CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-779

Statistical Review(s)

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

Statistical Review and Evaluation

Clinical Studies

NDA/Serial Number:	NDA 21-779
Drug Name:	Ventavis (iloprost) inhalation solution
Indication(s):	Pulmonary hypertension
Applicant:	CoTherix
Date(s):	Received: 06/30/04
	DFS: 10/27/04
Review Priority:	Priority
Biometrics Division:	DB1 (HFD-710)
Statistical Reviewer:	Valeria Freidlin, Ph.D.
Concurring Reviewers:	James Hung, Ph.D., Team Leader,
	Kooros Mahjoob, PH. D., Acting Division Director
Medical Division:	Cardio-Renal, HFD-110
Clinical Team:	Maryann Gordon, MD (HFD-110)
Project Manager:	Melissa Robb (HFD-110)

Keywords: CMH test, stratification, non-parametric ANCOVA, baseline control, missing data, sensitivity analyses, survival analysis, Kaplan-Meier product limit, log-rank test, logistic regression, covariates, odds ratio, Breslow-Day test.

Table of Contents

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES 1

FOOD AND DRUG ADMINISTRATION 1

STATISTICAL REVIEW AND EVALUATION 1

1. EXECUTIVE SUMMARY 3

1.1 CONCLUSIONS AND RECOMMENDATIONS 3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES 3
1.3 STATISTICAL ISSUES AND FINDINGS 3

2. INTRODUCTION 4

2.1 OVERVIEW 4
2.2 DATA SOURCES 4

3. STATISTICAL EVALUATION 4

3.1 EVALUATION OF EFFICACY 4
3.2 EVALUATION OF SAFETY 13

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 13

4.1 COMBINED RESPONDER CRITERION IN THE PREDEFINED SUBGROUPS 13
4.2 CHANGE FROM BASELINE IN WALKING DISTANCE AND MORTALITY IN PREDEFINED SUBGROUPS. 14

5. SUMMARY AND CONCLUSIONS 15

5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE 15
5.2 CONCLUSIONS AND RECOMMENDATIONS 15

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Primary efficacy analysis of Study ME97218 showed that iloprost is statistically significantly better than placebo relative to the combined responder rate at week 12 based on the ITT-observed cases (p=0.0067). Sensitivity analyses of the combined responder rate based on the ITT-LOCF and the Per Protocol populations support the primary efficacy results ($p \leq 0.016$). Secondary analyses of the change from baseline to week 12 in walking distance and of other secondary endpoints were consistent with the primary efficacy analysis. Although Study ME97218 seems to be a solid confirmatory study, for the non-mortality endpoints the usual standard requires efficacy evidence to be confirmed in at least one more study.

As the sponsor did not pre-specify any statistical rule for dealing with multiple secondary endpoints, there is no basis for selective inclusion of favorable p-values for some secondary endpoints in the sponsor's labeling.

1.2 Brief Overview of Clinical Studies

In this NDA, the efficacy of inhaled iloprost is primarily based on the data from a single Phase 3, European Study ME97218 in patients with pulmonary arterial hypertension classified as NYHA Class III or IV. Study ME97218 is a randomized, double-blind, 2-arm, parallel-group, placebo-controlled, multi-national, multi-center study in patients with primary or secondary pulmonary hypertension (PHT).

1.3 Statistical Issues and Findings

In general, the reviewer agrees with the sponsor's analytical methods. The reviewer's primary efficacy results are very close to the sponsor's primary efficacy results. As the sponsor did not pre-specify any statistical rule for dealing with multiple secondary endpoints, there is no basis for selective inclusion of favorable p-values for some secondary endpoints in the sponsor's labeling. For the secondary efficacy endpoints, p-values in this review are shown only to support consistency of the results, because there was no adjustment for multiple comparisons with respect to the large number of secondary endpoints, number of subgroups, and number of statistical methods.

