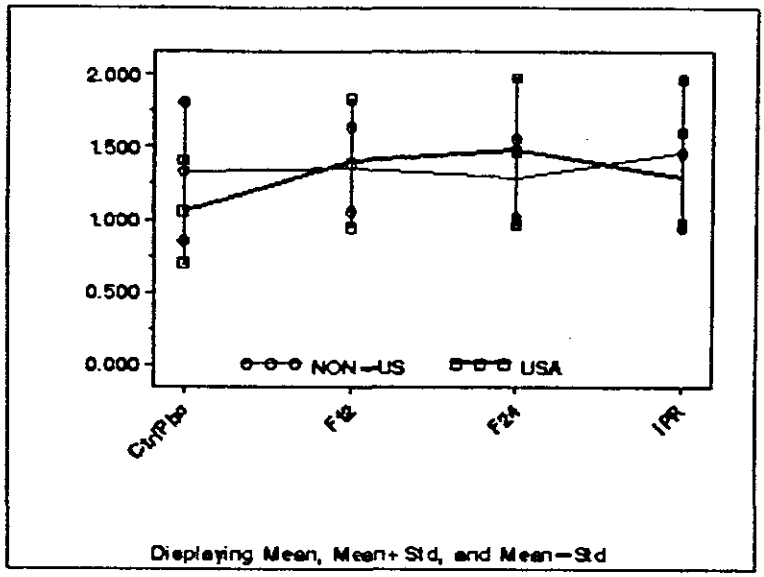
The test for country-by-treatment interaction (Table 24) indicated that such interaction is statistically significant with p-value of 0.0549.

Table 24. Test for country-by-treatment interaction on U.S. patients (Study 056)

COUNTRY	10	2.4788663	0.2478866	3.88	0.0001
CENTER(COUNTRY)	33	2.9687851	0.0899632	1.41	0.0672
COUNTRY*TREAT	30	2.8032099	0.0934403	1.46	0.0549
TREAT	3	2.5903088	0.8634363	13.51	0.0001
BASEFEV	1	113.0468999	113.0468999	1768.86	0.0001

Figure 8 shows mean AUC of FEV_1 at Visit 5 for U.S. and non-U.S. patients, indicated by legends. The points of mean plus and minus one STD are also shown. Lines connect treatment means. The crossing lines indicate the existence of country-treatment interaction.

Figure 8. AUC of FEV_1 at Visit 5 (Study 056)



Following graphs (Figure 9 to Figure 12) shows AUC of FEV_1 at Visit 5, by treatment and country. Country differences are observed, however, the U.S. data do not seem to be much different from the rest of the participating countries.

Figure 9. AUC of FEV_1 at Visit 5 by country: Placebo control (Study 056)

