CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

209279Orig1s000

STATISTICAL REVIEW(S)

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 #:	NDA 209279
Drug Name:	Tracleer (bosentan)
Indication(s):	Pulmonary arterial hypertension (PAH)
Applicant:	Actelion
Date(s):	Received: 08/05/2016 PDUFA: 06/05/2017
Review Priority:	Standard
Biometrics Division:	Biometrics I, HFD-710
Statistical Reviewer:	Jialu Zhang, Ph.D.
Concurring Reviewers:	James Hung, Ph.D.
Medical Division:	Division of Cardiovascular and Renal Products, HFD-110
Clinical Team:	Maryann Gordon, M.D.
Project Manager:	Maryam Changi

Keywords:

Link to keywords:

http://intranetapps.fda.gov/scripts/ob_apps/ob/eWork/uploads/eWork/2009/Keywords-in-DFS.htm

Table of Contents

1 EXECUTIVE SUMMARY 4

2 INTRODUCTION 5

2.1 OVERVIEW 5
2.2 DATA SOURCES 6

3 STATISTICAL EVALUATION 6

3.1 DATA AND ANALYSIS QUALITY 6
3.2 EVALUATION OF EFFICACY 6
3.2.1 AC-052-356 (BREATHE-3) 6
3.2.2 AC-052-365 (FUTURE 1) and AC-052-367 (FUTURE 2) 15
3.2.3 AC-052-373 (FUTURE 3) 18
3.3 SAFETY 21

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 21

4.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION 21
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS 21

5 SUMMARY AND CONCLUSIONS 21

5.1 STATISTICAL ISSUES 21
5.2 COLLECTIVE EVIDENCE 22
5.3 CONCLUSIONS AND RECOMMENDATIONS 22

LIST OF TABLES

Table 1: List of all studies included in analysis....5
Table 2: Demographic and Baseline Characteristics in BREATHE-3....9
Table 3: PVR Change from Baseline to Week 12 in BREATHE-3....10
Table 4: WHO FC Change from Baseline to Week 12 in BREATHE-3....10
Table 5: Overview of PVR Data in Adult and Pediatric Trials....11
Table 6: Comparison of Summary Statistics between Our Datasets and Sponsor's....11
Table 7: Patient Demographic and Baseline Characteristics in FUTURE 1 and FUTURE 2....17
Table 8: WHO Functional Class Change from Baseline to EOS in FUTURE 1 and FUTURE 2....18
Table 9: Patient Demographic and Baseline Characteristics in FUTURE 3....20
Table 10: Change in WHO Functional Class in FUTURE 3 and FUTURE 4....20

LIST OF FIGURES

Figure 1: Individual-level Surrogacy Based on 3 Bosentan Adult Trials....12
Figure 2: Trial-level Surrogacy Based on 3 Bosentan Adult Trials....13
Figure 4: Relationship between Sample Size Per Arm and Prediction Interval....15
Figure 5: Patient Disposition in FUTURE 1 and FUTURE 2....16
Figure 6: Patient Disposition in FUTURE 3....19

1 EXECUTIVE SUMMARY

The main focus in this review is to bridge the adult trials and the non-controlled BREATHE-3 pediatric trial for predicting a potential treatment effect of bosentan on six-minute walk distance (6MWD) in pediatric pulmonary arterial hypertension (PAH) patients aged between 3–15 years. The key question is whether bosentan would show a significant treatment effect on improving 6MWD if there were a placebo arm in the pediatric trial and 6MWD were measured.

There are two parts in this question. First, how should we model the relationship between 6MWD and pulmonary vascular resistance (PVR) in the bosentan adult trials? Secondly, what should we do to explore the treatment effect of bosentan on PVR since BREATHE-3 did not have a placebo arm?

To answer the first question, a bivariate fixed-effects model was applied to 3 bosentan adult trials (AC-052-351, AC-052-364 and AC-052-405) to estimate the overall relationship between PVR and 6MWD.

To answer the second question, bootstrapping was used to generate PVR data for a putative placebo arm using the PVR data of the placebo arm in a sildenafil pediatric trial.

All the calculations and simulations are based on at least two critically important assumptions:

1. The relationship between PVR and 6MWD in the pediatric trial is the same as in the adult trials.
2. If there is a placebo arm in BREATHE-3, the distribution of PVR change from baseline in that placebo arm would be the same as in other pediatric trials (such as in the sildenafil pediatric trial)

The individual-level and trial-level surrogacy can be examined in the bivariate fixed-effects model. In this case, the individual-level R^2 is 0.37 and the trial-level R^2 is 0.65, which indicates that PVR may not be considered as a reliable surrogate for 6MWD in the bosentan adult trials.

Using bootstrap sampling to generate a placebo arm for BREATHE-3 based on the placebo data in the sildenafil pediatric trial, the simulation showed that 1) there is only at most 50% chance to demonstrate a significant treatment effect by bosentan on PVR if the sample size is 17 for each treatment arm, and 2) the bosentan's effect on 6MWD relative to a putative placebo is predicted to be 59.6 meters with 95% prediction interval (-25.7, 144.9). If bootstrap sampling were applied for the bosentan arm, then the prediction interval would have been even wider.

In summary, we are not able to predict that bosentan would have demonstrated a statistically significant treatment effect on 6MWD, had BREATHE-3 had a placebo arm of similar sample size and had 6MWD been measured in these pediatric patients.

2 INTRODUCTION

2.1 Overview

Bosentan was approved by the Agency on November 20, 2001 for treating pulmonary arterial hypertension (PAH). This NDA is to seek approval of bosentan for treating pediatric patients < 12 years of age with pulmonary arterial hypertension (PAH).

The efficacy evaluation in this NDA is based on 'bridging' data between the adult and pediatric studies primarily utilizing the pivotal, Phase 3 trial, AC-052-356 (BREATHE-3), in which both adult and pediatric patients were dosed with 62.5 mg and 125 mg bosentan film-coated tablets.

This review covers several pediatric studies. AC-052-356 (BREATHE-3) is a pediatric study with bosentan film-coated tablets. FUTURE1, FUTURE2 and FUTURE3 are pediatric studies with bosentan dispersible tablets. FUTURE2 is the extension of FUTURE1.

