CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

209471Orig1s000

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 209-471
Supplement #:	
Drug Name:	Maxigesic® 325 (COMBOGESIC®); Acetaminophen 325 mg + Ibuprofen 97.5 mg
Indication(s):	(b) (4)
Applicant:	AFT Pharmaceuticals
Date(s):	Date received: May 7, 2020 Primary/Secondary reviews due: October 7/October 14, 2020 PDUFA: November 6, 2020
Review Priority:	Standard
Biometrics Division:	Division I
Statistical Reviewer:	John Lawrence, PhD
Concurring Reviewers:	Jinglin Zhong, PhD Jim Hung, PhD
Medical Division:	Division of Anesthesiology, Addiction Medicine, and Pain Medicine
Clinical Team:	Christina Fang
Project Manager:	Sandy Truong

Keywords:

Combination drug; factorial design

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

A single study was submitted to support the approval. The study had 4 arms (placebo, two monotherapy arms, and combination arm). The primary efficacy objective of this trial was to compare the time-adjusted Sum of Pain Intensity Differences (SPID) from baseline derived from the VAS pain intensity scores recorded over the 48 hours double blind period among the four treatment groups. The combination of acetaminophen 325 mg + ibuprofen 97.5 mg was demonstrated to be effective at reducing pain intensity in comparison to the other 3 treatment arms. All three pairwise comparisons of the combination compared to the other three arms were statistically significant ($p<0.001$). In addition, the combination reduced the use of rescue therapy, an important secondary endpoint.

The safety profile of Maxigesic® 325 in this study is consistent with what has been observed in the previous studies of this combination. Gastrointestinal disorders (50.3%) were the most frequently reported AEs followed by nervous system disorders (21.4%). The risk of experiencing gastrointestinal disorders was comparable among the four study groups and no significant differences were observed. There were no serious adverse events associated with study drug in the study.

From the statistical team's point of view, efficacy and safety of the combination for this indication was established.

2 INTRODUCTION

2.1 Overview

A single study was conducted for this indication. The study randomized 408 subjects in 3:3:3:2 ratio with 110 patients in the combination arm, 112 ibuprofen, 111 acetaminophen 75 placebo. Inclusion criteria were:
males and females aged at least 18 years and not more than 60 years old on the day of consent;
undergoing dental surgery for the extraction of at least two impacted third molar teeth;
a resting VAS pain intensity score at baseline (within 6 hours after the completion of surgery) of ≥40mm on a 100mm VAS scale with 0 = no pain and 100 = worst pain imaginable.
the key details of the study are shown in Table 1.

Table 1. List of all studies included in analysis

Study Number	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
AFT-MX-6	3	48 hours	48 hours	110 Maxigesic 112 ibuprofen 111 acetaminophen/ 75 placebo	pain following dental surgery

2.2 Data Sources

Electronic analysis datasets and study reports:

\\CDSESUB1\evsprod\NDA209471\0017\m5\datasets\aft-mx-6\analysis\adam\datasets

\\CDSESUB1\evsprod\NDA209471\0017\m5\53-clin-stud-rep\535-rep-effic-safety-stud\pain\5351-stud-rep-contr\aft-mx-6

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Studies of this type are complicated by missing data and rescue therapy. To some extent, this is unavoidable. In particular, denying rescue medication would raise ethical concerns. The number of subjects with missing data not due to rescue therapy was relatively low. A mixture of different methods depending on the reason for missing was used to handle missing data. Pre-rescue VAS scores were carried forward up to 6 hours to ensure that the outcome was not

influenced by the analgesic effect of the rescue medication (oxycodone 5-10 mg every 4-6 hours as require). A time interval of 6 hours was considered sufficient for the added analgesic effect of rescue medication to have worn off. Intermittent missing VAS scores and/or times were imputed by linear interpolation

For the 10 subjects that stopped recording VAS pain intensity scores within the first 12 hours of the first dose of study medication (excluded from EPP), their time-adjusted SPID48 were estimated using multiple imputation. The multiple imputation approach incorporated a multiple regression estimation procedure that used pre-discontinuation measures for the outcome, age, gender, stratum and number of molars extracted.

It was planned in both the protocol and Statistical Analysis Plan to test the model assumptions for the mixed effects model used in the primary analysis. If the normality assumption was found to be violated, then a log transformation or a nonparametric analysis would be used. However, I could not find any such assessments of the model assumptions and I could not find any discussion about log transformation or nonparametric analyses results shown in the Study Report. For this type of mixed effects model, the type I error may still be controlled even if the normality assumption is violated if the sample size is large. But, there remain good reasons for doing the assessment and the alternative analyses beyond the type I error rate argument. The assessment of normality planned in the protocol and SAP was a good idea and it should have been shown in the Study Report regardless of how it turned out. The failure to do that, particularly when it was stated in both the protocol and SAP, detracts from the analysis quality.

No adjustment was made for multiple comparisons for the secondary endpoints. This was unfortunate because that makes it difficult to interpret the secondary endpoint analyses. In a study of a combination drug, there are complexities because the hypotheses are of the intersection-union type. In other words, the alternative hypothesis is the combination is better than monotherapy A and the combination is better than monotherapy B. That's the intersection part. The multiple comparisons among different endpoints in an ordinary two arm trial (test vs. control) do not have this added complexity. In summary, although the adjustment for multiple endpoints is more difficult in this setting, it is still important for the interpretation of the results. The FDA guidance document on multiple endpoints requires the familywise error rate control for all primary and secondary endpoints.

Despite the limitations described above, I would still judge the quality of the data and the analysis as very good. Some sensitivity analyses of the primary endpoint were done to assess the impact of different assumptions on the results. All showed consistent results with the primary analysis. The data and analysis quality could also have been better if the missing data not due to rescue was held to 0 patients and/or if a more conservative way of handling the missing data was used such as worst case or tipping point analyses.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

408 subjects with dental pain of 40 mm or higher on a 100 mm VAS were randomized in 3:3:3:2 ratio to combination, monotherapy, or placebo arms. Randomization occurred at the research facility on the day of surgery, once participant eligibility for the study was confirmed postoperatively. The randomization was stratified based on the type of anesthesia used for the dental surgery (general anesthetic or local anesthetic) and study site. Double blinding was achieved by the use of matching tablets packaged into identical blisters and cardboard outer containers.

On the day of surgery, the participant was randomized postoperatively after meeting the last inclusion criterion of scoring at least 40 mm on the pain intensity VAS measured at rest. This VAS pain score was considered the baseline pain score (measured within 6 hours following completion of the surgery and prior to the first dose of study medication). The first dose of study medication was administered once the participant was randomized under the supervision of the study staff.

The Double-blind Treatment Period was from the start of study drug administration up to 48 hours after the first dose. During the study period, participants were supplied with the double-blinded study medication to be taken approximately every 6 hours. To facilitate dosing compliance participants received a text message reminding them to take their study drug. The text message was sent approximately 15 minutes before their dose was due. Participants were required to stay at the research facility (hospital) or at the study site for up to 6 hours after surgery. During their hospital stay, participants were supplied with two stopwatches to record the time to onset of pain relief (perceptible and meaningful pain relief).

VAS pain intensity scores were assessed by participants and recorded in their patient diary at scheduled time points listed below:
• During their hospital stay, VAS pain scores were assessed at 15, 30, 45 minutes and 1, 1.5, 2, 3, 4, 5, 6 and 7 hours after the first dose of study medication.
• After discharge from the hospital, VAS pain intensity scores were assessed immediately before taking each dose and 2 hours after taking each dose while the participant was awake.
• If participants needed extra pain relief, an extra VAS pain intensity score was assessed immediately before taking the rescue medication.