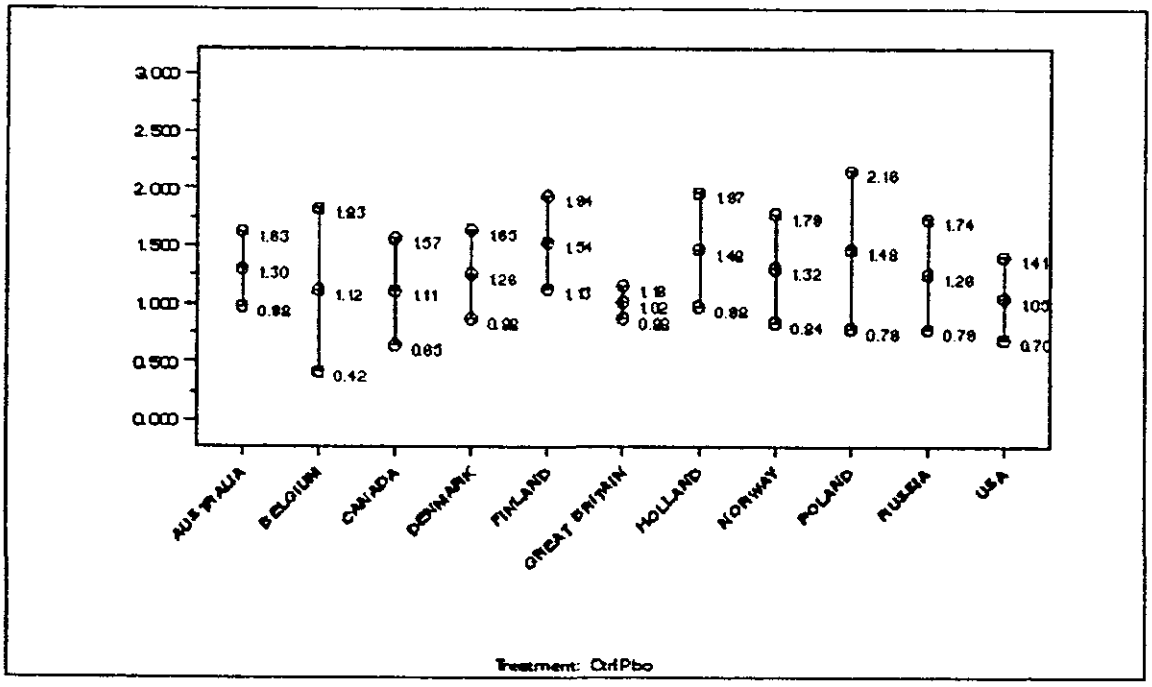


Figure 10. AUC of FEV_1 at Visit 5 by country: Foradil 12 µg (Study 056)

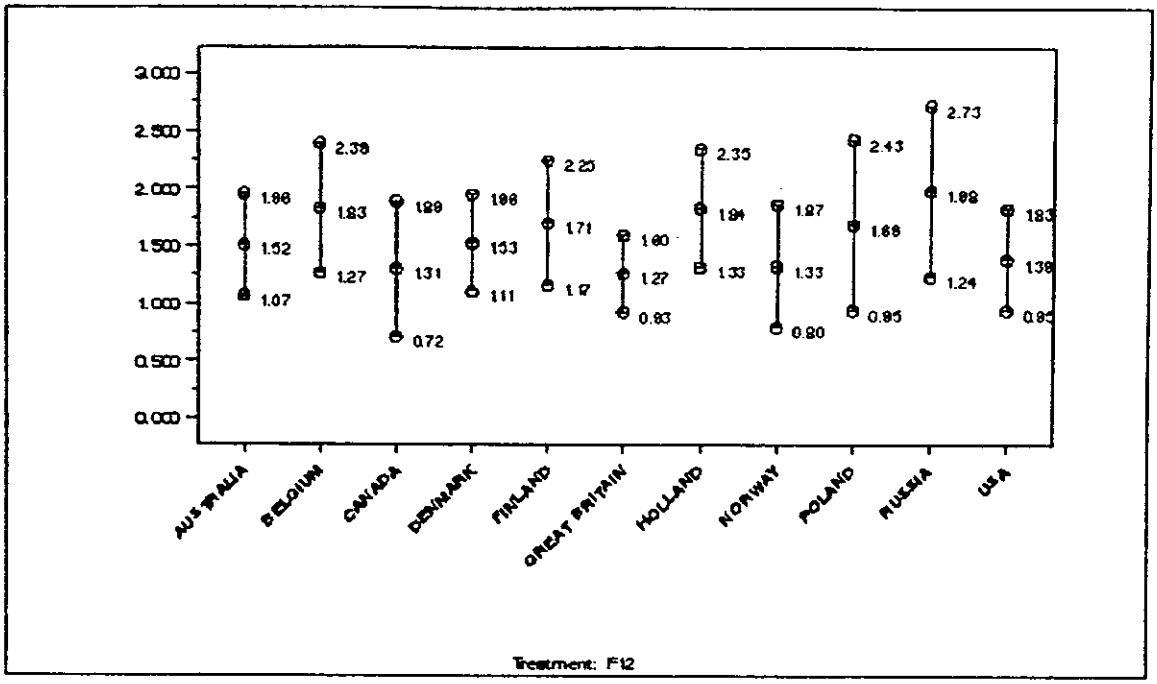


Figure 11. AUC of FEV_1 at Visit 5 by country: Foradil 24 µg (Study 056)