Table 1: List of all studies included in analysis

	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
AC-052-356 (BREATHE-3)	Pediatric trial, open-label, non-controlled	12 weeks	Month 6 and Month 12, and every 3 months until EOS	19 total	pediatric PAH patients 3–15 years (primary or related to scleroderma or congenital heart defects and of WHO FC II–III)
AC-052-365 (FUTURE 1)	Pediatric trial, single-arm, open-label	12 weeks	NA	36 total	Pediatric patients aged ≥ 2 and < 12 years of age with WHO FC II or III idiopathic or familial PAH
AC-052-367 (FUTURE 2)	open-label, uncontrolled extension	continued until the formulation was approved	28 days after last dose	33 total	Pediatric patients aged ≥ 2 and < 12 years of age with WHO FC II or III idiopathic or familial PAH
AC-052-373 (FUTURE 3)	Pediatric trial, open-label, non-controlled	24 weeks	60 days	64 total	children with PAH from ≥ 3 months to < 12 years of age

2.2 Data Sources

BREATHE-3 datasets are located at \\CDSESUB1\evsprod\NDA209279\0000\m5\datasets\ac-052-356\analysis\legacy\datasets.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The reviewer was able to replicate the main efficacy results reported by the sponsor.

3.2 Evaluation of Efficacy

The main focus in this review is on bridging from the adult trials to the non-controlled BREATHE-3 pediatric trial to make a prediction of bosentan's effect. The question is whether bosentan might show a statistically significant effect on improving 6MWD if there was a placebo arm in the trial and 6MWD was measured.

3.2.1 AC-052-356 (BREATHE-3)

3.2.1.1 Study Design and Endpoints

BREATHE-3 is a multicenter, open-label, non-controlled study in pediatric PAH patients.

3.2.1.2 Statistical Methodologies

To predict the treatment effect of bosentan on 6MWD in the pediatric study BREATHE-3 based on the treatment effect of bosentan on PVR, I used bivariate fixed-effects model following the approach of Buyse et al (2000)[1]. The relationship between 6MWD and PVR is based on 3 bosentan adult trials. There are 4 bosentan adult trials, e.g., AC-052-351, AC-052-355 (BREATHE-2), AC-052-364 (EARLY) and AC-052-405 (BREATHE-5). Out of the 4 adult trials, three trials included a placebo arm. BREATHE-2 has epoprostenol rather than placebo as the control arm. In the bivariate fixed-effects model, the intercepts in the individual linear

[1] Buyse M., Molenberghs G., Burzykowski, T., Renard, D. and Geys, H. (2000). The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics, 1,1, pp. 49-67

regression models in Stage 1 are assumed to come from the same distribution. Therefore, BREATHE-2 was excluded from the analysis.

A bivariate fixed-effects model as following is fitted for the first stage

$$S_{ij} = \mu_{si} + \alpha_i Z_{ij} + \varepsilon_{sij}$$
$$T_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij}$$

S_{ij} is the surrogate endpoint value for patient j in trial i. T_{ij} is the clinical endpoint value for patient j in trial i. In this case, S is PVR and T is 6MWD. The first step is to look at the trial-[2]specific treatment effect on both the surrogate endpoint (PVR) and the clinical endpoint (6MWD). μ_{Si} and μ_{Ti} are the PVR value and 6MWD value in the control group in trial i. α_i and β_i are bosentan's effects on PVR and 6MWD, respectively.

In the second stage, the estimates based on the individual trials are used to provide an estimate of trial-level treatment effects.

$$\hat{\beta}_i = \lambda_0 + \lambda_1 \hat{\mu}_{Si} + \lambda_2 \hat{\alpha}_i + \varepsilon_i$$

$\hat{\beta}_i$, $\hat{\mu}_{Si}$, and $\hat{\alpha}_i$ are parameter estimates based on the model fitted in Stage 1.

Let D matrix be the following

$$D = \text{var}\begin{pmatrix} \hat{\mu}_{Si} \\ \hat{\mu}_{Ti} \\ \hat{\alpha}_i \\ \hat{\beta}_i \end{pmatrix} = \begin{pmatrix} d_{SS} & d_{ST} & d_{Sa} & d_{Sb} \\ & d_{TT} & d_{Ta} & d_{Tb} \\ & & d_{aa} & d_{ab} \\ & & & d_{bb} \end{pmatrix}$$

The treatment effect on T based on the treatment effect on S in a new trial to be predicted is given by

$$E(\beta_0 \mid \mu_{S0}, \alpha_0) = \beta + \begin{pmatrix} d_{Sb} \\ d_{ab} \end{pmatrix}^T \begin{pmatrix} d_{SS} & d_{Sa} \\ d_{Sa} & d_{aa} \end{pmatrix}^{-1} \begin{pmatrix} \mu_{S0} - \mu_S \\ \alpha_0 - \alpha \end{pmatrix}$$

$$Var(\beta_0 \mid \mu_{S0}, \alpha_0) = d_{bb} + \begin{pmatrix} d_{Sb} \\ d_{ab} \end{pmatrix}^T \begin{pmatrix} d_{SS} & d_{Sa} \\ d_{Sa} & d_{aa} \end{pmatrix}^{-1} \begin{pmatrix} d_{Sb} \\ d_{ab} \end{pmatrix}$$

A $(1-\gamma)100\%$ prediction interval can be obtained as $E(\hat{\beta}_0 \mid \mu_{S0}, \alpha_0) \pm z_{1-\gamma/2}\sqrt{Var(\hat{\beta}_0 \mid \mu_{S0}, \alpha_0)}$.

Simulation in Tibaldi 2003[3] showed that fixed effect approach has minor loss in statistical efficiency over a full random-effects approach.