These VAS pain intensity measurements were recorded in a patient diary and were collected at the end of 48 hours study period. Instructions on how to complete the patient diary and how to use the stopwatches were given to the participant before the assessment was undertaken.

The sample size was calculated to have 90% power to detect a difference in mean time-adjusted SPIDs of 10.5mm. The pooled standard deviation was assumed to be 19.5mm. Using a two-tailed α=0.05, a sample size of 96 evaluable participants per active group in the ITT population are required for 90% power. A sample size of 64 evaluable participants in the placebo group will

provide 90% power to detect differences of 12.0mm or larger as statistically significant (two-tailed α=0.05), between placebo and any of the active treatment groups.

3.2.2 Statistical Methodologies

The time-adjusted SPID is defined as {the sum of the observed pain scores multiplied by the time interval length} divided by the total time. Missing data were imputed as described above in section 3.2.1.

A general linear model including randomized treatment and stratum as factors with planned pair-wise comparisons between Maxigesic 325 and each of the three other study treatments will be used to compare the efficacy of the four treatments. If the data do not meet assumptions for parametric analyses as determined by for example Lilliefors test, then in the first instance the data will be log transformed and if this does not achieve adequate normality then the groups will be compared using the Kruskal-Wallis non-parametric test.

The secondary efficacy endpoints as other indicators of pain relief are:
-The time to onset of pain relief after the first dose of study drug defined as
(i) perceptible and (ii) meaningful pain relief using the two-stopwatch method.
-The maximum VAS pain scores up to 48 hours after the first dose of study medication.
-The response rates (response rate to be defined as the percentage of participants who reduce their pain intensity scores by at least 50% compared with the baseline VAS measure).
-The time to peak reduction in VAS pain intensity scores following the first dose of study medication.
-The time to requirement for rescue medication.
-The percentage of participants who use rescue medication.
-The amount of rescue medication used (defined as number of tablets).
-The categorical global pain rating.

The differences between Maxigesic 325 and each of the other treatments with regards to the time to onset of pain relief after the first dose of study drug defined as (i) perceptible and (ii) meaningful pain relief will be used as the key secondary endpoint. The statistical significance of this key secondary endpoint will only be assessed if the primary endpoint shows a statistically significant advantage to Maxigesic 325 over the other three treatments.

A general linear model including randomized treatment and stratum as factors with pair-wise comparisons between Maxigesic 325 and each of the three other study treatments, if the treatment is significant in the ANOVA model, will be used to compare the maximum VAS pain intensity scores. If the data do not meet assumptions for parametric analyses, as determined by for example Lilliefors test, then in the first instance the data will be log transformed and if this does not achieve adequate normality then the groups will be compared using the Kruskal-Wallis non-parametric test.

The percentage of participants responding and the percentage of participants requiring rescue medication will be compared among the four treatments using a Mantel-Haenszel Chi-square test, with stratum as the stratification factor. If the overall effect of treatment is significant, then pair-wise comparisons (Maxigesic 325 vs each of the other treatments) will be undertaken using the Mantel-Haenszel Chi-square test. The amount of rescue medication and the categorical global scores will be compared amongst study groups using a Kruskal-Wallis non-parametric ANOVA. If the Kruskal-Wallis test is significant then pair-wise comparisons (Maxigesic 325 vs each of the other treatments) will be undertaken using Mann-Whitney U tests.

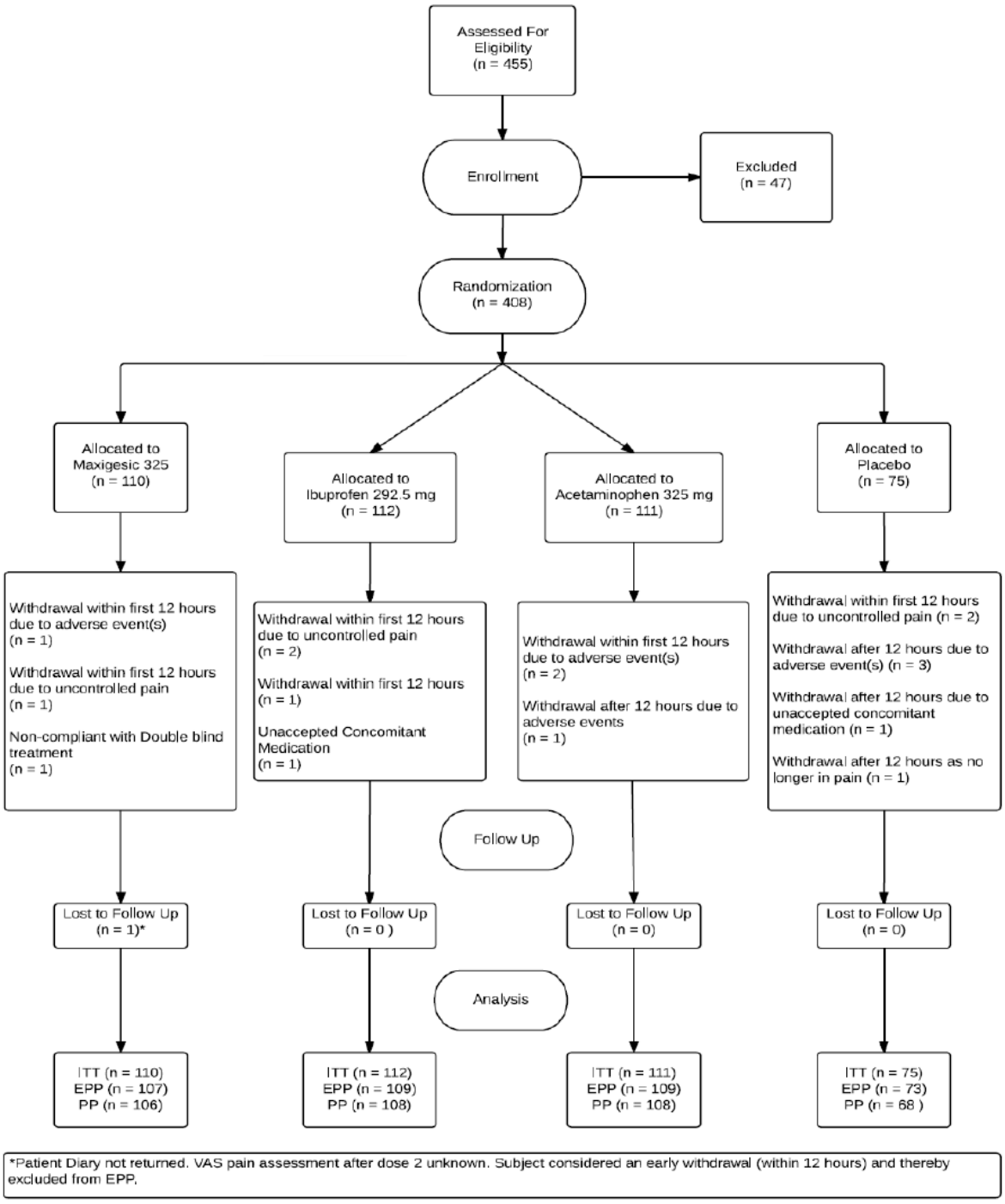
The time to peak response, the time to use of rescue medication and the times to perceptible and meaningful pain relief will be summarized using Kaplan-Meier curves and compared between groups using stratified log-rank tests. If these comparisons are significant then comparisons between pairs of treatments (Maxigesic 325 vs each of the other treatments) will be undertaken using log-rank tests. For these time-to-event analyses any participants who do not experience the event, (e.g. do not take any rescue medication or have perceptible and meaningful pain relief) will be censored at the end of the relevant observation period.