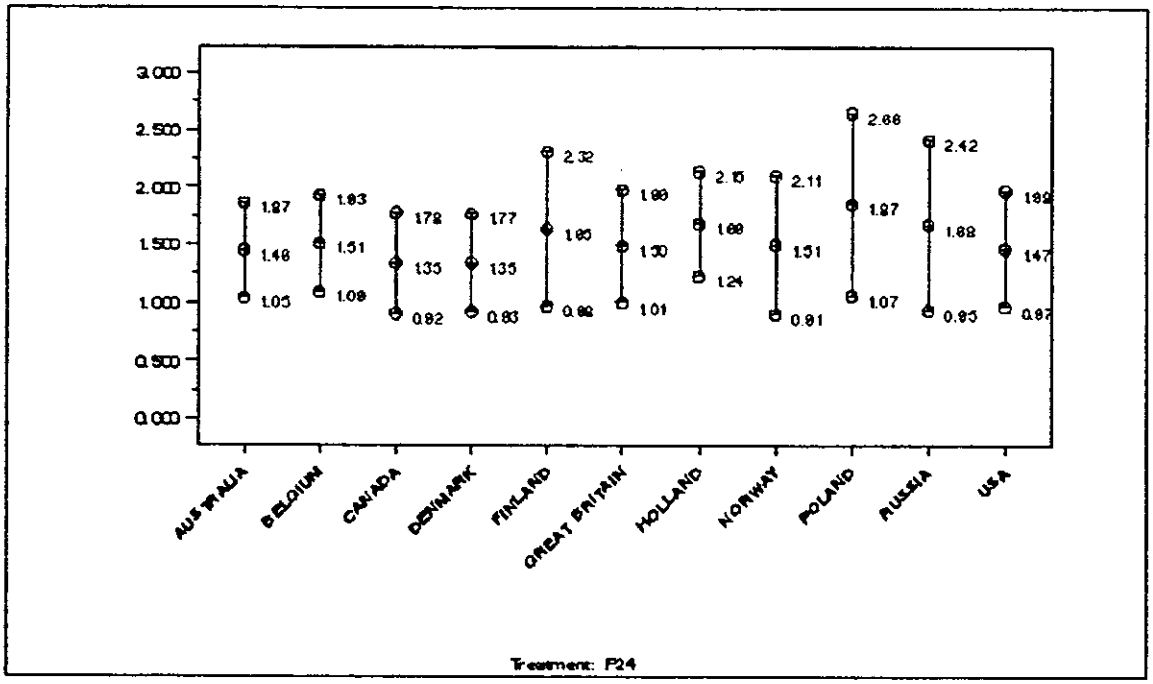
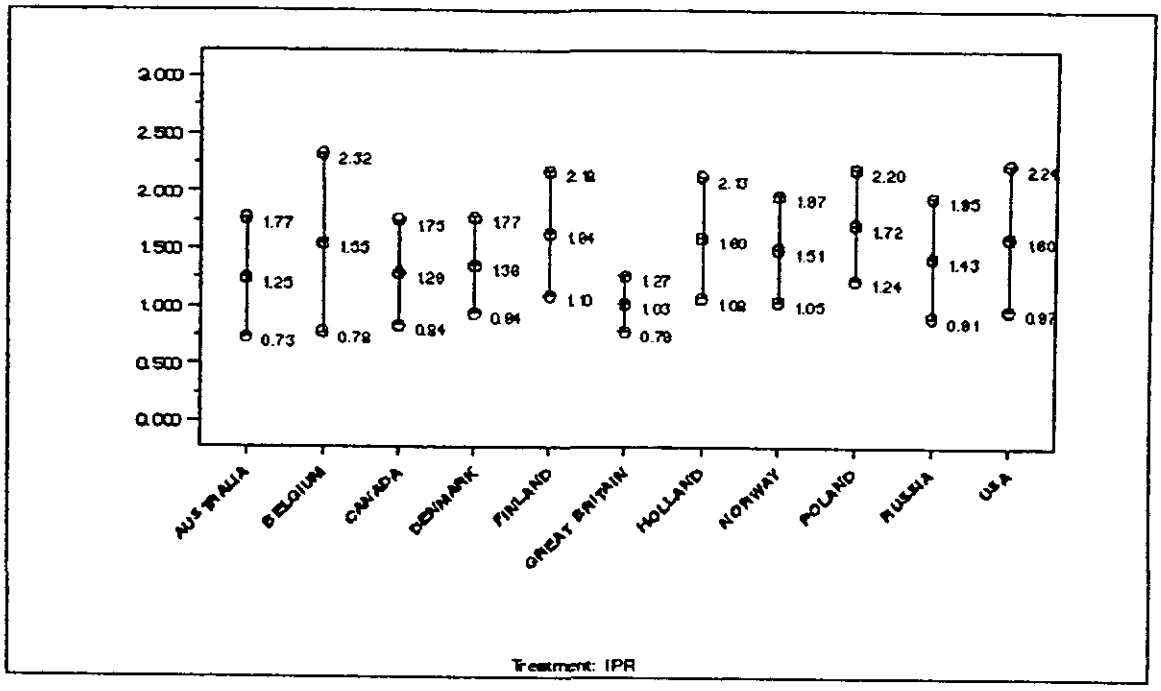


Figure 12. AUC of FEV_1 at Visit 5 by country: Ipratropium (Study 056)



Examination of Hourly FEV_1 for U.S. and Non-U.S. Patients

Figure 13 and Figure 14 show FEV1 vs. time at Visit 5 for U.S. and non-U.S. patients, respectively.