Due to the limited number of trials, only reduced bivariate fixed model can be used. In the reduced model, instead of assuming each trial has its own intercept (PVR effect in placebo arm), the reduced model assumes that all trials have the same intercept. The following model is fitted for the first stage:

$$S_{ij} = \mu_s + \alpha_i Z_{ij} + \varepsilon_{sij}$$

$$T_{ij} = \mu_T + \beta_i Z_{ij} + \varepsilon_{Tij}$$

The second stage linear model was reduced to:

$$\hat{\beta}_i = \lambda_0 + \lambda_1 \hat{\alpha}_i + \varepsilon_i$$

In order to predict the 6MWD in the pediatric trial, the estimate of PVR in the placebo and the estimate of the treatment effect of PVR in the pediatric study are needed. BREATHE-3 does not have a placebo arm. Several critically important assumptions are needed:

1. The relationship between PVR and 6MWD in pediatrics is the same as in adults
2. If there were a placebo arm in BREATHE-3, the PVR distribution in the placebo arm would be the same as that in other pediatric trials (such as in sildenafil pediatric trial)

A number of analyses were performed to look at the data in multiple ways.

First, the mean PVR change from baseline in placebo arm in the sildenafil pediatric trial and the mean PVR change from baseline in BREATHE-3 are used to predict the treatment effect on 6MWD in BREATHE-3 trial.

The reviewer also used bootstrap to generate 17 PVR measurements for placebo pediatric patients (i.e., assuming that sample sizes for the treatment and placebo groups are in 1:1 ratio), had there been a placebo group. The generation uses the PVR distribution of the 50 placebo patients in sidenafil pediatric trial.

Let's assume that M sets of placebo patients were generated by bootstrapping. By Rubin's formula, for each estimate $\hat{\beta}_{0k}$ with standard error s_k, k =1,2,…M, the overall combined estimate is $\overline{\beta}_0$ =mean ($\hat{\beta}_{0k}$), and the standard error is given by

$$\sqrt{\frac{1}{M}\sum_k {s_k}^2 + \left(1+\frac{1}{M}\right)\left(\frac{1}{M-1}\right)\sum_k \left(\hat{\beta}_{0k} - \overline{\beta}_0\right)^2}$$

[3] Tibaldi, F. S., Abrahantes, J. C., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., Stijnen, T. and Wolfinger, R. (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. Journal of statistical computation and simulation, vol 73, No. 9, pp 643-658.

Bootstrap was also used to generate both placebo and bosentan pediatric patients with different sample sizes. Placebo patients were sampled from placebo arm in the sildenafil pediatric trial and bosentan patients were sampled from BREATHE-3. The relationship between prediction interval and sample size in BREATHE-3 trial was examined by increasing hypothetical sample size in both placebo and bosentan arms.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 19 pediatric patients were enrolled in the study. Of the 19 enrolled patients, 2 patients prematurely discontinued the study due to elevated liver aminotransferases. Only 17 patients had values in change from baseline in PVR. The majority of patients were female (52.6%), Caucasian (78.9% overall), and had a median age of 10 years (range: 3–15 years). The median time from diagnosis to enrollment was 61.1 months (range: 0.5–184.5 months). All patients belonged to either WHO FC II (78.9%) or FC III (21.1%) at baseline.

Table 2: Demographic and Baseline Characteristics in BREATHE-3

	Patients 10 - 20 kg N=7	Patients >20 - 40 kg N=6	Patients > 40 kg N=6	All N=19
SEX [n (%)]				
n	7	6	6	19
Males	4 57.1%	2 33.3%	3 50.0%	9 47.4%
Females	3 42.9%	4 66.7%	3 50.0%	10 52.6%
AGE (years)				
n	7	6	6	19
Mean	5.7	10.0	14.2	9.7
Standard deviation	1.9	2.4	1.2	4.0
Median	6.0	10.5	14.5	10.0
Min , Max	3.0 , 8.0	6.0 , 13.0	12.0 , 15.0	3.0 , 15.0
AGE [n (%)]				
n	7	6	6	19
< 8 years	5 71.4%	1 16.7%	-	6 31.6%
8 - 12 years	2 28.6%	4 66.7%	1 16.7%	7 36.8%
13 - 17 years	-	1 16.7%	5 83.3%	6 31.6%
RACE [n (%)]				
n	7	6	6	19
Caucasian/white	6 85.7%	6 100%	3 50.0%	15 78.9%
Other	1 14.3%	-	3 50.0%	4 21.1%
Etiology of PAH [n (%)]				
n	7	6	6	19
Primary Pulmonary Hypertension	3 42.9%	4 66.7%	3 50.0%	10 52.6%
Congenital Heart Defect	4 57.1%	2 33.3%	3 50.0%	9 47.4%
WHO grade at Baseline [n (%)]				
n	7	6	6	19
II	7 100%	5 83.3%	3 50.0%	15 78.9%
III	-	1 16.7%	3 50.0%	4 21.1%

[Source: Sponsor's Summary of Clinical Efficacy Table 3-5, confirmed by the reviewer]

3.2.1.4 Results and Conclusions

The mean PVR change from baseline to Week 12 was -389 with 95% CI (-706, -72). During the 12 weeks of treatment, 5 patients (26.3%) had improved WHO FC. One patient worsened in WHO FC and the rest of 13 patients remained unchanged.

Table 3: PVR Change from Baseline to Week 12 in BREATHE-3

	Patients 10 - 20 kg N=7		Patients >20 - 40 kg N=6		Patients > 40 kg N=6		All N=19	
Baseline								
n	5		6		6		17	
Mean	1668		1060		934		1194	
Standard deviation	1151		435		496		755	
95% CL of mean	239 ,	3097	604 ,	1517	413 ,	1455	806 ,	1582
Median	1479		1102		932		1194	
95% CL of median	408 ,	3100	343 ,	1674	283 ,	1438	613 ,	1438
Min , Max	408 ,	3100	343 ,	1674	283 ,	1438	283 ,	3100
Week 12								
n	5		6		6		17	
Mean	822		730		866		805	
Standard deviation	312		493		695		505	
95% CL of mean	435 ,	1210	213 ,	1247	137 ,	1596	546 ,	1065
Median	831		708		726		794	
95% CL of median	366 ,	1224	106 ,	1538	193 ,	2173	467 ,	928
Min , Max	366 ,	1224	106 ,	1538	193 ,	2173	106 ,	2173
Change from baseline								
n	5		6		6		17	
Mean	-846		-330		-68		-389	
Standard deviation	906		168		440		616	
95% CL of mean	-1971 ,	280	-507 ,	-154	-529 ,	394	-706 ,	-72
Median	-537		-292		-111		-266	
95% CL of median	-1876 ,	-26	-620 ,	-136	-523 ,	751	-523 ,	-90
Min , Max	-1876 ,	-26	-620 ,	-136	-523 ,	751	-1876 ,	751