The Statistical Analysis Plan states "A two-tailed p-value <0.05 will be taken to indicate statistical significance for all the secondary outcomes. There will be no correction for multiple comparisons for the analyses of the secondary endpoints." However, the first secondary endpoint was designated as the key secondary endpoint.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

The patient disposition is summarized in Figure 1.

Figure 1. Patient disposition.



Source: Figure 4 of Study Report

The baseline demographic information is shown in Table 2. There were no important differences between groups.

Table 2. Baseline demographic variables.

	Maxigesic® 325 N=110	Ibuprofen N=112	Acetaminophen N=111	Placebo N=75	Total N=408
Sex					
Male	41 (37.3%)	30 (26.8%)	35 (31.5%)	27 (36.0%)	133 (32.6%)
Female	69 (62.7%)	82 (73.2%)	76 (68.5%)	48 (64.0%)	275 (67.40%)
Age (years)					
Mean (SD)	25.4 (7.0)	24.3 (5.9)	25.0 (6.9)	24.3 (6.3)	24.8 6.6
Weight (kg)					
Mean (SD)	71.3 (14.1)	71.5 (14.9)	71.2 (13.6)	73.1 (14.0)	71.6 (14.1)
Race					
White	93 (84.5%)	88 (78.6%)	95 (85.6%)	57 (76%)	333 (81.6%)
Black or African American	2 (1.8%)	3 (2.7%)	0 (0.0%)	5 (6.7%)	10 (2.5%)
Asian	1 (0.9%)	2 (1.8%)	2 (1.8%)	1 (1.3%)	6 (1.5%)
Native Hawaiian or Other Pacific Islander	1 (0.9%)	1 (0.9%)	0 (0.0%)	1 (1.3%)	3 (0.7%)
Other	10 (9.1%)	14 (12.5%)	11 (9.9%)	8 (10.7%)	43 (10.5%)
Mixed	3 (2.7%)	4 (3.6%)	3 (2.7%)	3 (4.0%)	13 (3.2%)

Baseline VAS Pain

Mean (SD)	56.2 (13.9)	56.1 (14.3)	55.4 (13.2)	58.4 (15.3)	56.4 (14.0)
Median	51.5	51.5	51.0	54.0	52.0
Min ; Max	40 ; 98	40 ; 94	40 ; 94	40 ; 100	40 ; 100

Source: Table 14 and 22 of Study Report.

3.2.4 Results and Conclusions

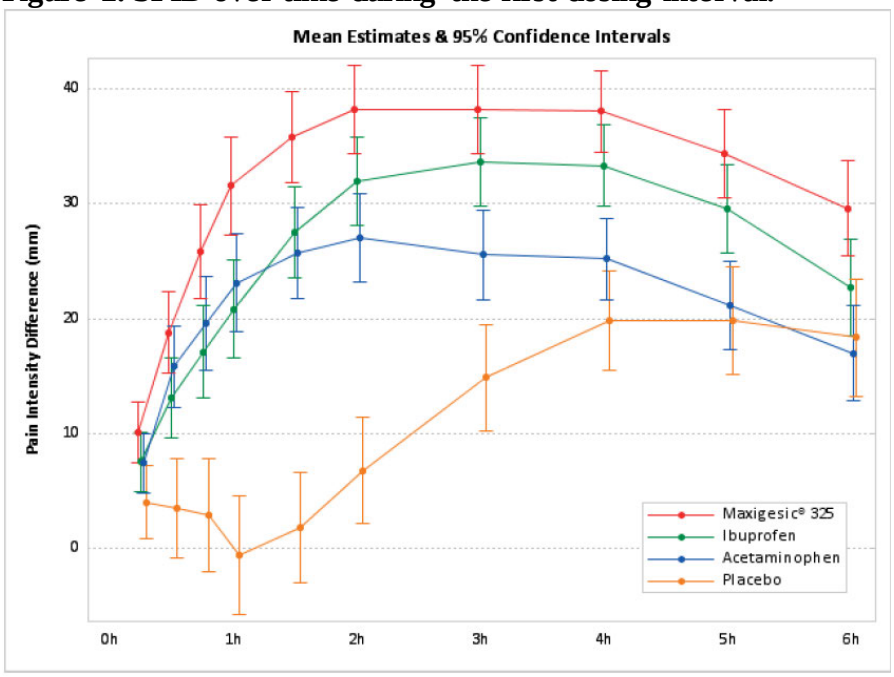
The results for the primary endpoint are shown in Table 3. There was a statistically significant difference between the combination arm and all three other arms ($p<0.001$). The mean SPID for each treatment arm over time is shown in Figure 2 (during first 6-hour dosing interval) and Figure 3 (during entire 48 hours of follow up). The mean SPID curves show the effect of the combination persists throughout the entire 48 hours of follow up. However, note that the last measurement of VAS pain was 2 hours post-final dose. The final dose was at 42 hours and the final VAS pain measurements was at 44 hours. In other words, there were no measurements at 48 hours and therefore, it may be more accurate to say that the duration of follow-up is 44 hours.

Table 3. Summary of time-adjusted SPID by study arm (primary endpoint, ITT population).

	Maxigesic® 325 N=110	Ibuprofen N=112	Acetaminophen N=111	Placebo N=75
Initial Data				
N	107	109	109	73
Mean (SD)	35.88 (21.57)	28.23 (22.71)	22.63 (24.10)	19.72 (21.76)
Median	35.58	28.27	23.79	19.50
Min ; Max	-14.25 ; 78.49	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Statistics by Imputation				
Imputation #1				
N	110	112	111	75
Mean (SD)	36.82 (22.36)	28.16 (22.41)	22.33 (24.21)	20.02 (21.54)
Median	36.66	27.80	23.79	19.65
Min ; Max	-14.25 ; 102.97	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Imputation #2				
N	110	112	111	75
Mean (SD)	36.40 (21.57)	27.89 (22.82)	22.17 (24.12)	19.89 (21.49)
Median	36.66	27.80	23.77	19.65
Min ; Max	-14.25 ; 78.49	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Imputation #3				
N	110	112	111	75
Mean (SD)	36.58 (22.09)	28.13 (22.57)	22.56 (23.89)	19.85 (21.55)
Median	36.66	27.80	23.77	19.50
Min ; Max	-14.25 ; 97.12	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Imputation #4				
N	110	112	111	75
Mean (SD)	36.34 (21.90)	28.37 (22.44)	22.66 (23.89)	19.56 (21.49)
Median	35.61	28.36	23.79	19.47
Min ; Max	-14.25 ; 90.17	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Imputation #5				
N	110	112	111	75
Mean (SD)	36.94 (22.33)	28.30 (22.41)	22.44 (23.95)	19.97 (21.53)
Median	36.66	28.36	23.77	19.65
Min ; Max	-14.25 ; 87.70	-21.67 ; 88.41	-39.16 ; 82.60	-31.97 ; 70.68
Analysis after Imputations				
Mean Estimate (SE)	31.56 (1.94)	23.18 (1.89)	17.71 (1.89)	14.86 (2.26)
95% Confidence Interval	27.76 ; 35.37	19.47 ; 26.89	14.00 ; 21.43	10.43 ; 19.30
P value vs. Maxigesic® 325	-	<.001	<.001	<.001
P value vs. Ibuprofen		-	0.026	0.003
P value vs. Acetaminophen			-	0.302