Figure 13. FEV_1 vs. time at Visit 5 for U.S. patients (Study 056)

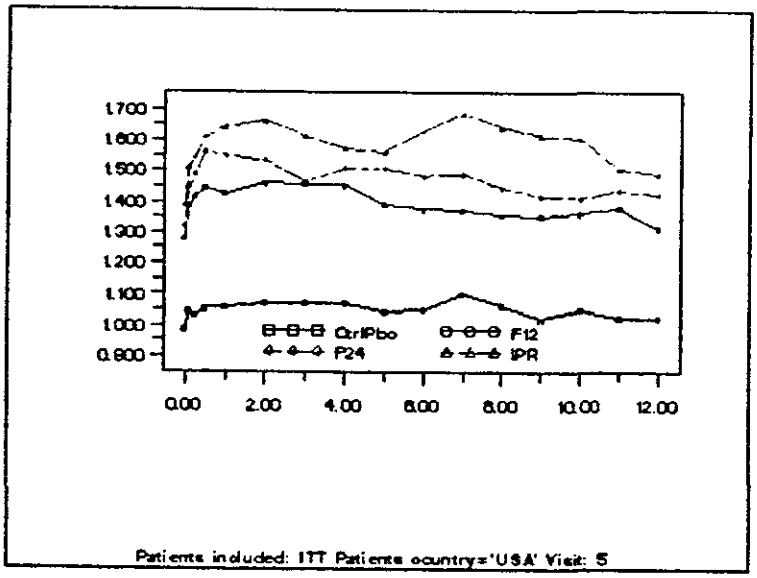
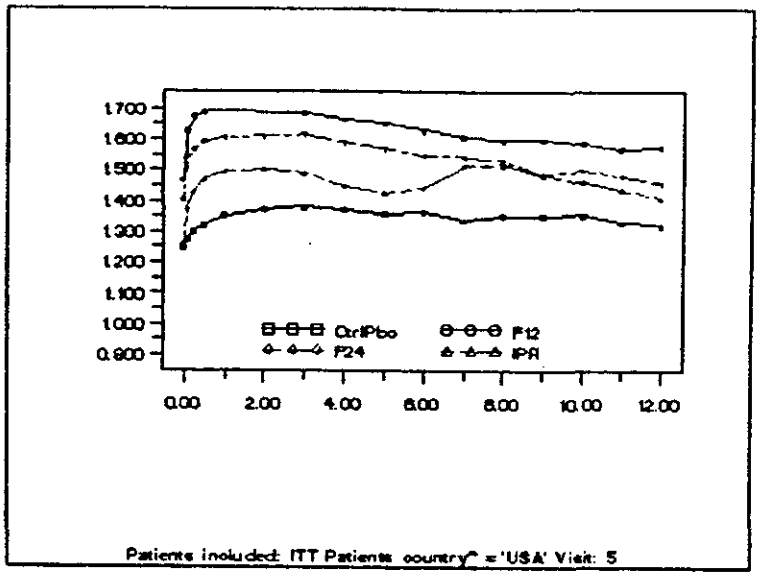


Figure 14 FEV_1 vs. time at Visit 5 for non-U.S. patients (Study 056)



Based on the above two graphs (Figure 13 and Figure 14), the following points are worth noting:

1. Based on examination of the AUC of FEV_1 for U.S. and non-U.S. patients at Visit 5, it appears that Foradil at 24 µg b.i.d. provided more benefit to the U.S. patients than did Foradil at 12 µg b.i.d., while such a difference was not demonstrated for non-U.S. patients.
2. The same examination also indicates that the U.S. patients treated with ipratropium had a greater AUC than did those treated with Foradil. On the contrary, among the non-U.S. patients, Foradil at 12 µg b.i.d. appeared to be more effective than Foradil at 24 µg b.i.d., which appeared superior to ipratropium. Therefore, the country-by-treatment interaction might be significant.
3. For confirmation purposes, a test for country-by-treatment interaction (Table 24) indicated that such interaction is statistically significant with p-value 0.0549[8].
4. The average FEV_1 for U.S. patients in the placebo group at Visit 5 was much lower than for non-U.S. patients, while such a difference was very small at the beginning of the treatment period. In addition, U.S. patients dropped out at a higher percentage than non-U.S. patients.