[Source: Sponsor's AC-052-356 Final Study Report Appendix 11, confirmed by the reviewer]

Table 4: WHO FC Change from Baseline to Week 12 in BREATHE-3

				Week 12							
Treatment	n	Baseline WHO class	n	I No.	%	II No.	%	III No.	%	IV No.	%
All	19	I	-	-		-		-		-	
		II	15	2	10.5%	12	63.2%	1	5.3%	-	
		III	4	-		3	15.8%	1	5.3%	-	
		IV	-	-		-		-		-	

[Source: Sponsor's Summary of Clinical Efficacy Table 3-18, confirmed by the reviewer]

PVR and 6MWD in Adult Trials

There are 4 bosentan adult trials, e.g., AC-052-351, AC-052-355 (BREATHE-2), AC-052-364 (EARLY) and AC-052-405 (BREATHE-5). The mean change in PVR from baseline in each of the four adult bosentan trials is shown in **Table 5**.

Table 5: Overview of PVR Data in Adult and Pediatric Trials

Studies (n/N)	**PVR change** Mean change from baseline, 95% CLs (dyn·sec/cm^5) or mean ± SD (dyn·sec-m^2/cm^5)
Adult studies	
AC-052-351 (19/ 21)[1]	**–223.5** (–341.4, –105.6)
EARLY (80/93)[1]	**–69.4** (–175.0, 36.3)
BREATHE-5 (36/37)[1]	**–253.6** (–478.2, –28.9)
BREATHE-2 (20/22)*	**–563.5** (–800.3, -326.7)
Pediatric studies	
BREATHE-3 (17/19)[1]	**–389.2** (–706.0, –72.4)
FUTURE 3 (4/10)[2, 3]	**–55.0** ± 77.7 (b.i.d.) **–160.5** ± 226.98 (t.i.d.)
AC-052-377 (Japanese 6/6)[2, 3]	**–4.0** ± 258.6

[1] Film-coated tablets, [2] dispersible tablets, [3] PVR index calculated.
*Results to be interpreted taking into consideration the fact that the start of bosentan was concomitant to epoprostenol administration in the study.
[Source: Sponsor's Summary of Clinical Efficacy Table 1-3]

In the bivariate fixed-effects model, the intercepts in the individual linear regression models in Stage 1 are assumed to come from the same distribution. Therefore, BREATHE-2 was excluded from the model since its control arm is epoprostenol rather than placebo. The relationship between 6MWD and PVR is based on 3 bosentan adult trials.

The adult trial datasets we used to develop the relationship between 6MWD and PVR are mostly the same as reported by the sponsor. There does have some differences between the summary statistics based on the dataset that the reviewer used and what the sponsor reported in their CSR on BREATHE-5.

Table 6: Comparison of Summary Statistics between Our Datasets and Sponsor's

Study	results in Sponsor's CSR		Our datasets	
	n/N	PVR change from baseline	n	PVR change from baseline
AC-052-351	19/21	-223.5 (-341.4, -105.6)	19	-223.5 (-341.4, -105.6)
AC-052-364 (EARLY)	80/93	-69.4 (-175.0, 36.3)	80	-69.4 (-175.0, 36.3)
AC-052-405 (BREATHE-5)	36/37	-253.6 (-478.2, -28.9)	36	-316.9 (-597.8, -36.1)

Figure 1 and Figure 2 showed the individual-level and trial-level surrogacy using the above three bosentan adult trials. Individual-level surrogacy, which is quantified as individual-level R^2, measures the association between both endpoints after adjustment for the treatment effect. Trial-level surrogacy is quantified as trial-level R^2. A large trial-level surrogacy would indicate that a good portion of treatment effect can be captured by the surrogate.

A good surrogate endpoint would be able to predict the treatment effect on the true endpoint with sufficient precision. In this case, the individual-level R^2=0.37 and the trial-level R^2=0.65. From this point of view, PVR may not be considered as a reliable surrogate for 6MWD in the adult trials.

