U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

# STATISTICAL REVIEW AND EVALUATION

## CLINICAL STUDIES

| | |
|---|---|
| **NDA/BLA Serial Number:** | 203109 |
| **Drug Name:** | Revatio (sildenafil citrate) |
| **Indication(s):** | pediatric pulmonary arterial hypertension |
| **Applicant:** | Pfizer |
| **Date(s):** | 11/30/2011 |
| **Review Priority:** | priority |
| **Biometrics Division:** | Division of Biometrics I |
| **Statistical Reviewer:** | John Lawrence, PhD |
| **Concurring Reviewers:** | Jim Hung, PhD |
| **Medical Division:** | Cardiorenal HFD-110 |
| **Clinical Team:** | Maryann Gordon, MD |
| **Project Manager:** | Dan Brum |

**Keywords:**

multiple endpoints, survival analysis, surrogate outcomes

# Table of Contents

# 1 EXECUTIVE SUMMARY

There is one pediatric study where children were randomized to placebo or 2 doses of sildenafil (3 doses for middle and higher weight strata) and followed for 16 weeks for efficacy. The efficacy endpoints measured included exercise tolerance, hemodynamics and change in functional class. Subjects who completed the study were enrolled in an extension study to measure long term survival and other outcomes. In the short term study, the higher doses of sildenafil were associated with improvements in the short term efficacy endpoints, but none were statistically significant according to the pre-specified analysis plan. In the long term survival follow-up, the higher doses were associated with an increased rate of mortality. No dose studied was shown to be safe and effective in children.

My recommendation is to not approve any dose in children nor the new formulation which is intended for children. Furthermore, a long term dose response study of survival in adults is recommended.

# 2 INTRODUCTION

Sildenafil is approved in adults for the treatment of pulmonary arterial hypertension (PAH). The recommended dose in adults is 20 mg tid. The studies examined in this review were conducted to determine whether there is a safe and effective dose in children.

## 2.1 Overview

This review is for two studies of sildenafil in children with PAH. The studies are summarized in Table 1.

**Table 1: List of all studies included in analysis**

| Study | Phase and Design | Treatment Period | Follow-up Period | # of Subjects per Arm | Study Population |
|---|---|---|---|---|---|
| A1481131 | Phase 3 parallel, double-blind | 16 Weeks | 16 Weeks | 60 (placebo)<br>42 (low dose)<br>56 (middle dose)<br>77 (high dose) | children ages 1-17 with PAH |
| A1481156 | Phase 3, parallel, open label | indefinite | indefinite | 55 (low dose)<br>74 (middle dose)<br>100 (high dose) | children ages 1-17 with PAH |

## 2.2 Data Sources

Sponsor's study report and electronic datasets:
\\Cdsesub1\evsprod\NDA203109\0000\m5\datasets\a1481131\analysis\datasets

# 3 STATISTICAL EVALUATION

## 3.1 Data and Analysis Quality

No issues were identified with the quality of the datasets.

## 3.2 Evaluation of Efficacy

### 3.2.1 Study Design and Endpoints

Subjects were randomized to placebo or a dose of sildenafil for the short term study. The primary endpoint of the 16 week study was change from baseline in PVO2. Hemodynamic measurements and WHO functional class was ascertained at randomization and at the end of 16 weeks in most subjects.

At the end of the 16 week period, those subjects who were randomized to sildenafil stayed on their randomized dose and those subject randomized to placebo were randomized to a dose of sildenafil. Nearly all subjects from the short term study chose to remain in the long term study. The purpose of the long term study was to obtain long term safety and mortality data.

### 3.2.2 Patient Disposition, Demographic and Baseline Characteristics

The demographic and baseline characteristics are summarized in Table 1.