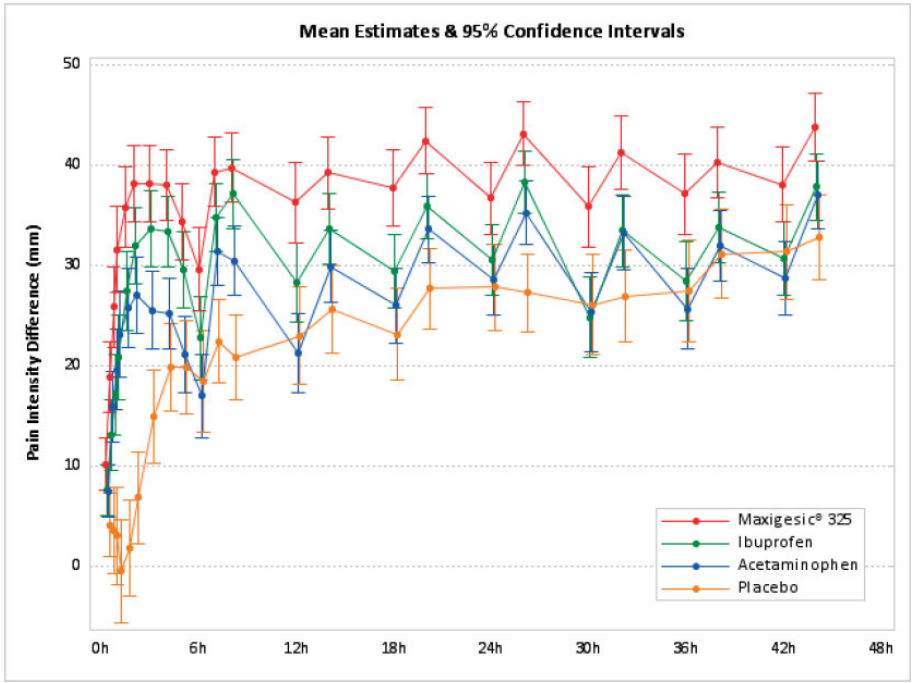
Source: Table 23 of Study Report and confirmed by FDA statistical reviewer.

Figure 2. SPID over time during the first dosing interval.



Source: Figure 14 of Study Report.

Figure 3. SPID over time during entire 48 hours.



Source: Figure 15 of Study Report.

The test for normality indicates that the residuals from the general linear model are not normal. The p-value for Lilliefors test is 0.036 and the p-value from the Shapiro-Wilk test is 0.030. The Kruskal-Wallis test has a p-value of less than 0.0001. This is a global test of whether there is

any difference among the four treatment arms. In order to compare the arms pairwise, the Wilcoxon-Mann-Whitney test should be used. The comparison of the combination arm to both acetaminophen and ibuprofen arms remain significant using this nonparametric test (p=0.00002 and p=0.011 respectively).

For the secondary endpoints, time to perceptible and meaningful pain relief was designated as the key secondary endpoint. The results are shown in Table 4. Without adjustment for multiple comparisons, the combination appears to be superior to all other arms for time to perceptible pain relief but not for time to meaningful pain relief. It is difficult to interpret these results since superiority was not demonstrated for both endpoints.

Table 4. Results for time to perceptible and meaningful pain relief (Key secondary analysis, ITT population).

	Maxigesic® 325 N=110	Ibuprofen N=112	Acetaminophen N=111	Placebo N=75
Time to Perceptible Pain Relief (min)				
N	110	112	111	75
Mean (SE)	27.06 (2.17)	37.13 (3.24)	37.66 (3.56)	64.95 (6.30)
Median Estimate	19.24	23.78	19.77	34.22
95% Confidence Interval	16.02 ; 21.05	17.25 ; 29.47	15.73 ; 25.48	22.37 ; 119.00
No. of Subjects Achieved (%)	102 (92.7%)	97 (86.6%)	94 (84.7%)	48 (64.0%)
P value vs. Maxigesic® 325	-	0.078	0.130	<.001
Time to Meaningful Pain Relief (min)				
N	110	112	111	75
Mean (SE)	64.04 (5.28)	78.18 (5.23)	90.92 (7.65)	173.61 (9.44)
Median Estimate	42.98	61.86	48.62	-
95% Confidence Interval	34.20 ; 49.17	51.20 ; 74.85	38.97 ; 61.75	184.08 ; -
No. of Subjects Achieved (%)	94 (85.5%)	84 (75.0%)	83 (74.8%)	32 (42.7%)
P value vs. Maxigesic® 325	-	0.022	0.030	<.001

Source Table 24 of Study Report.

The remaining secondary endpoints cannot be interpreted because there was no plan to control the familywise error rate. The p-values are not adjusted for multiple comparisons and none of the results should be called statistically significant. But, they may provide supportive evidence for the effectiveness of the primary endpoint. The results from the Study Report for these analyses are copied below:

-Time to Meaningful Pain Relief:
Median time to meaningful pain relief was shortest for the Maxigesic® 325 group (42.98 min), followed by acetaminophen (48.62 min), and ibuprofen (61.86 min) groups. Median time to meaningful pain relief could not be estimated for the placebo group as less than 50% of subjects in this group did not achieve meaningful pain relief.

-The Maximum VAS Pain Intensity Score:
Mean maximum VAS pain intensity score was lowest in the Maxigesic® 325 group (Mean=55.06, SE=1.83), followed by the ibuprofen group (Mean=61.28, SE=1.81) and acetaminophen group (Mean=64.46, SE=1.81). The highest maximum VAS pain intensity score was observed in the placebo treatment group (Mean=72.05, SE=2.15).

-The Response Rate:
The percentage of participants that responded to treatment (defined as achieving a 50% reduction in baseline VAS pain prior to the consumption of rescue medication) was highest in the Maxigesic® 325 group (87%) followed by the ibuprofen (77%), acetaminophen (69%) and placebo (37%) groups.

-Time to Peak Reduction/Response:
The median time to response was shortest in the Maxigesic® 325 group (0.75 hours), followed by the acetaminophen (1.00 hours), ibuprofen (1.50 hours) and placebo (7.00 hours) groups.

-Time to the requirement of rescue medication:
Mean time to rescue was longest for Maxigesic® 325 (37.53 hours) followed by ibuprofen (28.52 hours), acetaminophen (22.98) hours and placebo (9.90 hours). As fewer than 50% of subjects in the Maxigesic® 325 and ibuprofen groups require rescued medication, the median time to rescue was not estimated for these groups and was 20.33 hours for the acetaminophen group and 1.75 hours for the placebo group.

-The percentage of participants using rescue medication:
The percentage of subjects who took at least one dose of rescue medication was lowest in the Maxigesic® 325 group (23.9 %), followed by the ibuprofen group (43.2 %), acetaminophen group (53.2%) and the highest in placebo group (81.3%).

-The amount of rescue medication consumed:
Any rescue medication consumed beyond 48 hours after the first dose of study medication was not included in the analysis of the amount of rescue medication consumed. The mean quantity (mg) of rescue medication (oxycodone) consumed was 3.7 mg (SD=9.11 mg) in the Maxigesic® 325 group, 7.1 mg (SD=11.57 mg) in the ibuprofen group, 11.0 mg (SD=14.92 mg) in the acetaminophen, and 17.9 mg (SD=18.29 mg) in the placebo group.

-The Categorical Global Pain Rating:
The pairwise comparisons of Maxigesic® 325 group and the other three treatment groups revealed that patients in this group achieved significantly greater pain relief according to the Mann-Whitney U test.