Figure 1: Individual-level Surrogacy Based on 3 Bosentan Adult Trials

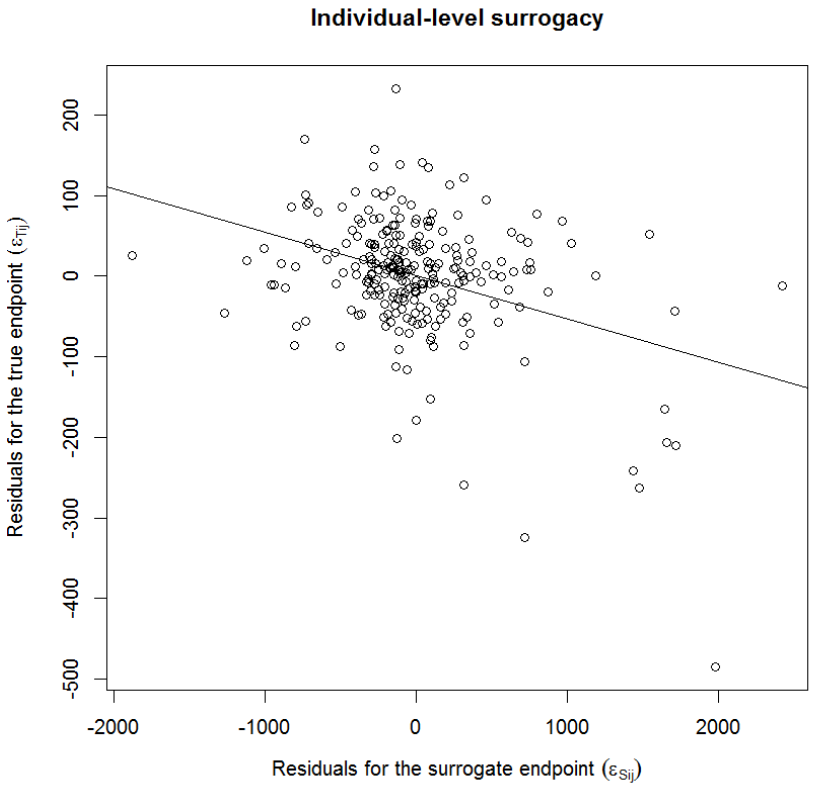
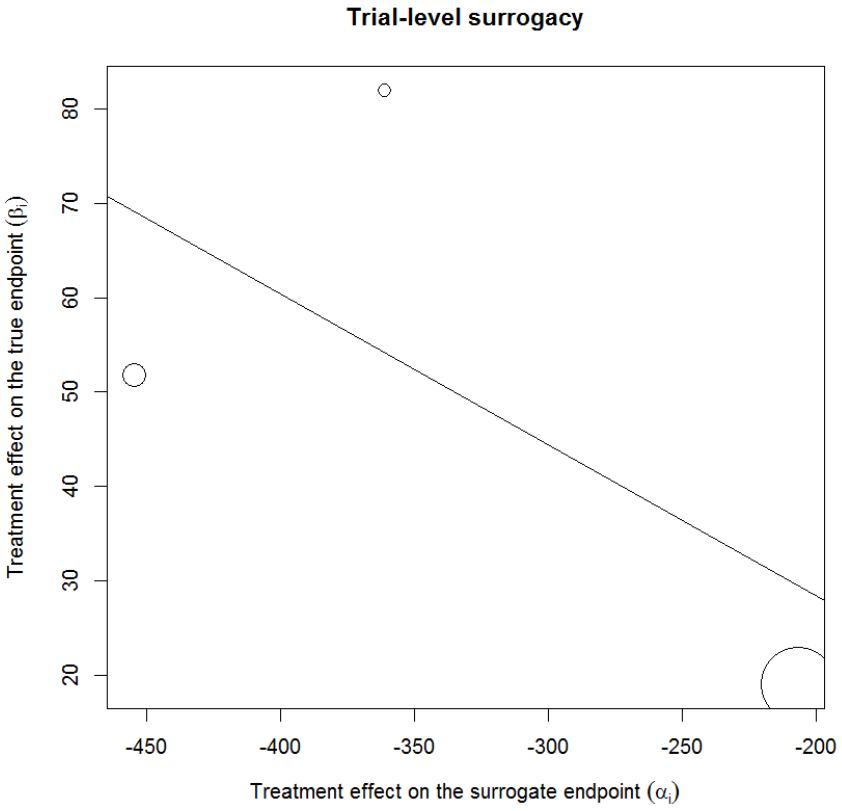


Figure 2: Trial-level Surrogacy Based on 3 Bosentan Adult Trials



Using point estimate from placebo arms in adult trials

The mean PVR change from baseline in placebo arm in the sildenafil pediatric trial (NDA 203109 Study A1481131) is 6. The mean change from baseline in PVR in BREATHE-3 is -389 (i.e., for bosentan). If bosentan's effect in PVR relative to the putative placebo (assuming the mean PVR change from baseline for placebo is exactly zero with no error in the pediatric trial) were -389 and the relationship between 6MWD and PVR in the pediatric trial is the same as in the adult trials, then bosentan would have a treatment effect of 58.7 meters on 6MWD with 95% prediction interval (11.0, 106.4). But this is an unrealistic argument because the point estimate always has an error.

Using bootstrap to generate placebo pediatric patients

In another approach, PVR values for placebo pediatric patients in bosentan trial (BREATHE-3) are simulated by bootstrapping from the 50 placebo pediatric patients in the sildenafil trial. There were 60 placebo patients in that pediatric study, but only 50 patients had change from baseline in

PVR values. In each run, 17 placebo patients were generated to be compared with the 17 bosentan pediatric patients. 10,000 placebo samples were generated using bootstrap.

A simple t-test showed that 4970 out of 10,000 times, bosentan showed a significant treatment effect in PVR when compared with placebo. This indicates that even if the true treatment effect of bosentan on PVR is -389, there is only about 50% chance to show that the bosentan's effect on PVR is significant if the sample size is 17 for each treatment arm.

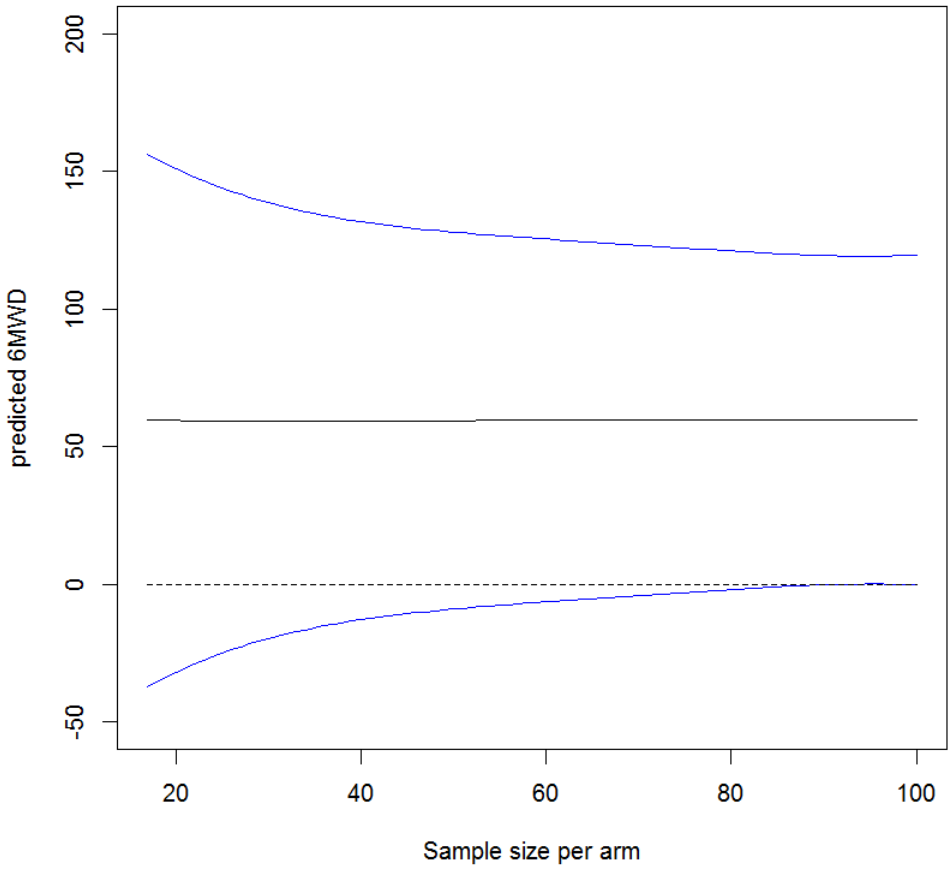
A linear regression model was used on the simulated datasets to estimate the treatment effect on PVR and the change from baseline in PVR in placebo arm. The estimates from each simulated dataset were used to predict the treatment effect on 6MWD. Rubin's formula was used to combine all 10,000 sets of results into an overall prediction. The overall predicted treatment effect on 6MWD is 59.6 meter with 95% prediction interval (-25.7, 144.9). The overall prediction of 6MWD does not change much by adding a simulated placebo arm (the mean change from baseline in PVR in sildenafil placebo patients is close to zero). However, by accounting for the uncertainty in placebo arm, the prediction interval is much wider and includes zero even though the mean change from baseline in PVR in the placebo arm remains the same. If bootstrap sampling were applied to the bosentan arm, the prediction interval would have been even wider.