In summary, the magnitudes of the effects of the active treatments (Foradil and Ipratropium) on the U.S. patients differed from the effects on non-U.S. patients. Please note that this finding was based on an exploratory analysis.

[8] The test for country-by-treatment interaction includes terms of COUNTRY, CENTER(COUNTRY), COUNTRY*TREAT, TREAT, BASEFEV in the linear model.

Reviewer's Evaluation of Study 058

This reviewer examined the completeness of the patient data. Table 25 shows the percentages of patients who stayed on study by treatment and by visit. A reasonably large proportion (ranging from approximately 85% to 94%) of the patients were followed as of Week 12 (Visit 3). The theophylline-treated patient group had the lowest percentage of patients (85.2%) who completed 12 weeks of the study. Even though the study lasted for 12 months, the primary efficacy study was based on the data from the first 12 weeks' of treatment.

Table 25. Numbers and percentages of patients who stayed in the study by visit (Study 058)

Treatment	#Total Patients	#Patients	Percent Completed	Visit
CtrlPbo	220	220	100.0%	2
CtrlPbo	220	205	93.2%	3
CtrlPbo	220	182	82.7%	4
CtrlPbo	220	170	77.3%	5
CtrlPbo	220	165	75.0%	6
F12	211	211	100.0%	2
F12	211	200	94.8%	3
F12	211	178	84.4%	4
F12	211	167	79.1%	5
F12	211	165	78.2%	6
F24	214	214	100.0%	2
F24	214	203	94.9%	3
F24	214	190	88.8%	4
F24	214	179	83.6%	5
F24	214	177	82.7%	6
THE	209	209	100.0%	2
THE	209	178	85.2%	3
THE	209	148	70.8%	4
THE	209	134	64.1%	5
THE	209	130	62.2%	6

Source: Reviewer's analysis data set spi58rec.sd2

Table 26 describes the mean values of FEV_1 at baseline for the treatment groups. The baseline is defined as the pre-treatment FEV_1 values at Visit 2. The overall difference in FEV_1 appears to be small among the treatment groups (p=0.4912).

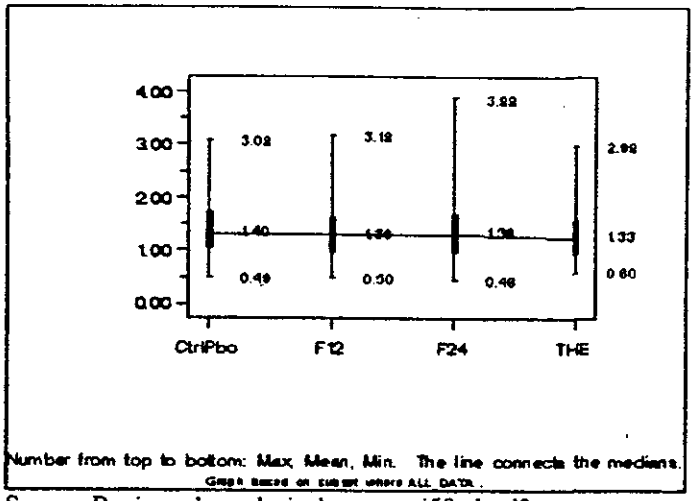
Table 26. Baseline FEV_1 (Study 058)

Treat	N	MEAN	STD	MIN	MAX
CtrlPbo	220	1.4008	0.4896	0.4900	3.0800
F12	211	1.3576	0.4599	0.5000	3.1800
F24	214	1.3875	0.5346	0.4600	3.8800
THE	209	1.3335	0.4814	0.6000	2.9800
Test for Group Diff: F				DF	P
0.8051				3	0.4912

Source: Reviewer's analysis data set spi58tab.sd2

A graphic representation of baseline FEV_1 values is depicted in Figure 15, below. It shows the distributions of the baseline FEV_1 values for each of the treatments. The bottom and top edges of the boxes mark the 25th and 75th percentiles of the sample. A line connects the medians. The maximums, minimums, and means are printed on the right hand side of the box plot.

Figure 15. Baseline FEV_1 (Study 058)



Source: Reviewer's analysis data set spi58tab.sd2

Table 27 shows selected statistics for the AUC values of FEV_1 at Visit 3 (Week 12) for the four treatment groups.