Primary efficacy analysis of Study ME97218 showed that iloprost is statistically significantly better than placebo relative to the combined responder rate at week 12 based on the ITT-observed cases (p=0.0067). Sensitivity analyses of the combined responder rate based on the ITT-LOCF and the Per Protocol populations support the primary efficacy results ($p \leq 0.016$). Secondary analyses of the change from baseline to week 12 in walking distance and of other secondary endpoints were consistent with the primary efficacy analysis. Although Study

ME97218 seems to be a solid confirmatory study, for the non-mortality endpoints the usual standard requires efficacy evidence to be confirmed in at least one more study.

As the sponsor did not pre-specify any statistical rule for dealing with multiple secondary endpoints, there is no basis for selective inclusion of favorable p-values for some secondary endpoints in the sponsor's labeling.

2. INTRODUCTION

2.1 Overview

In this NDA, the efficacy of inhaled iloprost is primarily based on the data from a single European Phase 3 Study ME97218 in patients with pulmonary arterial hypertension classified as NYHA Class III or IV. Study ME97218 is a randomized, double-blind, 2-arm, parallel-group, placebo-controlled, multi-national, multi-center study in patients with primary or secondary pulmonary hypertension (PHT).

2.2 Data Sources

Efficacy and safety data sets for Study ME97218 were provided by the sponsor on 6/30/04 and are stored in the EDR: \\CDSESUB1\n21779\N_000\2004-06-30\ crt\datasets\a02997.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

Study ME97218 is a randomized, double-blind, 2-arm, parallel-group, placebo-controlled, multi-center study in 203 patients with primary or secondary pulmonary hypertension (PHT). Eligible patients were randomly assigned to one of the following treatment groups for 12 weeks:

1. Inhalation of iloprost aerosol with an individually adapted total daily dose or
2. Inhalation of a corresponding aerosol placebo to be administered under the same conditions as iloprost.

The patients were assigned to treatment groups and stratified for pulmonary hypertension (primary or secondary) and for the NYHA class III or IV. The study was planned as a multinational study in 40 centers in Europe. The study medication was to be administered over a period of 12 weeks.

Double-blind treatment period (weeks 1-12). After the individually tolerated dose and dosing times had been determined during the initial 7 days, they were to be maintained over the remaining treatment period, i.e. for 11 weeks. After the first week the patients

had to be seen for further assessments on day 8 (± 2 days), at week 4, at week 8 and at week 12 (± 5 days, respectively) for the end of treatment visit, or in the event of discontinuation of treatment at the time point of withdrawal.

Follow-up period (weeks 13-16). An additional **follow-up visit** was scheduled four weeks after the end of the treatment period for assessment of safety variables. In case of premature withdrawal from treatment, the patient was to continue to follow the regular schedule of study visits but could receive any therapy.

Randomization. The patients were assigned to either the iloprost or the placebo treatment group by central telephone randomization. The patients were stratified for primary or secondary pulmonary hypertension (PPH or SPH) and for NYHA class (III or IV). Of the planned sample size of n = 100 patients per treatment group, it was ensured by central randomization that at least 25% of patients in either treatment group belonged to NYHA class IV. Patients were allocated a random number once the baseline examinations had been performed and compliance with the entry criteria had been established. Investigators were given blinded packs of study medication, each having a unique number for each patient printed on the label of each unit.

Efficacy variables
Primary efficacy variable

The primary efficacy endpoint was the combined responder criterion defined as:
• Improvement in exercise capacity (6-minute walking test) at 12 weeks by at least 10% versus baseline and
• Improvement by at least one NYHA class at 12 weeks versus baseline, and
• No deterioration of PHT, or death at any time before 12 weeks.

The components of this combined responder criterion were also defined as the secondary efficacy variables. Patients having developed the criteria of deterioration and discontinuing study treatment were to be counted as failure. Death by any cause was to be counted as failure. If efficacy assessments at baseline and week 12 were not possible due to reasons unrelated to PHT, the respective assessments may have been repeated once within 4 weeks after the originally scheduled time of assessment.