Sample size and prediction interval

The small sample size certainly played a part in the wide prediction interval in addition to the uncertainty in the relationship between PVR and 6MWD.

One can try to increase the sample size of both bosentan and placebo arms by bootstrapping sample data for bosentan pediatric patients in BREATHE-3 and pediatric patients in placebo group from sildenafil trial. Figure 4 illustrates how the prediction interval behaves as the hypothetical sample size per arm increases in the pediatric trial under the critical assumptions given above.

Figure 3: Relationship between Sample Size Per Arm and Prediction Interval



3.2.2 AC-052-365 (FUTURE 1) and AC-052-367 (FUTURE 2)

3.2.2.1 Study Design and Endpoints

FUTURE 1 was a 12-week, multicenter, prospective, open-label, single-arm study designed to evaluate the PK and safety of the dispersible tablet formulation of bosentan in male and female patients aged ≥ 2 and < 12 years of age with WHO FC II or III idiopathic or familial PAH.

AC-052-367 (FUTURE 2) was an open-label, uncontrolled extension to FUTURE 1. Patients continued to receive the maintenance dose of bosentan they were receiving at the end of FUTURE 1.

Efficacy endpoints include:

Change from baseline to EOS or premature study drug discontinuation in:

– WHO FC
– QoL questionnaire score (10-item short form health survey for children)
– GCIS according to the parents/legal representatives
– GCIS according to the physician

PAH worsening endpoints:
– Worsening of PAH: time to first occurrence of death, transplantation, or hospitalization for PAH worsening.
– Worsening of PAH (broader definition): time to first occurrence of worsening of PAH (as defined above), or initiation of new therapy for PAH, or new right heart failure, or worsening of right heart failure

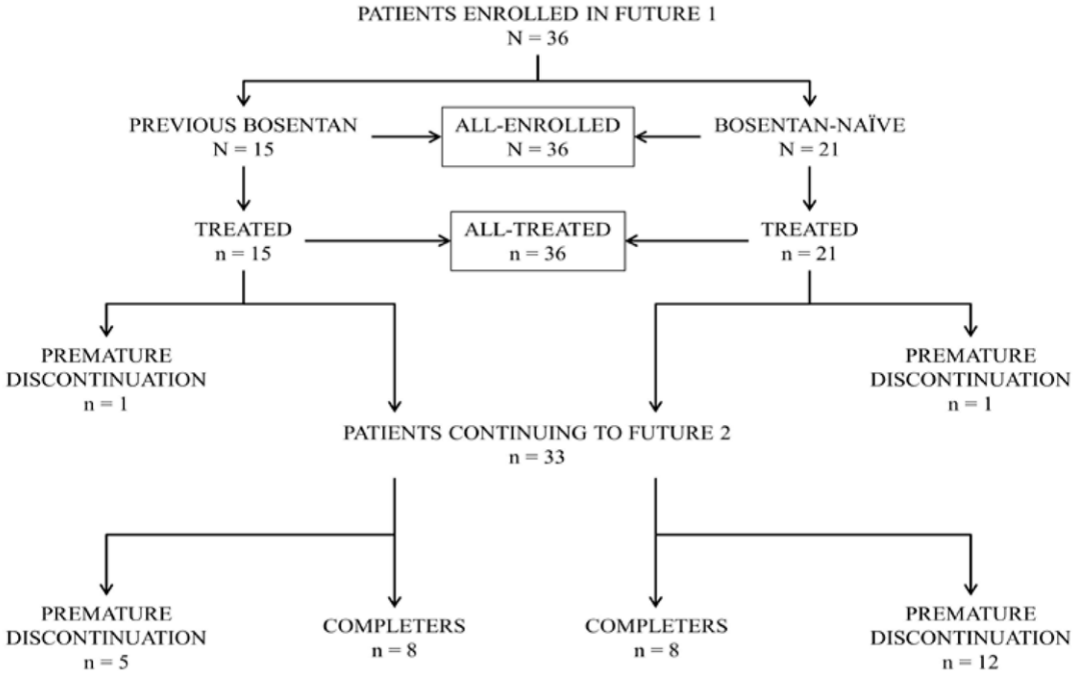
3.2.2.2 Statistical Methodologies

No formal hypothesis testing was defined. The efficacy variables were analyzed descriptively.

3.2.2.3 Patient Disposition, Demographic and Baseline Characteristics

36 patients were enrolled in the FUTURE 1 trial. The majority of patients were female (58.3%), and the median age was 7 years (range 2–11 years). About 90% of patients were Caucasian, and the majority (86.1%) had idiopathic PAH. The discontinuation rate in FUTURE 2 is high.

Figure 4: Patient Disposition in FUTURE 1 and FUTURE 2



[Source: Figure 3-1 in Sponsor's Summary of Clinical Efficacy]

Table 7: Patient Demographic and Baseline Characteristics in FUTURE 1 and FUTURE 2

	All patients N=36	Patients with previous Bosentan N=15	Patients Bosentan naive N=21
SEX [n (%)]			
n	36	15	21
Males	21 58.3%	10 66.7%	11 52.4%
Females	15 41.7%	5 33.3%	10 47.6%
AGE (years)			
n	36	15	21
Mean	6.8	6.9	6.6
Standard deviation	2.72	2.05	3.15
Median	7.0	7.0	7.0
Q1 , Q3	4.5 , 9.5	5.0 , 8.0	4.0 , 10.0
Min , Max	2.0 , 11.0	3.0 , 10.0	2.0 , 11.0
AGE [n (%)]			
n	36	15	21
2 - 3 years	4 11.1%	1 6.7%	3 14.3%
4 - 5 years	9 25.0%	3 20.0%	6 28.6%
6 - 11 years	23 63.9%	11 73.3%	12 57.1%
RACE [n (%)]			
n	36	15	21
Caucasian/white	32 88.9%	14 93.3%	18 85.7%
Black	1 2.8%	-	1 4.8%
Hispanic	2 5.6%	-	2 9.5%
Other	1 2.8%	1 6.7%	-
ETIOLOGY OF PAH [n (%)]			
n	36	15	21
Idiopathic	31 86.1%	12 80.0%	19 90.5%
Familial	5 13.9%	3 20.0%	2 9.5%

[Source: Sponsor's Summary of Clinical Efficacy Table 3-6, confirmed by the reviewer]

3.2.2.4 Results and Conclusions

Table 5 showed that 11 out of 28 patients in FUTURE 1 and FUTURE 2 had improved WHO functional class. 2 patients has worsening WHO functional class.

Table 8: WHO Functional Class Change from Baseline to EOS in FUTURE 1 and FUTURE 2

			End of Study/Premature Discontinuation of Study Drug visit				
	n	Baseline	n	I	II	III	IV
All patients	28	I	-	-	-	-	-
		II	17	6 21.4%	10 35.7%	1 3.6%	-
		III	11	2 7.1%	3 10.7%	5 17.9%	1 3.6%
		IV	-	-	-	-	

[Source: Sponsor's Summary of Clinical Efficacy Table 3-19, confirmed by the reviewer]

The KM event-free estimate for worsening of PAH (defined as time to first occurrence of death, lung transplantation or hospitalization for PAH worsening) was 78.9% with 95% CI (60.7%, 89.3%) at 2 years. For the broader definition of worsening of PAH (defined as time to first occurrence of death, lung transplantation, hospitalization for PAH worsening or initiation of new therapy for PAH, or new/worsening right heart failure), the KM event-free estimate was 56.2% with 95% CI (38.0%, 71.0%) at 2 years.

3.2.3 AC-052-373 (FUTURE 3)

3.2.3.1 Study Design and Endpoints

The FUTURE 3 study was a prospective, open-label, randomized study to investigate the PK of the dispersible tablet formulation of bosentan at doses of 2 mg/kg b.i.d. and 2 mg/kg t.i.d. in children with PAH from ≥ 3 months to < 2 years of age and from 2 years to < 12 years of age. This study was followed by FUTURE 3 extension study which enrolled patients who had completed the FUTURE 3 core study (AC-052-373). 58 out of 64 patients in FUTURE 3 core study entered FUTURE 3 extension study.

Patients were randomized in a 1:1 ratio to receive oral doses of bosentan as an aqueous dispersion at doses of 2 mg/kg b.i.d. or 2 mg/kg t.i.d. The treatment period was 24 weeks.

Efficacy endpoints include

- Change from baseline to 3, 6, 9, 12, 15, and 18 months of treatment over FUTURE 3 core and extension study in WHO FC
- PAH worsening endpoints:
    - Worsening of PAH: time to death, lung transplantation or hospitalization for PAH progression1 up to End-of-Treatment (EOT) + 7 day
    - Worsening of PAH (broader definition): time to death, lung transplantation, hospitalization for PAH progression, initiation of new therapy for PAH or new/worsening right heart failure.

- Overall survival: time to death (any cause) up to EOS
- Change from baseline to 3, 6, 9, 12, 15, and 18 months of treatment over the FUTURE 3 core and extension study in GCIS assessed by the physician and parents

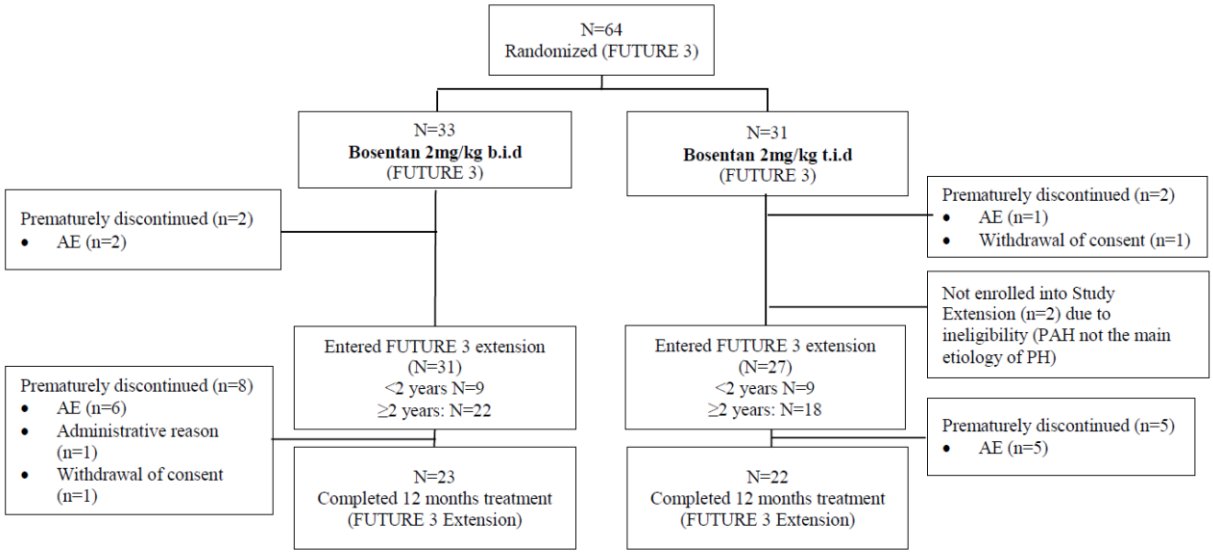
3.2.3.2 Statistical Methodologies

No formal hypothesis testing was defined. The efficacy variables were analyzed descriptively.