Table 27. AUC of FEV_1 at Visit 3 (Study 058)

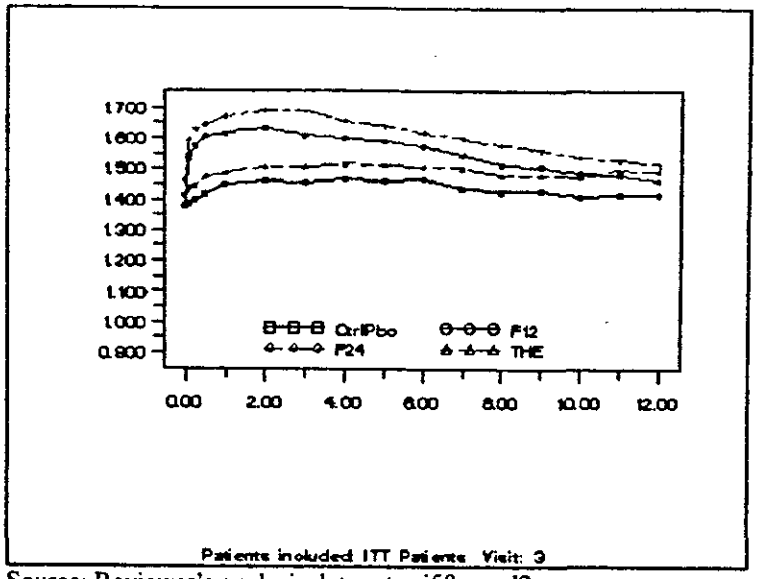
Treatment	N	Mean	Std.	#Missing	Min	Max
-All-	728	1.5231	0.5566	113	0.5226	4.3203
CtrlPbo	186	1.4238	0.5823	31	0.5226	4.3203
F12	191	1.5589	0.4960	19	0.6286	2.8834
F24	197	1.6121	0.5794	14	0.6865	3.4837
THE	154	1.4846	0.5486	49	0.5539	3.3938

Source: Reviewer's analysis data set spi58tab.sd2

Note that the sponsor did not report the AUC of FEV_1 for Visit 2 (baseline). In addition, the mean AUC of FEV_1 in the theophylline group was greater than that of the placebo group, but smaller than the Foradil groups.

The values of FEV_1 over time, by treatment, for both visits are shown in the following Figure 16. This graphic demonstrates that the FEV_1 values in Foradil groups (12 and 24 μg doses) were clearly greater than those in the theophylline group, which, in turn, was superior to the placebo group for up to 10 hours post dosing. Foradil at 24 μg appeared to provide more improvement in FEV_1 than did Foradil at 12 μg.

Figure 16. FEV_1 vs. time at Visit 3 (Study 058)



Source: Reviewer's analysis data set spi58rec.sd2

This reviewer's statistical analysis was based on a statistical model including the factors: treatment, country, center within country, and baseline FEV_1 as the covariate.

Table 28 and Table 29 describe the statistical analysis of AUC (0-12hours) of FEV_1 at Visit 3. Foradil at 12 and 24 μg was superior to two comparators, placebo and theophylline. Note that theophylline failed to demonstrate superiority to placebo, based on Tukey's Test (also based on Dunnett's T Test and Scheffe's F test, not shown in this report). The difference between the two Foradil regimens does not appear to be statistically significant.

Table 28. Estimated Mean of AUC of FEV_1 at Visit 3 (Study 058)

TREAT	95% Lower Confidence Limit	AUC of FEV_1 LS_MEAN	95% Upper Confidence Limit
CtrlPbo	1.331707	1.385407	1.439108
F12	1.540208	1.592332	1.644455
F24	1.556569	1.608117	1.659665
THE	1.460785	1.517130	1.573474

Source: Reviewer's analysis data set spi58tab.sd2

Table 29. Comparisons in AUC of FEV_1 at Visit 3 (Study 058)