Secondary efficacy variables:

• Exercise capacity,
• NYHA class,
• Dyspnea index,
• Hemodynamic parameters (PVR, mPAP, CO) and gas exchange (SVO2),
• Deterioration of PHT,
• Mortality,
• Need for transplantation,

• Quality of life (QoL).

Statistical methods planned in the protocol

Statistical analyses are to be based on the cases actually observed. Additional LOCF analyses might be performed for supportive reasons. Statistical tests for efficacy variables were to be two-sided using a 5% level of significance.

Analysis populations

The primary population for the evaluation of efficacy was to be ITT population defined as all randomized patients with at least one study drug application and at least some post-baseline efficacy data.

Supportive efficacy analyses were to be performed on a **per protocol population.** A patient was included in the **per protocol population** for analysis:
- if the treatment period had been completed, or
- if the patient had discontinued due to the defined criteria of deterioration, or
- if the patient had died, and no major protocol violation had occurred for the respective patient.

The decision on assigning a patient to the per protocol population was to be made by the International Trial Manager before database closure and un-blinding. The safety population was to include all randomized patients with at least one study drug application and at least some post-baseline safety data.

Analysis of the primary efficacy endpoint

The hypotheses to be tested were:
H_0: The responder rates of both treatments are equal
H_1: The responder rates of both treatments are not equal.

For the primary efficacy variable, the method for treatment comparison was to be the two-sided CMH test stratified for PHT (PPH or SPH) and NYHA (III or IV) at the 0.05 significance level. Patients who prematurely discontinued the study medication because of the defined criteria of deterioration or because of treatment with not allowed concomitant prostanoid medication during the study period were to be counted as non-responders. Patients with missing information for the primary endpoint were to be counted as non-responders. Patients assigned a value of 0 meters for their baseline exercise capacity assessment who showed any improvement at 12 weeks were to be considered as having satisfied the first part of the combined responder criterion (improvement by 10%). Patients who received a lung or heart and lung transplantation during the study period were to be considered not assessable for efficacy if this was only due to transplant availability. Logistic regression methods were to be used to describe the relationship between response and prognostic factors, i.e. treatment, demographic data and other baseline characteristics, including the factors used for stratification.

The methods for secondary efficacy analyses:

Absolute and percent changes in 6-minute walking distance as compared to the baseline value were to be calculated for each patient. The main comparison for exercise capacity was to be between the baseline value and the peak value obtained at week 12. Patients unable to walk because of defined deterioration or death were to be assigned a value of 0 meters. Absolute and percent changes in exercise capacity were to be described for each treatment group and each visit by descriptive statistics and p-values from t-statistics for comparing the mean to zero. Comparison between the two treatment groups for changes in exercise capacity was to be based on non-parametric analysis of covariance.

NYHA class.
For analyses of changes in the NYHA functional classification between baseline and post baseline visits, shift tables were to be given. Treatment comparison for the NYHA class was to be performed with CMH statistics adjusting for baseline NYHA class (III or IV).

Dyspnea index.
The three individual categories as well as the scored data were to be summarized by descriptive statistics and frequency tables for each treatment group and visit.

Hemodynamic parameters and gas exchange.
Absolute and percent changes to the pre-inhalation value of hemodynamic parameters were to be described for each treatment group and each visit by descriptive statistics and p-values from t-test for comparing the mean to zero. Long-term effects (after 12 weeks) were to be outlined by non-parametric analysis of covariance as described for exercise capacity.

Deterioration of PHT, mortality, need for transplantation
All time-to-event data were to be described for each treatment group by Kaplan-Meier estimates. Treatment comparison was to be made with log-rank tests stratified for PHT (primary or secondary) and NYHA (III or IV).

Determination of sample size
The sample size calculation was based on the primary efficacy endpoint (responder criterion). Assuming a 10% placebo response and a 30% iloprost response, 100 patients per treatment arm provided more than 90% power (alpha = 0.05, two-sided).

Results of Study 97218

Disposition of Patients

Of the screened 235 patients, 203 patients were randomized. Table 1 shows premature discontinuations from the study.