Table 1. Baseline and demographics of short term study.

| | Sildenafil | | | | |
|---|---|---|---|---|---|
| Dose | Low | Medium | High | Combined | Placebo |
| Number of subjects | 42 | 55 | 77 | 174 | 60 |
| Male, n (%) | 17 (40.5) | 24 (43.6) | 26 (33.8) | 67 (38.5) | 22 (36.7) |
| Female, n (%) | 25 (59.5) | 31 (56.4) | 51 (66.2) | 107 (61.5) | 38 (63.3) |
| Age (years), n (%): | | | | | |
| 1-4 | 0 | 9 (16.4) | 19 (24.7) | 28 (16.1) | 7 (11.7) |
| 5-12 | 25 (59.5) | 28 (50.9) | 36 (46.8) | 89 (51.1) | 37 (61.7) |
| 13-17 | 17 (40.5) | 18 (32.7) | 22 (28.6) | 57 (32.8) | 16 (26.7) |
| ≥18 | 0 | 0 | 0 | 0 | 0 |
| Race, n (%): | | | | | |
| White | 19 (45.2) | 26 (47.3) | 28 (36.4) | 73 (42.0) | 24 (40.0) |
| Black | 1 (2.4) | 1 (1.8) | 1 (1.3) | 3 (1.7) | 2 (3.3) |
| Asian | 6 (14.3) | 13 (23.6) | 15 (19.5) | 34 (19.5) | 7 (11.7) |
| Other | 18 (38.1) | 15 (27.3) | 33 (42.9) | 64 (36.8) | 27 (45.0) |
| Region, n (%): | | | | | |
| America[a] | 10 (23.8) | 11 (20.0) | 16 (20.8) | 37 (21.3) | 17 (28.3) |
| Asia | 6 (14.3) | 13 (23.6) | 15 (19.5) | 34 (19.5) | 7 (11.7) |
| Europe | 16 (38.1) | 18 (32.7) | 22 (28.6) | 56 (32.2) | 16 (26.7) |
| South America | 10 (23.8) | 13 (23.6) | 24 (31.2) | 47 (27.0) | 20 (33.3) |
| Mean weight (range), kg | 38.2 (20.0-105.0) | 32.1 (8.6-106.0) | 25.8 (8.2-61.0) | 30.8 (8.2-106.0) | 29.3 (9.1-60.0) |
| Mean height (range), cm | 141.6 (111.0-172.0) | 130.5 (77.0-192.5) | 120.8 (72.0-180.0) | 128.9 (72.0-192.5) | 128.4 (78.0-173.0) |
| Mean BMI (SD), kg/m$^2$ | 18.2 (4.8) | 17.6 (3.9) | 16.3 (3.4) | 17.2 (4.0) | 16.9 (3.6) |

BMI=body mass index
[a] America=includes USA, Canada and Mexico

Source: Table S3 of study report.

### 3.2.3 Statistical Methodologies

Missing PVO2 data was intended to be imputed by LOCF. However, most subjects who had missing Week 16 PVO2 had no post-baseline measurement and were not included in the analysis in any way. There were not many of these subjects. Out of 115 subjects developmentally able to perform the exercise test, 9 subjects had a baseline PVO2 but were not included in the analysis because they had no usable post-baseline measurement (one of these subjects had a Week 8 measurement, but it was not used as it was not measured at trough plasma concentration).

The number of changes to the study design and analysis during and after the study were vast and a detailed listing and criticism of all of them is beyond the scope of this review. But, particular mention should be made about the sample size and primary analysis plan. Initially, the sample size was planned to enroll 224 subjects developmentally able to exercise out of a total sample size of 332. The analysis was to compare PVO2 between each dose to placebo using Hochberg's multiple comparison procedure. At some point, the primary analysis was changed to compare the combined doses versus placebo and to reduce the number of subjects who could exercise to 204. Later (early 2007 according to the study report based on blinded interim estimate of variability of 54 subjects PVO2 data), the number was dropped to about 90 subjects developmentally able to exercise with a minimum of 200 total.

Doses were pooled to compare to placebo. This is in general not recommended because results are difficult to interpret and does not allow testing of whether specific doses are effective.

Hemodynamics were complicated because of the use of two different methods of calculating cardiac output. Four subjects (Subjects 10429, 10435, 11625 and 11621) had both baseline and Week 16 measurements of PVRI, but were not included in the analysis because the cardiac output method was not the same at both time points. The data from those subjects that were excluded appeared to go against the effectiveness of the drug, i.e. the sildenafil subjects tended to have worsening of PVRI while the placebo subjects tended to have improvement in PVRI.
Subject 10429: high dose, baseline = 2.6 Wood units*$m^2$, Week 16 = 3.89
Subject 10435: medium dose, baseline = 4.3, Week 16 = 6.4
Subject 11621: low dose, baseline not provided in dataset, Week 16 = 9.5
Subject 11625: placebo, baseline = 20.3, Week 16 = 8.9

Analyses were appropriately stratified or adjusted by weight strata.