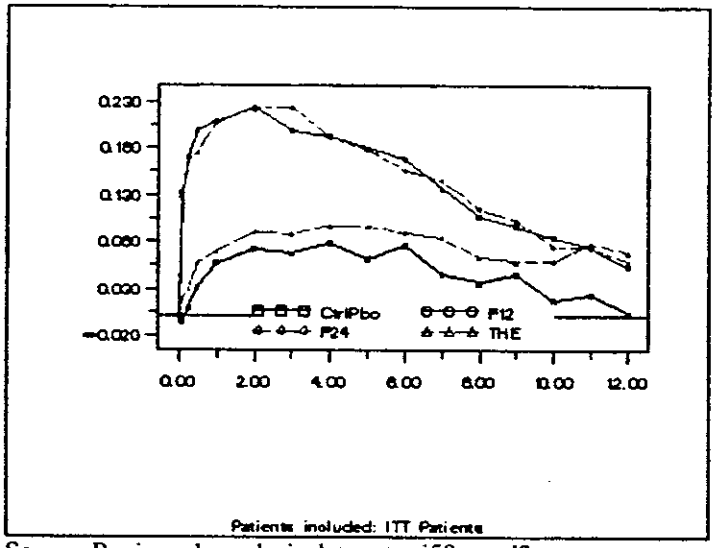
General Linear Models Procedure
Tukey's Studentized Range (HSD) Test for variable: AREAFEV2
NOTE: This test controls the type I experimentwise error rate.
Alpha= 0.05 Confidence= 0.95 df= 668 MSE= 0.076659
Critical Value of Studentized Range= 3.642
Comparisons significant at the 0.05 level are indicated by '***'.

TREAT Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
F24 - F12	-0.01301	0.05921	0.13144	
F24 - THE	0.05550	0.13212	0.20874	***
F24 - CtrlPbo	0.11839	0.19131	0.26423	***
F12 - F24	-0.13144	-0.05921	0.01301	
F12 - THE	-0.00423	0.07291	0.15005	
F12 - CtrlPbo	0.05863	0.13210	0.20556	***
THE - F24	-0.20874	-0.13212	-0.05550	***
THE - F12	-0.15005	-0.07291	0.00423	
THE - CtrlPbo	-0.01860	0.05919	0.13698	
CtrlPbo - F24	-0.26423	-0.19131	-0.11839	***
CtrlPbo - F12	-0.20556	-0.13210	-0.05863	***
CtrlPbo - THE	-0.13698	-0.05919	0.01860	

Source: Reviewer's analysis data set spi56ta2.sd2

For the data exploration purpose, and for the interest of the reviewing medical officer, Dr. Sullivan, it is useful to visualize the FEV_1 changes from test-day pre-dose values.

Figure 17. Changes in FEV_1 from test-day pre-dose values vs. time at Visit 3 (Study 058)



Source: Reviewer's analysis data set spi58rec.sd2

The changes in FEV_1 from test-day pre-dose values in the Foradil groups at Visit 3 appeared to be greater than those in the placebo and theophylline groups.

Additional Analysis (Study 058)

In Study 58, U.S. accounts for 22.8% (13/57) of the centers and 23.3% (199/854) of all patients. Regarding regulatory decisions based on international study as such, it is useful to compare the response of U.S. patients to the study drug, Foradil 12 and 24 μg b.i.d., with that or non-U.S. patients.

Figure 18 shows FEV_1 vs. time at Visit 3 for U.S. patients.

Figure 18. FEV_1 vs. time at Visit 3 for U.S. patients (Study 058)

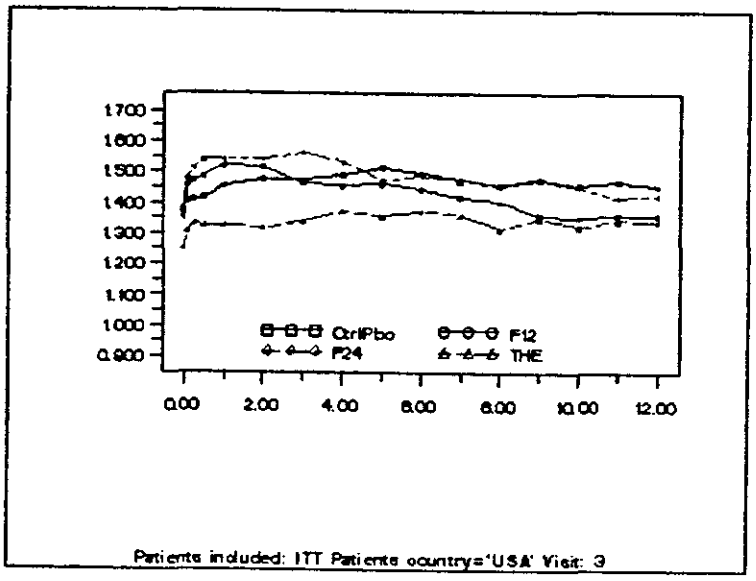
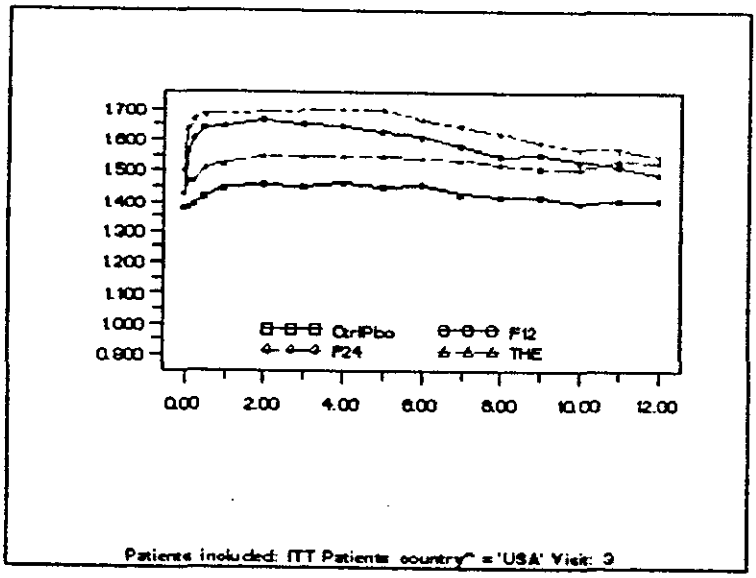