3.3 Evaluation of Safety

See clinical review.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

The results for the primary endpoint by gender, race and age are shown in Table 5. There were no important differences in the results within these subgroups compared to the overall analysis. The results by country are shown in Figure 4. There were only two geographical regions represented in the study: US and New Zealand. There appears to be a large difference in the

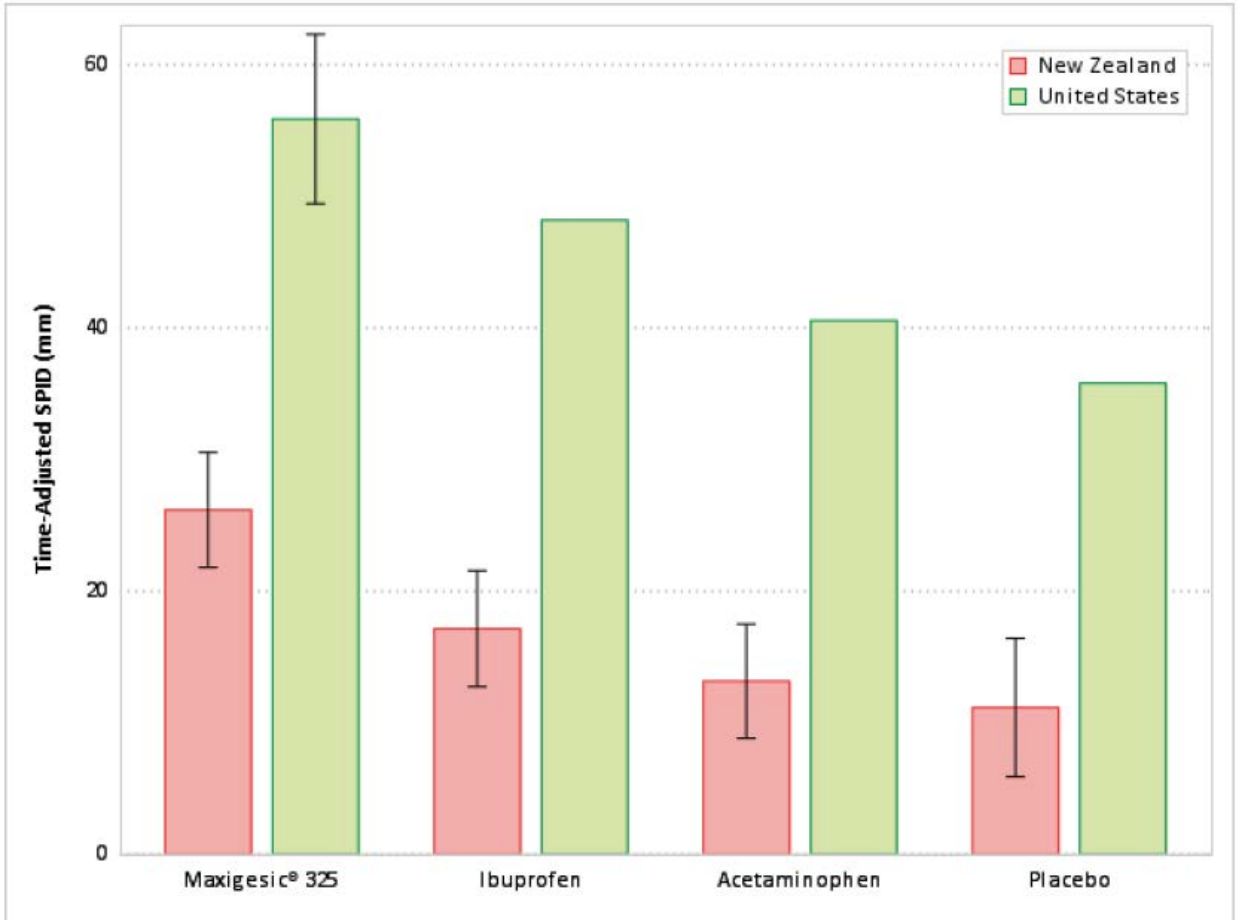
mean SPID between the two regions. However, the effects of each monotherapy and the combination appear consistent in both countries.

Table 5. Results for SPID by gender, age, race (primary endpoint, ITT population).

	Maxigesic® 325 N=110	Ibuprofen N=112	Acetaminophen N=111	Placebo N=75
By Sex				
Male				
N	41	30	35	27
Mean Estimate (SE)	32.47 (3.44)	25.52 (4.03)	24.82 (3.62)	16.39 (3.90)
95% Confidence Interval	25.66 ; 39.28	17.54 ; 33.49	17.65 ; 31.98	8.67 ; 24.10
Mean Difference Estimate (SE)	-	6.95 (4.56)	7.66 (4.33)	16.09 (4.71)
95% Confidence Interval	-	-2.06 ; 15.97	-0.92 ; 16.23	6.77 ; 25.40
P Value	-	0.129	0.080	<.001
Female				
N	69	82	76	48
Mean Estimate (SE)	29.56 (2.37)	21.01 (2.18)	13.66 (2.26)	12.87 (2.84)
95% Confidence Interval	24.89 ; 34.23	16.72 ; 25.30	9.21 ; 18.10	7.28 ; 18.47
Mean Difference Estimate (SE)	-	8.55 (3.00)	15.90 (3.05)	16.68 (3.45)
95% Confidence Interval	-	2.63 ; 14.46	9.89 ; 21.91	9.88 ; 23.48
P Value	-	0.005	<.001	<.001
By Age				
< 25 Years				
N	71	81	75	53
Mean Estimate (SE)	28.50 (2.41)	21.40 (2.28)	18.00 (2.37)	13.17 (2.77)
95% Confidence Interval	23.75 ; 33.24	16.92 ; 25.88	13.33 ; 22.68	7.72 ; 18.62
Mean Difference Estimate (SE)	-	7.10 (3.04)	10.50 (3.10)	15.33 (3.40)
95% Confidence Interval	-	1.11 ; 13.09	4.39 ; 16.60	8.63 ; 22.03
P Value	-	0.020	<.001	<.001
≥ 25 Years				
N	39	31	36	22
Mean Estimate (SE)	32.88 (3.33)	23.53 (3.56)	15.05 (3.25)	14.08 (4.23)
95% Confidence Interval	26.28 ; 39.48	16.47 ; 30.59	8.62 ; 21.47	5.70 ; 22.46
Mean Difference Estimate (SE)	-	9.35 (4.47)	17.83 (4.35)	18.80 (5.07)
95% Confidence Interval	-	0.50 ; 18.20	9.23 ; 26.44	8.76 ; 28.84
P Value	-	0.039	<.001	<.001
By Race[1]				
White				
N	93	88	95	57
Mean Estimate (SE)	30.35 (2.03)	19.91 (2.06)	15.94 (1.97)	12.16 (2.47)
95% Confidence Interval	26.36 ; 34.34	15.86 ; 23.95	12.07 ; 19.82	7.31 ; 17.02
Mean Difference Estimate (SE)	-	10.44 (2.67)	14.41 (2.62)	18.19 (3.03)
95% Confidence Interval	-	5.18 ; 15.70	9.25 ; 19.56	12.23 ; 24.15
P Value	-	<.001	<.001	<.001
Non-White				
N	17	24	16	18
Mean Estimate (SE)	36.04 (5.42)	32.93 (4.96)	27.14 (6.27)	22.19 (6.01)
95% Confidence Interval	25.21 ; 46.87	23.03 ; 42.84	14.62 ; 39.66	10.20 ; 34.19
Mean Difference Estimate (SE)	-	3.10 (6.41)	8.90 (7.20)	13.85 (6.98)
95% Confidence Interval	-	-9.70 ; 15.91	-5.47 ; 23.26	-0.08 ; 27.78
P Value	-	0.630	0.221	0.051

Source Table 45 of Study Report.

Figure 4. Results for SPID by region (primary endpoint, ITT population).



Source: Figure 28 of Study Report.