Long term survival was appropriately analyzed using the group as randomized (ITT analysis) and Kaplan-Meier analyses were used to handle subjects at the end of follow-up.

### 3.2.4 Results and Conclusions

Table 2 shows the results for change in PV02- the primary efficacy endpoint. The pre-specified analysis plan was to compare the combined doses versus placebo. The p-value was 0.056, which was not statistically significant. However, there appeared to be a numerical dose response trend and the middle dose (but not the low or high dose) had an unadjusted 95% confidence interval that excluded 0 difference from placebo. There was no provision to interpret the individual doses compared to placebo if the combined analysis failed and to the sponsor's credit, no p-values for individual doses were reported in this table. As an aside, in the original analysis plan (using Hochberg's procedure), none of the doses including the medium dose would have been better than placebo.

Table 2. Change from Baseline in Peak Volume of Oxygen Consumed at Week 16.

| | Sildenafil | | | | |
|---|---|---|---|---|---|
| Dose | Low | Medium | High | Combined | Placebo |
| Number of subjects[a] | 24 | 26 | 27 | 77 | 29 |
| Mean (SD) $VO_2$, mL/kg/minute | | | | | |
| Baseline[b] | 17.37 (4.36) | 18.03 (4.70) | 17.43 (3.70) | 17.61 (4.22) | 20.02 (3.80) |
| Week 16 | 18.40 (5.61) | 20.39 (6.16) | 19.00 (3.59) | 19.28 (5.21) | 20.01 (4.44) |
| Change from baseline | 1.03 (3.41) | 2.36 (3.36) | 1.57 (2.56) | 1.67 (3.13) | -0.01 (3.34) |
| Percentage change from baseline | 6.44 (20.16) | 13.40 (19.50) | 10.58 (15.51) | 10.24 (18.39) | 0.53 (15.91) |
| Mean difference versus placebo (SE)[c] | 3.81 (5.00) | 11.33 (4.84) | 7.98 (4.85) | 7.71 (3.98) | NA |
| 95% Confidence interval[c] | -6.11, 13.73 | 1.72, 20.94 | -1.64, 17.60 | -0.19, 15.60 | NA |
| P-value[c] | NA | NA | NA | 0.056 | NA |

LOCF=last observation carried forward; ITT=intention-to-treat population; SE=standard error; SD=standard deviation; NA=not applicable

[a] ITT subset of developmentally able subjects

[b] Baseline was the average of all assessments on or before the first day of study treatment

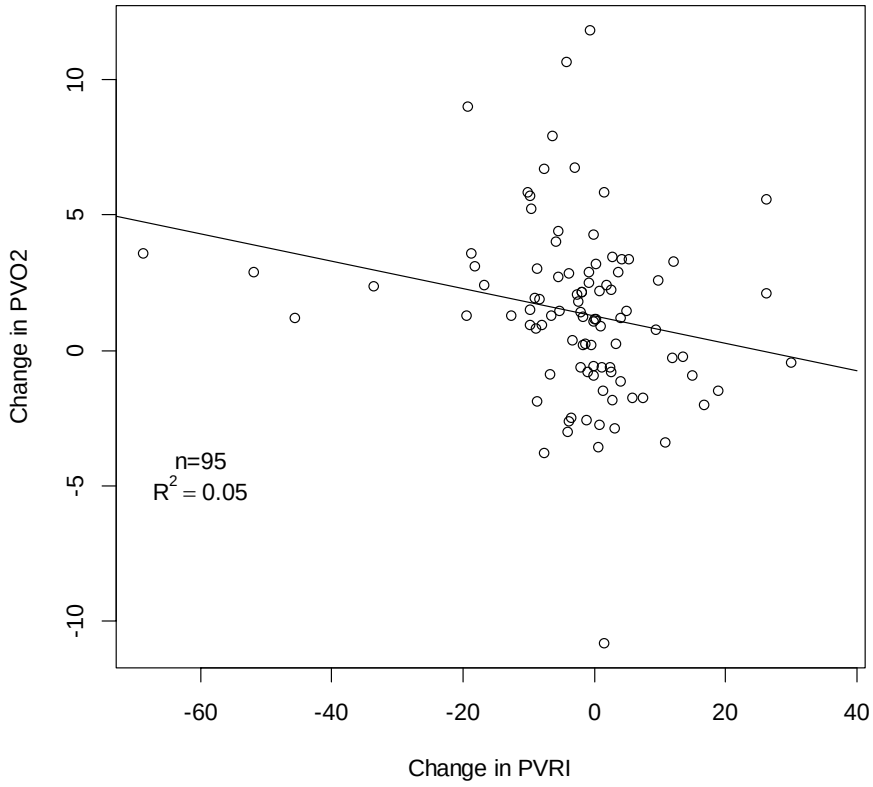
[c] Analyses were performed using analysis of covariance with etiology, weight and baseline peak $VO_2$ as the covariates