Table 1. Premature discontinuation from Study ME97218

	Iloprost	Placebo	P-value*
Premature discontinuation of the study	9 (9%)	23 (23%)	0.011
due to: discontinuation of study medication	5 (5%)	10 (10%)	0.28
lost to follow up	0	1 (1%)	1.00
death	2 (2%)	6 (6%)	0.28
withdrawal of consent	0	2 (2%)	0.50
other	3 (3%)	4 (4%)	1.00

Reviewer's analysis. Fisher's exact test.

Intent-to-treat (ITT) population

All randomized patients (N = 203; iloprost: n = 101; placebo: n = 102) fulfilled the ITT criteria. Thus, the set of 203 randomized patients is also the ITT population which was used for the primary analysis of efficacy.

Per protocol (PP) population

The PP population consists of 192 patients (iloprost: n = 98, placebo: n = 94). The allocation of patients to strata in the PP analysis is based on the baseline CRF entries.

Safety population

The safety population is the same as the ITT population.

Stratification

The patients enrolled in the study were stratified by randomization for the type of PHT (PPH or SPH) and for the NYHA functional class (III or IV). Table 2 shows patient distribution over the four strata.

Table 2. Patient distribution over predefined strata after randomization. ITT population.

	PPH	SPH

NYHA functional class III	70		50	
	Iloprost: 34	Placebo: 36	Iloprost: 26	Placebo: 24
NYHA functional class IV	38		45	
	Iloprost: 19	Placebo: 19	Iloprost: 22	Placebo: 23

Source: Sponsor's Table TT 8.

Baseline main characteristics

The treatment groups were comparable with regard to weight (mean weight: iloprost: 71.3 ± 14.6 kg; placebo: 72.6 ± 13.9 kg), height (mean height: iloprost: 166.1 ± 9.3 cm; placebo: 165.6 ± 9.0 cm), and ethnic group (96.1% of all patients were Caucasian). In both groups, female to male ratio was 2:1. With respect to age, the treatment groups were also comparable (mean age: iloprost: 51.2 ± 13.2 years; placebo: 52.8 ± 12.0 years), with the exception of the PPH/IV groups where patients tended to be younger in the iloprost group than in the placebo group. In both treatment groups, SPH patients were older than PPH patients with the SPH/IV group showing the highest mean age.

Efficacy results

Primary endpoint

Primary efficacy analysis showed that the iloprost group was statistically significantly (p=0.0067) better than the placebo group relative to the combined responder criterion:

- Improvement in exercise capacity (6-minute walking test) at 12 weeks by at least 10% versus baseline and
- Improvement by at least one NYHA class at 12 weeks versus baseline, and
- No deterioration of PHT, or death at any time before 12 weeks.

The primary analysis of the combined responder criterion is a two-sided CMH test stratified for the four groups: type of PHT (PPH or SPH) and NYHA functional class (III or IV). Table 3 shows that the odds ratio (OR) is 3.97 with the 95% confidence interval (1.39, 11.31).

Table 3. Primary efficacy analysis* of Study ME97218 Combined responder criterion. ITT population

Iloprost N=101	Placebo N=102	p-value Odds Ratio (95% CI)
Responders	Responders	p=0.0067
17 (16.8%)	5 (4.9%)	OR=3.97 (1.39, 11.31)

** Reviewer's analysis. CMH test adjusted for NYHA class and PHT type.*

To support the primary efficacy analysis, the combined responder criterion is also examined by analysis of the:

- Predefined strata;

• Time course of responder rates.

Combined responder criterion in the predefined strata

Table 4. Combined responder criterion by predefined strata. ITT population				
Stratum	**Iloprost (N=101)**		**Placebo (N=102)**	
	Responders		**Responders**	
PPH/III	5/34	14.7%	2/36	5.6%
PPH/IV	6/19	31.6%	1/19	5.3%
SPH/III	5/26	19.2%	2/24	8.3%
SPH/IV	1/22	5.5%	0/23	0%
All	17/101	16.8%	5/102	4.9%

Source: Sponsor's Table TT 33.