4.2 Other Special/Subgroup Populations

No other subgroups were analyzed.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The overall conclusions are that the combination was found to be effective at reducing VAS pain over the followup period for the patients studied. The sensitivity analyses performed by the sponsor and myself do not change any of those conclusions.

5.2 Collective Evidence

Not applicable.

5.3 Conclusions and Recommendations

Based on the collective evidence, from the statistical viewpoint the combination was demonstrated to be superior to each component on the primary endpoint SPID. The combination was not demonstrated to be superior on both key secondary endpoints, but the trend was in the right direction. Although there was no plan to control the familywise error rate for the secondary endpoints, all of the secondary endpoints provide supportive evidence to the claim for the primary endpoint.

5.4 Labeling Recommendations (as applicable)

P-values for the secondary endpoints should not be included in the label. Some summaries of the clinically relevant findings in the secondary endpoints can be included. Although the assessment of model assumptions for the primary analysis shows the residual errors are not normally distributed, it is acceptable to include the results of the mean differences for SPID in the label.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

JOHN P LAWRENCE
09/30/2020 09:42:31 AM

JINGLIN ZHONG
09/30/2020 09:59:45 AM

HSIEN MING J HUNG
09/30/2020 10:17:34 AM

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 209-471
Related IND #:	IND 107,435
Drug Name:	Combogesic Tablets (acetaminophen 325 mg / ibuprofen 97.5 mg)
Indication(s):	Short term management of mild to moderate acute pain
Applicant:	AFT Pharmaceuticals
Date(s):	Submitted: March 1, 2017 PDUFA date: January 1, 2018
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Yi Ren, Ph.D.
Concurring Reviewers:	David Petullo, M.S.
Medical Division:	Division of Anesthesia, Analgesia, and Addiction Products
Clinical Team:	Christina Fang, M.D., Medical Reviewer Joshua Lloyd, M.D., Team Lead
Project Manager:	Allison Meyer

Keywords: NDA review, clinical trials

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

AFT Pharmaceuticals Ltd submitted a new drug application for Combogesic tablet which contains 325 mg of acetaminophen and 97.5 mg of ibuprofen for the short term management of mild to moderate acute pain in adult patients. This review focused on a multicenter, randomized, double-blind, placebo-controlled phase 3 study (AFT-MX-6). Patients in this study reported moderate to severe pain after undergoing removal of at least two impacted third molars within 6 hours. This study randomized 408 patients to four treatment groups (FDC 325/97.5, APAP 325, IBU 97.5, and placebo) in a 3:3:3:2 ratio. The primary efficacy endpoint was time-adjusted summed pain intensity difference of the pain intensity scores up to 48 hours after the first dose of study medication (SPID48). The key secondary efficacy endpoints were time to onset of perceptible and meaningful pain relief, respectively.

Since the datasets submitted by the applicant in the original submission were not of sufficient quality to confirm their primary and secondary analyses, I was not able to confirm that the results from study AFT-MX-6 provided substantial evidence of efficacy. I was also not able to examine the impact of missing data on the primary and secondary analyses. When I requested the applicant submit all programs used to derive the primary and secondary analyses, they were unable to do so because all preprocessing and derivations of clinical endpoints were performed manually within Microsoft Excel. In light of this, the applicant will contract a third party to produce SAS programs to reproduce the primary and secondary analyses from the submitted ADaM domains. AFT will also contract a third party to produce SAS programs for the derivation of new ADaM datasets as well as primary and secondary endpoint analyses from the submitted SDTM domains. The applicant stated that revised datasets would be submitted by October or November. Since these datasets were submitted on November 22, 2017, I did not have sufficient time to review the data during this cycle.

2 INTRODUCTION

2.1 Overview

AFT Pharmaceuticals Ltd submitted a 505(b)(2) new drug application (NDA) for Combogesic tablet, which is a fixed dose combination (hereafter referred to as FDC 325/97.5) of 325 mg of acetaminophen (APAP 325) and 97.5 mg of ibuprofen (IBU 97.5) for the short term management of mild to moderate acute pain. It was developed for the United States (US) market as a lower strength formulation of AFT's proprietary Maxigesic® product (acetaminophen 500 mg/ibuprofen 150 mg tablets), which has been on the market in New Zealand, Australia, Singapore, and some European countries since 2009. In 2011 due to the safety concerns about liver damage, FDA retricted the maximum dose of acetaminophen per tablet from 500 mg to 325 mg while retaining the maximum recommended daily dose of 4,000 mg (see 76 FR 2691[1]). In December 2012, the dosing regimen in study protocol has been changed from "three tablets four times a day" to "three tablets every 6 hours for 48 hours".

All relevant communications with the FDA are summarized below:

In the Advice Letter in June 2010, FDA stated that one adequate and well-controlled, full factorial study in patients with acute pain will be required.

The protocol for this study (AFT-MX-6) was submitted as a special protocol assessment (SPA) on June 15 and September 23, 2011 and March 9, 2012. The proposed statistical analysis plan (SAP) was not adequate to assess the efficacy due to several reasons: 1) inappropriate analysis population for efficacy, defined as patients who have at least 3 pain scores in the first 12 hours; 2) no imputation for missing data; 3) sample size was not calculated based on a two-sided test using a significance level of 0.05; 4) the general linear model did not include treatment as a factor.

In the Type B Meeting in February 2012, FDA disagreed on the proposed baseline observation carried forward (BOCF) approach to handling missing data for patients with no post-baseline data. FDA recommended approaches that are consistent with the National Academy of Science report. Also, FDA advised the applicant must record and analyze all use of rescue medication.

In the Pre-NDA Meeting dated July 9, 2015, FDA agreed that the phase 3 Study AFT-MX-6 would be sufficient to support an NDA for an acute pain indication.

(b) (4)

A second placebo-controlled phase 3 study (AFT-MX-6E) was conducted using Maxigesic, a higher strength formulation (acetaminophen 500 mg/ibuprofen 150 mg tablets) approved in New Zealand and could be considered supportive. However, according to the applicant, Maxigesic

[1] https://www.gpo.gov/fdsys/granule/FR-2011-01-14/2011-709

only showed a significant difference over placebo with respect to the primary endpoint, SPID24. There was no difference from acetaminophen or ibuprofen. Additionally, acetaminophen and ibuprofen were not significantly different from placebo.

My review focuses on Study AFT-MX-6, a multicenter, randomized, double-blind, placebo-controlled phase 3 study. I briefly discuss study AFT-MX-6E in Section 3.2.2.

2.2 Data Sources

All documentation including the study protocol, SAP, clinical study report, and literature referenced, as well as the SDTM and ADaM datasets were submitted under the network path \\CDSESUB1\evsprod\NDA209471\0000. Datasets were submitted by the applicant to the CDER electronic data room in SAS transport format.

In response to the email correspondence dated March 16, 2017 regarding the incorrect naming of dataset files in their initial submission, the applicant resubmitted datasets and define files under the network path \\CDSESUB1\evsprod\NDA209471\0002.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Upon receipt of applicant's submission, I found several issues regarding the quality of the dataset submitted which prevented confirmation of the applicant's analyses. Three separate information requests (IR) were sent to the applicant on April 20, June 7, and July 12, 2017. Concerns noted with the datasets were inadequately documented to allow a through review of the data. Specifically, the SDTM dataset QS.xpt submitted on March 20, 2017 was inconsistent with respect to planned analysis time (QSTPT) and actual assessment time (QSDTM). The discrepancy in time was up to 744 hours. and pre-rescue pain assessment. The actual assessment time did not follow the pre-specified dosing intervals for study drug and rescue medication. The ADaM dataset ADSPID.xpt has several significant errors with respect to elapsed time interval from previous pain assessment (TIMINT), analysis value from pre-rescue pain assessment carried forward (AVAL). Details are discussed further in Section 3.1.1. In response to the second IR sent on June 7, 2017 requesting all programs used to derive and generate the analyses datasets, the applicant informed FDA that they were unable to provide these programs because all preprocessing and derivations of clinical endpoints were done manually within Microsoft Excel and this manual replacement strategy introduced errors. Therefore, neither the primary nor the secondary efficacy analyses can be confirmed.