Source: Table S4 of study report and confirmed by reviewer.

PVRI had a similar numerical trend towards a dose response. Refer to the Statistical Review for IND 63,175 from November 2010 for a detailed discussion about the hemodynamic measurements and their relationship to exercise measurements in adults with PAH. In summary, when comparing relationships among individuals, the coefficient of determination ($R^2$) for adult subjects with idiopathic etiology (the only data looked at) is about 4% and there are many trials where an effect was shown on PVRI but not on 6 minute walk distance or vice versa (effect on 6MWD but not PVRI). Both of these make PVRI a poor surrogate marker for 6MWD. In this study, the estimated mean difference from placebo for the low, middle, and high doses in change in PVRI were -0.6, -4.5, and -7.2 Wood units*$m^2$. The unadjusted p-value for the combined doses compared to placebo was 0.041.

Figure 1 shows the relationship between PVRI and PVO2 for the 95 subjects who had both measurements. The figure shows there is very little correlation ($R^2 = 0.05$) between change in PVRI and change in PVO2. This small correlation is similar to what was observed in the adult data (correlation between PVRI and 6 minute walk distance in adults).

Figure 1. Relationship between change from baseline in PVRI (Wood units*$m^2$) versus change from baseline in PVO2 (mL/kg/min).




Source: FDA analysis.

Similar trends for short term efficacy of the middle and high dose (and no efficacy of the low dose) were seen with other endpoints measured.

In the long term follow-up, some subjects had a 1 year PVO2 measurement. For those subjects, the results are shown in Table 3. This is only a subset of the subjects who had PVO2 measured at Week 16, but in this subset, the trend reversed so that the High and Middle doses were worse than the Low dose on average.

Table 3. Results of PVO2 (mL/kg/min) at 1 year.

| | Sildenafil Low Dose (N=33) | Sildenafil Medium Dose (N=32) | Sildenafil High Dose (N=35) |
|---|---|---|---|
| **Baseline** | | | |
| Mean (SD) | 18.30 (4.54) | 18.11 (4.44) | 17.78 (3.65) |
| **Year 1** | | | |
| Mean (SD) | 19.97 (5.17) | 18.69 (5.92) | 17.93 (4.02) |
| Mean (SD) Change from Baseline | 1.67 (3.64) | 0.58 (5.22) | 0.15 (3.44) |
| Mean (SD) % Change from Baseline | 11.19 (22.98) | 5.37 (31.62) | 2.56 (21.46) |
| Comparison with Low Dose: | | | |
| Mean Difference (SE) | | -7.02 (6.10) | -9.84 (5.92) |
| 95% Confidence Interval | | -19.13, 5.09 | -21.60, 1.93 |
| p-value | | 0.253 | 0.100 |
| Comparison with Medium Dose: | | | |
| Mean Difference (SE) | | | -2.82 (6.01) |
| 95% Confidence Interval | | | -14.75, 9.11 |
| p-value | | | 0.640 |