The Breslow-Day test showed homogeneity (p=0.79) of odds ratios over the four strata (PPH/III, PPH/IV, SPH/III and SPH/IV). Table 4 shows the combined responder rates in the four predefined strata. Treatment was the only significant prognostic factor (p = 0.0095) found in the logistic regression model.

Time course of responder rates

Over time, the rate of the combined responder endpoint was rising in the iloprost group from 10.9% at week 4 to 16.8% at week 12. In the placebo group, combined responder rate was rising from 3.9% at week 4 to 4.9% at week 12 (Table 5).

Table 5. Combined responder rate by visit. ITT population		
Week	**Responder Iloprost**	**Responder Placebo**
Week 4	10.9%	3.9%
Week 8	12.9%	3.9%
Week 12	16.8%	4.9%

Source: Sponsor's Table TT 34.

Sensitivity analyses for the combined responder rate

The reason for the sensitivity analyses was the fact that at week 12, much more placebo patients had missing information of at least one of the components of the combined responder criterion and thus were counted as non-responders for the primary efficacy analysis. Namely, 9 placebo patients (#176, #411, #387, #11, #30, #53, #124, #240 and #267) were counted as non-responders for the combined responder endpoint as compared to two iloprost patients (#238 and #55). The following sensitivity analyses were performed to examine the robustness of the primary efficacy result:

- Analysis based on ITT-LOCF
- Analysis based on the Per Protocol population

ITT-LOCF results

A LOCF procedure imputing at week 12 the last known data for the patients in the ITT population with missing values, produced the same result as the primary analysis based on observed cases: the same 17 iloprost patients and 5 placebo patients met the combined responder criterion with p=0.0067.

Per Protocol (PP) Population results

All 17 iloprost patients and 5 placebo patients who met the combined responder criterion in the ITT population, also belong to the PP population. The CMH test results for the combined responder criterion based on the PP population are similar to those for the ITT population (treatment effect: p = 0.016; odds ratio: 3.487; 95% CI: (1.261, 9.646). Breslow-Day test showed homogeneity (p=0.813) of odds ratios over the four strata.

Secondary endpoints

Exercise capacity (6-minute walking test)

At week 12, the 6-minute absolute walking distance improved statistically significantly (p=0.032) more in the iloprost treated patients (mean= -22.2 meters) than in the placebo treated patients (mean = -3.3 meters). Similar result was for the percent change from baseline (p=0.032). As the baseline in the iloprost group was greater than in the placebo group, an ANCOVA adjusting for the baseline means was performed. ANCOVA adjusting for baseline walking distance found that iloprost was statistically significantly better than placebo: p=0.021.

NYHA class

Relative to the improvement by at least one NYHA functional class at week 12 versus baseline, the iloprost group was statistically significantly better than the placebo group (p = 0.032, stratified CMH test). The distribution of the NYHA functional class at week 12 compared with baseline was borderline significantly better for iloprost (p = 0.053, two-sided CMH statistics).

Dyspnea index

Relative to the Mahler focal score, the iloprost group was numerically better than the placebo group with respect to the mean absolute change at each time point. At week 12, p = 0.734 in the two-sided Kruskal-Wallis test. Concerning the Mahler transition score, the iloprost group improved statistically significantly better than the placebo group at week 12 (p=0.015, Kruskal-Wallis test).

Hemodynamic parameters (PVR, mPAP, CO) and gas exchange (SVO_2)

Effects on hemodynamic parameters and the gas exchange were measured during the week 12 catheter examination and compared to baseline catheter findings. With respect to the trough values of the acute drug effect obtained before inhalation, over the observation period the iloprost treated patients remained stable as compared to baseline (p>0.2) while the placebo group deteriorated with respect to PVR, CO and SVO_2. The deterioration in placebo group versus the baseline values was statistically significant ($p < 0.04$) except for mPAP (p=0.85). Concerning the treatment effect at trough, non-parametric ANCOVA found that Iloprost was statistically significantly better than placebo (p=0.045) relative to PVRI and numerically better relative to mPAP, CO and SVO_2 (p>0.0.26).

Deterioration of PHT

With respect to deterioration of PHT, numerically fewer iloprost patients developed the defined criteria of deterioration (p=0.24, log-rank test for treatment comparisons of time to the first deterioration).

Mortality

At 12 weeks, mortality was numerically lower in the iloprost group (1%, 1/101) than in the placebo group (4%, 4/102) with p=0.37 in Fisher's exact test, two-sided. Deaths were also analyzed beyond the cut-off at week 12. By the end of the 4-week safety follow-up, there were 5 more patients who died. Two of these patients had deterioration before or at week 12. In the placebo group 7 (7%) patients died and in the iloprost group 3 (3%) patients died (p = 0.33, Fisher's exact test).

The cases of death were homogeneously distributed over the four predefined strata (p = 0.99). With regard to the individual groups, no significant differences were observed in any predefined stratum (PPH/III: p = 0.32; PPH/IV: p = 0.99; SPH/III: p = 0.86; SPH/IV: p = 0.24; log-rank test).

Need for transplantation

By week 12, two iloprost patients (2.0%) and 4 placebo patients (3.9%) were newly scheduled for transplantation. Both iloprost patients belonged to the SPH/IV stratum. No patient was transplanted during the entire course of the study.

3.2 Evaluation of Safety

Extent of exposure and discontinuations are shown in Table 6. Statistically significantly (p=0.014) more iloprost patients completed the entire study. Statistically significantly (p=0.018) fewer iloprost patients discontinued study medication for any reason.

Table 6. Extent of exposure and discontinuations in Study ME97218

	Iloprost N=101	**Placebo N=102**	**P-value***
Completed the entire study (until week 16)	91	79	0.022
Discontinued study medication for any reason	7	18	0.031
Discontinued for adverse events	3	8	0.21

** Reviewer's analysis. Fisher's exact test.*

Overall the frequency of any adverse events reported at any time was high and similar in both groups up to week 12 and up to week 16 (Table 7). The rate of any adverse events during entire study was slightly higher as compared to the rate up to week 12. Mortality was lower in the iloprost group than in the placebo group for all time intervals. Up to week 12, mortality was: one iloprost patient as compared to 4 placebo patients. Up to week 16, mortality was: 3 iloprost patients as compared to 7 placebo patients.

Table 7. Overall rates of any adverse events during the treatment period (up to 12 weeks) and during the entire study period (up to 16 weeks)

	Iloprost N=101	Placebo N=102	P-value*
Treatment period	91 (90%)	90 (88%)	0.82
Entire study period	95 (94%)	93 (91%)	0.59

** Reviewer's analysis, Fisher's exact test*

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Combined responder criterion in the predefined subgroups

The Breslow-Day test showed homogeneity (p=0.79) of odds ratios over the four subgroups (PPH/III, PPH/IV, SPH/III and SPH/IV). In each of the four subgroups, Iloprost had a higher combined responder rate as compared to placebo. The most prominent treatment effect in terms of percent difference was seen between the PPH/IV groups and the smallest one between the SPH/IV groups (Table 8).

Table 8. Combined responder criterion by NYHA class and PHT type. ITT population				
Stratum	**Iloprost (N=101)**		**Placebo (N=102)**	
	Responders		**Responders**	
PPH/III	5/34	14.7%	2/36	5.6%
PPH/IV	6/19	31.6%	1/19	5.3%
SPH/III	5/26	19.2%	2/24	8.3%
SPH/IV	1/22	5.5%	0/23	0%
All	17/101	16.8%	5/102	4.9%

Source: Sponsor's Table TT 33

4.2 Change from baseline in walking distance and mortality in predefined subgroups.

Table 9 shows absolute change in walking distance from baseline to week 12 by type of PHT. Analysis of the change from baseline to week 12 in walking distance in the two PHT groups (PPH and SPH) suggests that in the PPH group, there is some beneficial effect of iloprost as compared to placebo. Non-parametric ANCOVA showed that for absolute change from baseline in walking distance, iloprost was numerically better than placebo in PPH/III subgroup (p=0.062) and in the PPH/IV subgroup (p=0.053). In the SPH group, there was no clear signal of the iloprost effect regarding the increase of the walking distance from baseline.