3.1.1 Review Issues and Information Requests

The three information requests (IRs) sent to the applicant and their responses are summarized as follows:

- **First IR sent on April 20, 2017**
  - Requested subgroup analyses by age, gender, race, and geographic region (US and New Zealand) subgroups.
  - Requested clarification of primary and secondary analyses, as well as derivation of the primary endpoint time-adjusted summed pain intensity difference (SPID) of the pain intensity scores up to 48 hours after the first dose of study medication (SPID48).
  - Since no imputed values found in ADaM datasets, requested details of handling missing data in the primary and secondary analyses, specifically the linear interpolation and multiple imputation in the primary analysis (see Section 3.2.2 for details).
  - Requested clarification of a statement made for secondary analyses in SAP
  - Requested all programs used to generate the efficacy analyses.
- Applicant's response received on April 26, 2017
  - Repeated the multiple imputation methodology pre-specified in the SAP without giving further explanation.
  - Explained the derivation of primary endpoint, calculated as area under the pain curve divided by the duration of planned time interval.
  - Provided SPSS programs for subgroup analyses only. However, the variable names were different from those in the output and ADaM datasets submitted. Therefore, these programs cannot be used to replicate the results.
- **Second IR sent on June 7, 2017**
  - Requested all programs used to generate and derive all ADaM datasets and produce the primary and secondary analyses
  - Requested explanation on how pre-rescue pain scores and missing pain scores were handled in the primary analysis.
- Applicant's response received on June 13, 2017
  - All data preprocessing including endpoint derivation were done manually within Microsoft Excel (i.e. no programs).
  - The manual replacement strategy introduced errors, such as negative time intervals calculated, wrong pre-rescue data replaced, pre-rescue scores carried forward inconsistently.
  - No explanation on multiple imputation.
  - AFT will contract a third party to
    - produce SAS programs to reproduce the primary and secondary analyses from the submitted ADaM and SDTM domains
    - produce SAS programs for the derivation of new ADaM datasets

    All relevant documentation (datasets, programs, etc.) will be shared with the FDA as soon as possible.
- **Third IR sent on July 12, 2017**
  - Data discrepancies found in SDTM dataset QS.xpt between the planned analysis time (QSTPT) and the actual assessment time (QSDTC) (see descriptive statistics in Table 1 – Table 3 and Appendix for details).

- Requested clarification on how withdrawals in the first 12 hours were handled in the primary analysis.

Table 1. Summary statistics of SDTM data discrepancy between planned and actual analysis times

Number of Subjects	Number of Observations	Time difference (hour)				
		Min	Max	Mean	Median	StdDev
66 (16%)	177	1	744	20.6	2	110.3

Table 2. Data discrepancy by treatment group

Treatment Group	Subjects with Discrepancy	Total Number of Subjects
FSC 325/97.5	21 (19%)	110
APAP 325 mg	28 (25%)	111
IBU 97.5 mg	8 (7%)	112
Placebo	9 (12%)	75

Table 3. Data discrepancy by study site and anesthesia type

Study Site and Anesthesia Type	Subjects with Discrepancy	Total Number of Subjects
CSL (local)	18 (31%)	58
CTNZ (general)	7 (10%)	71
CTNZ (local)	4 (15%)	26
Premier (local)	28 (21%)	135
SCT (local)	9 (9%)	101

Source: Reviewer

- Applicant's response received on July 19, 2017
  - Any of these discrepancies which are obvious transcription errors will be identified and corrected in the additional ADaM datasets, and will be documented for full traceability.
  - Subject level time-adjusted SPID48 values were imputed using multiple imputation regression method for early dropouts. This was not appropriate as it did not take within-subject variation into account**.**
  - All derivation and analysis of the endpoints from the submitted SDTM domains using SAS programs will be available in October/November.

In addition to the issues listed above, there are some minor review issues including but not limited to:

- SPIDs were calculated only up to 44 hours after the first dose of study medication (i.e. the primary endpoint is SPID44 instead of SPID48).

- The calculation of time-adjusted SPIDs used the duration of planned time interval instead of interval for actual assessment time.
- The actual assessment time did not follow the pre-specified dosing interval for study drug (every 6 hours) and dosing interval for rescue medication (4-6 hours as required and wait up to 2 hours after study drug administration).

3.2 Evaluation of Efficacy

3.2.1 Study AFT-MX-6

3.2.1.1 Study Design and Endpoints

Study AFT-MX-6 is a multicenter, randomized, double-blind, placebo-controlled phase 3 study that was conducted in New Zealand (study sites including CTNZ, SCT, and CSL) and US (study site Premier) in patients who had undergone dental surgery for the extraction of at least two molar teeth. General and local anaesthetic techniques for the dental surgery were pre-specified in the study protocol. Randomization was stratified by study site and type of anesthesia for the dental surgery.

The study was designed to evaluate the time-adjusted SPID48 among the four treatment groups (FDC 325/97.5, APAP 325, IBU 97.5, and placebo). Patients were eligible for enrollment if they were undergoing removal of at least two impacted third molars with moderate to severe post-surgical pain defined as a resting VAS pain intensity score ≥ 40mm within 6 hours after the completion of surgery and aged between 18 and 60 years old. Eligible patients were randomized to treatment groups in a 3:3:3:2 ratio, as shown in Table 4.