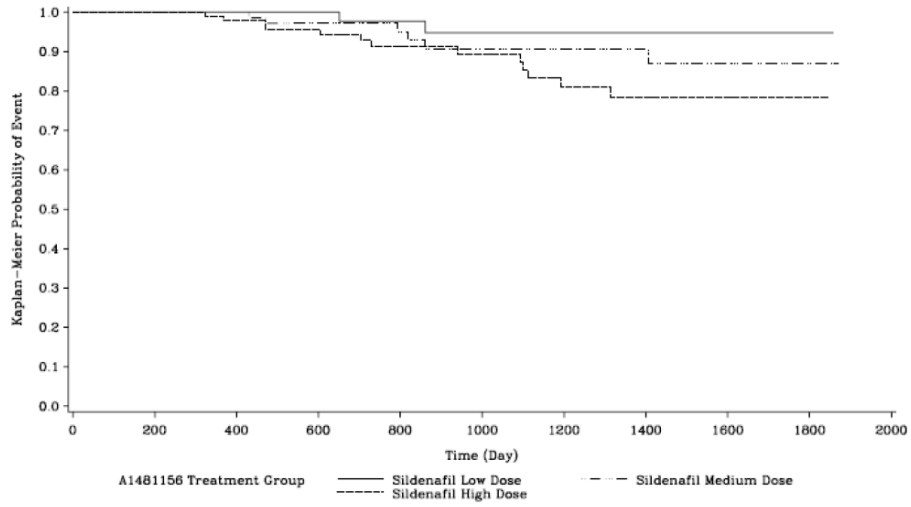
Source: Study Report Table T13.

## 3.3 Evaluation of Safety

See the Medical Officer review for safety endpoints other than mortality.

For long term survival, the Kaplan-Meier curves are shown in Figure 2. The low dose had the best survival rate, followed by the middle dose, while the high dose had the worst survival rate.

Figure 2. Kaplan-Meier estimates of survival curves.




Source: Figure F3 of study report.

The number of deaths in each of the treatment groups was 20/100 (20%) in the high dose treatment group, 10/74 (13.5%) in the medium dose treatment group, and 5/55 (9%) in the low dose treatment group, respectively.

In a proportional hazards model stratified by weight class assuming a linear relationship among the doses, the estimated hazard ratio of mortality comparing middle dose to low dose is 1.89 [p=0.008; confidence interval = (1.18, 3.03)]. Because of the assumed linear relationship, this is the same estimate of the hazard ratio for high dose compared to middle dose. The estimated hazard ratio of high dose compared to low dose is about 3.6. Note that this model assumes a linear relationship with dose level. Without assuming any model across doses and still stratifiying by weight class, the estimated hazard ratio of high dose compared to medium dose (using only the data from these two doses) is approximately 2.0 (not statistically significant); the estimated hazard ratio of medium dose compared to low dose is also approximately 2.0 (not statistically significant); the estimated hazard ratio of high dose compared to low dose is approximately 3.5 (p=0.015).

### 3.4 Benefit:Risk Assessment (Optional)

Not applicable.

## 4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

### 4.1 Gender, Race, Age, and Geographic Region

Not applicable.

### 4.2 Other Special/Subgroup Populations

Not applicable.

## 5 SUMMARY AND CONCLUSIONS

### 5.1 Statistical Issues and Collective Evidence

Refer to the Statistical Review for IND 63,175 from November 2010 for a detailed discussion about the utility of hemodynamic measurements and their relationship to exercise measurements in adults with PAH.

The study was complicated by not having a low dose arm for the lower weight strata. It would have been better to have a formulation used in the study for children that allowed three doses separated by factors of about three for each weight strata as was requested in the Written Request letters. Sensible analyses were planned that correctly stratified by weight strata, but the low weight group provided no information about the low dose because there was no low dose.

## 5.2 Conclusions and Recommendations

No dose studied was shown to be safe and effective for children. The short term efficacy results show no hint of efficacy for the low dose and a numerical trend toward efficacy for the middle and high doses studied. In the long term study, the middle and high doses had higher mortality rates than the low dose with no placebo group with which to compare any of the doses with respect to survival. The recommendation is to not approve any dose in children nor the new formulation which is intended for children. Furthermore, a long term dose response study of survival in adults is recommended.

---

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

---

/s/

---

JOHN P LAWRENCE
04/29/2012

HSIEN MING J HUNG
04/30/2012