Table 9. Absolute change in walking distance (meters) from baseline to week 12 by type of PHT. ITT population. Walking test after inhalation.					
	Treatment group	**N**	**Mean**	**SD**	**Median**
PPH	Iloprost	49	45	74	31
	Placebo	46	-7	90	2
SPH	Iloprost	46	-2	61	11
	Placebo	39	2	50	0

Source: Sponsor's Table TT 39

Mortality analysis in the four pre-specified groups (PPH/III, PPH/IV, SPH/III and SPH/IV) showed that the cases of death were homogeneously distributed over the four predefined groups ($p = 0.99$). With regard to the individual groups, no significant differences were observed in any predefined stratum (PPH/III: $p = 0.32$; PPH/IV: $p = 0.99$; SPH/III: $p = 0.86$; SPH/IV: $p = 0.24$).

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Statistical and analytical issues

In general, this reviewer agrees with the sponsor's statistical methods. The reviewer's primary efficacy results are very close to the sponsor's results.

Primary efficacy analysis of Study ME97218 showed that iloprost is statistically significantly better than placebo relative to the combined responder rate at week 12 based on the ITT-observed cases (p=0.0067). Sensitivity analyses of the combined responder rate based on the ITT-LOCF and the Per Protocol populations support the primary efficacy results ($p \leq 0.016$). Secondary analyses of the change from baseline to week 12 in walking distance and of other secondary endpoints were consistent with the primary efficacy analysis. Although Study ME97218 seems to be a solid confirmatory study, for the non-mortality endpoints the usual standard requires efficacy evidence to be confirmed in at least one more study.

The sponsor did not pre-specify any statistical rule for dealing with multiple secondary endpoints. For this reason, there is no basis for selective inclusion of favorable p-values for some secondary endpoints in the sponsor's labeling. In this review the p-values for the secondary endpoints and subgroups are provided only to show the consistency in the results because there was no adjustment for multiple comparisons regarding the number of secondary endpoints, the number of subsets, or the number of methods.

No analyses for possible interactions by centers were planned, because this study was randomized within the strata constructed by PHT diagnosis and the NYHA class. The assessment of responder rates per center and country did not suggest center effects.

Subgroup analyses were performed for two pre-specified PHT subgroups (PPH and SPH) and four pre-specified subgroups (PPH/III, PPH/IV, SPH/III and SPH/IV).

5.2 Conclusions and Recommendations

Primary efficacy analysis of Study ME97218 showed that iloprost is statistically significantly better than placebo relative to the combined responder rate at week 12 based on the ITT-observed cases (p=0.0067). Sensitivity analyses of the combined responder rate based on the ITT-LOCF and the Per Protocol populations support the primary efficacy results ($p \leq 0.016$). Secondary analyses of the change from baseline to week 12 in walking distance and of other secondary endpoints were consistent with the primary efficacy analysis. Although Study ME97218 seems to be a solid confirmatory study, for the non-mortality endpoints the usual standard requires efficacy evidence to be confirmed in at least one more study.

As the sponsor did not pre-specify any statistical rule for dealing with multiple secondary endpoints, there is no basis for selective inclusion of favorable p-values for some secondary endpoints in the sponsor's labeling.

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

/s/

Valeria Freidlin
10/27/04 11:44:52 AM
BIOMETRICS

James Hung
10/27/04 11:50:27 AM
BIOMETRICS

This needs your immediate attention.

Kooros Mahjoob
10/28/04 02:27:45 PM
BIOMETRICS