Figure 19 shows FEV_1 vs. time at Visit 3 for non-U.S. patients.

Figure 19. FEV_1 vs. time at Visit 3 for non-U.S. patients (Study 058)



Based on Figure 19, among non-U.S. patients, the AUC of FEV_1 in the Foradil at 24 μg group was greater than that of the Foradil 12 μg group. Both Foradil groups had greater AUC of FEV_1 than the theophylline and placebo groups. However, for U.S. patients, based on Figure 18, the differences between the two

Foradil groups and the placebo control group in terms of AUC of FEV_1 was too small to tell. Therefore, Foradil appeared to be less effective on the U.S. patients than on non-U.S. patients.

Conclusions

The review of this NDA was based on evaluation of two multinational studies, Studies 056 and 058. In Study 056, US patients, included in 4 (out of 44, or 9.1%) trial centers, comprised 6.4% (50 out of 780) of all patients. In Study 058, US patients, included in 13 (out of 57, or 22.8%) trial centers, comprised 23.3% (199 out of 854) of all patients. This reviewer reached the following conclusions:

- In Studies 056 and 058, Foradil at 12 and 24 μg b.i.d. was superior to the placebo control.
- In both studies the differences in AUC of FEV_1 between 12 and 24 μg b.i.d. of Foradil were not statistically significant.
- In Study 056, Foradil at 12 and 24 μg b.i.d. was statistically superior to Ipratropium MDI 40 μg q.i.d., which was also shown to be statistically superior to the placebo.
- In Study 058, Foradil at 24 μg b.i.d., but not at 12 μg b.i.d., proved to be statistically superior to theophylline, which, as an open-labeled treatment, failed to demonstrate its superiority to the placebo.
- In Study 056, subjects receiving Foradil at 12 μg b.i.d. appeared to have a greater average AUC value of FEV_1 than did those patients receiving Foradil 24 μg b.i.d. However, in US patients, there was a greater FEV_1 response to Foradil 24.
- The lack of statistical significance favoring one dose to another suggests no particular preference for the dosage of Foradil.
- Descriptive comparison of AUC (of FEV_1) data appears to indicate that the treatment effects were somewhat different for the U.S. and non-U.S. patients in the studies.

In conclusion, the effectiveness of Foradil at 12 and 24 μg b.i.d. was demonstrated in Studies 056 and 058 as a whole. However, a visual comparison of AUC (of FEV_1) between U.S. and non-U.S. patients suggested that Foradil is not as effective for the U.S. patients as for the non-U.S. patients.

Signoff Page

Reviewer: Ted Guo, Ph.D. Ted Guo 9/25/01
Concur: Steve Wilson, Ph.D.

CC:
Archival NDA 21-279

HFD-570/Division file
HFD-570/ESullivan
HFD-570/COstroff

HFD-715/Division file
HFD-715/SWilson
HFD-715/TGuo

HFD-700/Canello

TG/Nda21279.doc

APPEARS THIS WAY
ON ORIGINAL

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

/s/

Ted Guo
9/25/01 02:05:07 PM
BIOMETRICS

Steve Wilson
9/25/01 02:18:12 PM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL