CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

203496Orig1s000

STATISTICAL REVIEW(S)

JOINT CLINICAL/STATISTICAL REVIEW
(ADDENDUM)

NDA #:	203,496
Applicant:	United Therapeutics
Name of Drug:	Treprostinil tablets
Indication:	Pulmonary arterial hypertension
Date of submission:	October 23, 2011
Statistical Reviewer:	John Lawrence, Ph.D. (HFD–710)
Medical Reviewer:	Maryann Gordon, M.D. (HFD–110)

The purpose of this addendum is to provide some additional analyses of the pivotal Phase 3 efficacy study TDE-PH-302. As noted on p. 24 of the original review, at the end of study (Week 12), the estimated placebo-subtracted change in 6-minute walking distance was 25.5 m (p=0.0001) using the sponsor's single imputation method and their adjudication of reasons for dropout. However, 59 subjects (25%) in the UT-15C group did not have the week 12 walk test compared to 18 subjects (11%) in the placebo group. When the 59 UT-15C subjects are given worst rank, the p-value becomes 0.92. When the missing placebo subjects are assigned worst rank as well, the p-value becomes 0.21. It is of great concern that UT-15C had more than double dropout rate than placebo.

14 (6%) subjects in the treprostinil group and 9 (8%) subjects in the placebo group died during the course of the study before the walking distance could be measured at Week 12 (10 and 6 deaths respectively were listed as the reason for discontinuation but others died after discontinuation during 12 week period). Because of the 2:1 randomization, the percentage of deaths was approximately balanced between the two groups. There were an additional 45 (19%) subjects with missing data at week 12 in the treprostinil group and an additional 9 (8%) in the placebo group. Only 3 subjects total had missing data for the reason "In study, too ill to walk". 100% of the subjects that did not die and were not too ill to walk should have had a week 12 followup visit where the walking distance should have been measured and this value should have been used in the ITT analysis regardless of whether the subject took their randomized treatment.

The primary analysis was complicated, but essentially all subjects were assigned a score between 0 and 1 based on their change from baseline walking distance. Higher scores indicate better change in walking distance. The average imputed score for the 59 treprostinil subjects with missing Week 12 data is 0.36 while the average score for the 18 placebo subjects is 0.11. From this, it is seen that there was a large amount of missing data and the way it was handled seemed to substantially favor showing a treatment effect (i.e., 0.36 vs. 0.11). Since it is preferable to make decisions based on observed data rather than made up data and there was a substantial amount of missing data in this study with twice as much missing in the treatment group, it seems worthwhile to consider other ways of handling the missing data that do not favor showing a treatment effect.

In this addendum, three additional analyses are considered:

1. Analysis of change from baseline to Week 4 (earlier time point than used in the primary analysis) using other imputation methods; giving worst rank to all subjects (treprostinil and placebo) with missing data or only to all subjects in the trepostinil group.

In the sponsor's ITT analysis of change from baseline to Week 4, the estimated placebo subtracted change from baseline is 14 m with a p-value of 0.0025. 25 subjects in the treprostinil group and 8 in the placebo group had missing value sat Week 4. In the sponsor's analysis, the imputed scores for these 25 subjects with missing values had a mean of 0.34 while the mean score for the 8 placebo subjects was 0.08. Again, the sponsor's imputation seemed to substantially favor the treprostinil group.

When these 25 subjects in the treprostinil group are all given the worst score, the point estimate of the placebo subtracted change from baseline is 10 m and the p-value is 0.063. If, in addition, the 8 placebo subjects are also given the worst score, then the p-value remains 0.063 (note: 7 of them already had the worst rank, so it only changes the rank for 1 placebo subject).

2. Analysis of change from baseline to Week 12 giving fewer subjects from treprostinil group the worst rank (i.e. not all 59 are given worst rank).

When all 59 subjects in the treprostinil group are given the worst score, the p-value is 0.92. If only the top 23 of these 59 are given the worst score and the remaining 36 scores are left "as is", the average score for all 59 treprostinil subjects with missing data is 0.13 compared to an average score of 0.11 for the placebo subjects with missing data. The p-value for the analysis with this imputation is 0.051. 23 is the smallest number of subjects in the treprostinil group with missing data who were not already given the worst rank that would have to be given the worst rank to make the p-value above 0.05.

The reasons given for missing data at Week 12 for these 23 subjects were: Adverse event (14), Consent withdrawn (2), Discontinued for other reasons (1), In study, unblinded or other (3), Lost to follow-up (3).

3. Multiple imputation method for missing values at Week 12.

Another approach is to use multiple imputation to impute the missing data. For each imputed dataset, the worst rank was given to subjects who died as was done in the original single imputation analysis. For subjects with missing value who did not die, a random score was chosen uniformly between 0 and 0.25; this is based on the concept that anyone with missing data would have fallen in the lowest quartile had they been coerced into actually doing the walk test at Week 12. Note that there are 45 subjects in the treatment group who have missing data at Week 12 for reason other than death and 9 in the placebo group. The average score in the sponsor's analysis for the 45 subjects was 0.45 and the average score for the 9 placebo subjects was 0.16. From this imputed

dataset, the stratified treatment effect was estimated and then combined using a multiple imputation formula (Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.). The multiple imputation analysis will have two important differences from the original single imputation. First, the imputed scores will have the same average value in both groups (namely, 0.125). Second, the uncertainty in imputing a value will be reflected in the variance. This uncertainty is not factored in when a single imputation is used and treated like a fixed value. The result of the multiple imputation analysis done this way is a p-value of 0.056. However, the validity of multiple imputation analysis relies on model assumptions including the assumption that the values are missing at random. This assumption implies that the large imbalance in the rate of missing data between the two treatment groups does not matter in statistical analysis, which can be very problematic in my view.

There are other ways of imputing the missing values as part of the multiple imputation analysis. In general, those imputation schemes that tend to give better values to subjects with missing data will tend to favor showing a treatment effect because the number of subjects affected by these better imputed values is higher in the treprostinil group compared to the placebo group and vice versa. For example, if the missing data from the patients who did not die or too ill to walk are imputed by a random score smaller than 0.125 on average, then the treatment difference will likely not be statistically significant.

In summary, the robustness of the efficacy results depends heavily on how the missing data are treated in statistical analysis; the p-value can range from 0.0001 (from the sponsor's analysis) to 0.92 (from analysis giving all treatment subjects with missing data the worst score). So in my opinion, the efficacy of the treprostrinil tablets has not been convincingly demonstrated, based on this study.

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
10/10/2012

HSIEN MING J HUNG
10/10/2012

MARYANN GORDON
10/10/2012

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.