3.2.3.3 Patient Disposition, Demographic and Baseline Characteristics

64 patients were randomized in the FUTURE 3 core study. The majority patients were male (56.3%) and Caucasian (75.0%), with a median age of 3.8 years (range: 0.3–11.4 years). 42.2% and 28.1% of the overall patients were in WHO FC II and III, respectively.

Figure 5: Patient Disposition in FUTURE 3



[Source: Sponsor's Summary of Clinical Efficacy Figure 3-2]

Table 9: Patient Demographic and Baseline Characteristics in FUTURE 3

	b.i.d N=33	t.i.d N=31	Total N=64
GENDER [n (%)]			
n	33	31	64
Males	15 45.5%	21 67.7%	36 56.3%
Females	18 54.5%	10 32.3%	28 43.8%
AGE (years)			
n	33	31	64
Mean	4.5	5.2	4.8
Standard error	0.58	0.68	0.45
95% CL of mean	3.3 , 5.7	3.8 , 6.6	3.9 , 5.7
Median	3.7	4.8	3.8
Q1 , Q3	1.8 , 5.6	1.2 , 8.7	1.7 , 7.8
Min , Max	0.3 , 11.0	0.3 , 11.4	0.3 , 11.4
RACE [n (%)]			
n	33	31	64
Caucasian/white	25 75.8%	23 74.2%	48 75.0%
Black	1 3.0%	2 6.5%	3 4.7%
Asian	6 18.2%	4 12.9%	10 15.6%
Hispanic	-	1 3.2%	1 1.6%
Other	1 3.0%	1 3.2%	2 3.1%
Etiology of PAH			
n	33	30	63
Idiopathic (iPAH)	14 42.4%	15 50.0%	29 46.0%
Heritable (hPAH)	2 6.1%	-	2 3.2%
Associated (aPAH)	11 33.3%	13 43.3%	24 38.1%
PAH-CHD (open shunt)	6 18.2%	2 6.7%	8 12.7%
WHO functional Class [n (%)]			
n	33	31	64
I	9 27.3%	10 32.3%	19 29.7%
II	12 36.4%	15 48.4%	27 42.2%
III	12 36.4%	6 19.4%	18 28.1%

[Source: Sponsor's Summary of Clinical Efficacy Table 3-7, confirmed by the reviewer]

3.2.3.4 Results and Conclusions

Table 10 shows how WHO FC change from baseline at Month 6, Month 12 and Month 18 in FUTURE 3 and its extension study.

Table 10: Change in WHO Functional Class in FUTURE 3 and FUTURE 4

	Total N	Improved	Stable	Worsened
Month 6	62	10	51	1
Month 12	51	10	38	3
Month 18	37	4	33	0

Note: All WHO FC status were compared with baseline WHO FC

The KM estimates for event-free (PAH worsening) survival were 84.7% with 95% CI (72.6%, 91.7%) at Month 12 and 77.4% with 95% CI (64.2%, 86.2%) at Month 18.

3.3 Safety

Please refer to the medical review for safety evaluation.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Subgroup analyses were not performed due to the small sample size in the pediatric trials.

4.2 Other Special/Subgroup Populations

No other subgroups were analyzed.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The main focus in this review is to bridge the adult trials and the non-controlled BREATHE-3 pediatric trial for predicting a potential treatment effect of bosentan on six-minute walk distance (6MWD) in pediatric pulmonary arterial hypertension (PAH) patients aged between 3–15 years. The key question is whether bosentan would show a significant treatment effect on improving 6MWD if there were a placebo arm in the pediatric trial and 6MWD were measured.

There are two parts in this question. First, how should we model the relationship between 6MWD and pulmonary vascular resistance (PVR) in the bosentan adult trials? Secondly, what should we do to explore the treatment effect of bosentan on PVR since BREATHE-3 did not have a placebo arm?

To answer the first question, a bivariate fixed-effects model was applied to 3 bosentan adult trials (AC-052-351, AC-052-364 and AC-052-405) to estimate the overall relationship between PVR and 6MWD.

To answer the second question, bootstrapping was used to generate PVR data for a putative placebo arm using the PVR data of the placebo arm in a sildenafil pediatric trial.

All the calculations and simulations are based on at least two critically important assumptions:

1. The relationship between PVR and 6MWD in the pediatric trial is the same as in the adult trials.
2. If there is a placebo arm in BREATHE-3, the distribution of PVR change from baseline in that placebo arm would be the same as in other pediatric trials (such as in the sildenafil pediatric trial)

5.2 Collective Evidence

The individual-level and trial-level surrogacy can be examined in the bivariate fixed-effects model. In this case, the individual-level R^2 is 0.37 and the trial-level R^2 is 0.65, which indicates that PVR may not be considered as a reliable surrogate for 6MWD in the bosentan adult trials.

Using bootstrap sampling to generate a placebo arm for BREATHE-3 based on the placebo data in the sildenafil pediatric trial, the simulation showed that 1) there is only at most 50% chance to demonstrate a significant treatment effect by bosentan on PVR if the sample size is 17 for each treatment arm, and 2) the bosentan's effect on 6MWD relative to a putative placebo is predicted to be 59.6 meters with 95% prediction interval (-25.7, 144.9). If bootstrap sampling were applied for the bosentan arm, then the prediction interval would have been even wider.

The reviewer explored the relationship between sample size and prediction interval. Simulation showed that 100 patients per arm would translate to positive 6MWD prediction and prediction intervals.

5.3 Conclusions and Recommendations

In summary, we are not able to predict that bosentan would have demonstrated a statistically significant treatment effect on 6MWD, had BREATHE-3 had a placebo arm of similar sample size and had 6MWD been measured in these pediatric patients.

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

/s/

JIALU ZHANG
03/01/2017

HSIEN MING J HUNG
03/01/2017

(b)