Table 4. Study Design

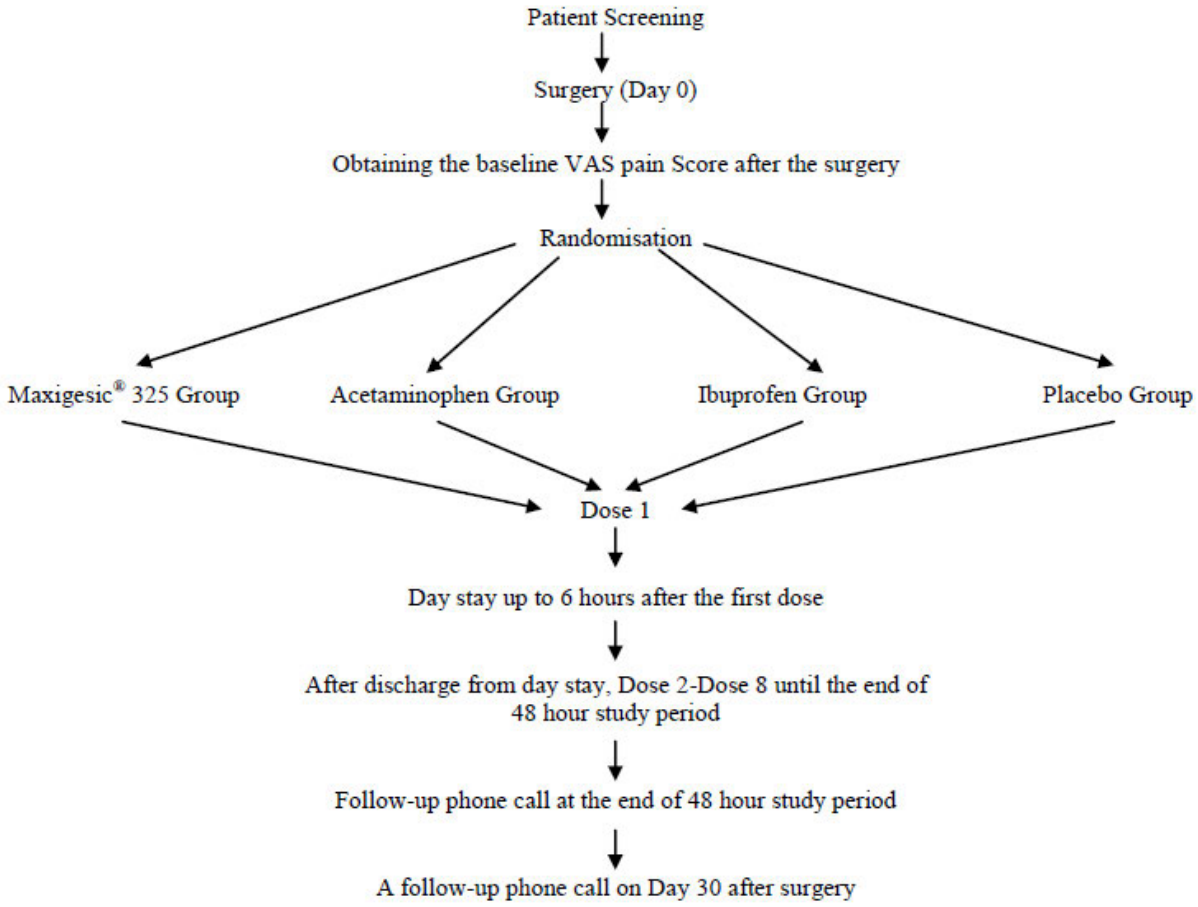
Study	Treatment	Sample Size		Study Site (Anesthesia Type)
AFT-MX-6	FDC 325/97.5	110	N=408	CTNZ (general/local)
	APAP 325	111		SCT (general/local)
	IBU 97.5	112		CSL (local)
	Placebo	75		Premier (local)

Source: Reviewer

As illustrated in Figure 1, this study consisted of patient screening, day of surgery, randomization, 48-hour double-blind treatment period, and a telephone follow-up at 30 days. The baseline pain score was measured within 6 hours following completion of the surgery and prior to the first dose of study drug. During the treatment period, patients were required to stay at the hospital or study site for up to 6 hours after surgery. During their stay, visual analogue scale (VAS) pain scores were assessed at 15, 30, 45 minutes and 1, 1.5, 2, 3, 4, 5, 6 and 7 hours after the first dose of study medication. The two stopwatch method was utilized record the time to onset of perceptible and meaningful pain relief. Three tablets of study drug were administered orally every 6 hours for 48 hours (up to 8 doses). After discharge from the hospital, pain scores

were assessed immediately before taking each dose of study drug and two hours after taking each dose while the patient was awake. If extra pain relief was needed, a pain score was assessed immediately before taking the rescue medication. Rescue medication was oxycodone 5-10 mg every 4-6 hours as needed. The primary efficacy endpoint was SPID48. The key secondary efficacy endpoints were time to onset of perceptible and meaningful pain relief, respectively.

Figure 1. Study Design of Study AFT-MX-6



Source: Applicant's Clinical Study Report (CSR) Figure 1

3.2.1.2 Statistical Methodologies

The intent-to-treat population (ITT) was defined as those patients who have been randomized and taken at least one dose of study medication. The SAP stated that the primary efficacy analyses will use the ITT population. The Per-protocol (PP) Population, defined as all dosed patients classified according to the actual treatment received who are compliant with the treatment protocol, will be used for secondary analyses if PP population is less than 90% of the ITT population. It also stated that the Evaluable Portion of the ITT Population (EPP) will be used

for a sensitivity analysis of the primary efficacy endpoint and analyses of other VAS pain intensity based outcomes.

For primary efficacy analysis, a general linear model including randomized treatment and stratification factors with pairwise comparisons between FDC 325/97.5 and each of the three other study treatments were used to compare the efficacy of the four treatments. In order for FDC 325/97.5 to be considered a superior treatment it needs to be statistically superior to all three other study treatments. The applicant did not adjust for baseline pain score in the general linear model for the primary analysis.

For secondary efficacy analysis, Kaplan-Meier curves and stratified log-rank tests were used to compare the times to onset of perceptible and meaningful pain relief, respectively, in the listing order. The statistical significance of key secondary endpoints will only be assessed if the primary endpoint shows a statistically significant advantage to FDC 325/97.5 over the other three treatments. There was no pre-specified testing hierarchy on secondary endpoints.

According to the statistical analysis plan (SAP), the approaches listed below were planned to be used on handling missing data.

For pain intensity:

- Subjects with some missing data points may have intermediate points interpolated and be time adjusted as necessary to allow analysis
- For subjects who discontinue the study during the first 12 hours (N=10), pain intensity scores were imputed with multiple regression method using pre-dropout pain intensity scores, age, gender, stratum and number of molars extracted
- The pre-rescue pain intensity scores were carried forward up to 6 hours.

For time to onset of pain relief:

- Subjects who did not take any rescue medication or have perceptible and meaningful pain relief will be censored at the end of the relevant observation period
- Subjects who took rescue medication and did not achieve the pain relief were censored at 6 hours following the first dose.

However, based upon my review of the submission, the applicant did not strictly follow the pre-specified plan for handling missing data. Since the datasets submitted in the original submission were not adequately documented and were not of sufficient quality, I was not able to conduct my own analyses.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

The ITT population included 408 subjects who randomized and received the first dose of the study medication (See Table 5). The PP population consisted of 390 subjects (106 in FDC; 108 in IBU; 108 in APAP; and 68 in Placebo) and represented 95.6% of the ITT population. Therefore, no PP analysis was considered necessary to be conducted by the applicant. Both primary and secondary analyses were conducted in ITT population. The EPP population was

used to conduct sensitivity analysis which included 398 subjects who randomized and had at least three pain intensity scores in the first 12 hours (10 subjects who withdrew from the study during the first 12 hours were excluded from EPP). There were 14 subjects (3.4%) discontinued the study within the 48-hour double-blind treatment period.

Table 5. Patient Disposition of Study AFT-MX-6

Number of Patients	FDC 325/97.5	IBU 97.5	APAP 325	Placebo	Total
Randomized (ITT)	110	112	111	75	408
Completed first 12hrs (EPP)	107	109	109	73	398
Discontinued first 12hrs based on CSR diagram:					
Adverse event	1	0	2	0	3
Lack of efficacy	1	2	0	2	5
Lost to follow up	1	0	0	0	1
Other	0	1	0	0	1
Total	3	3	2	2	10
Discontinued first 12hrs based on ADaM data ADSL.xpt:					
Withdrawal by subject	2	2	2	1	7
Protocol violation	0	1	0	1	2
Completed*	1	0	0	0	1
Total	3	3	2	2	10
Completed the study (48hrs)	108	109	108	69	394
Discontinued the study based on SDTM data DS.xpt:					
Withdrawal by subject	2	2	3	5	12
Protocol violation	0	1	0	1	2
Total	2	3	3	6	14

Source: Reviewer
*Subject (b) (6) (Premier site) in FDC 325/97.5 group, early withdrawal after Dose 2. Diary not returned, lost to follow up. Pain score data not covered for 12 hours.

The inconsistency highlighted in bold was found between the submitted data and the CSR. In the diagram of patient disposition from CSR, subjects discontinued during the first 12 hours due to adverse event, lack of efficacy, lost to follow up, and other reasons.

3.2.1.4 Results and Conclusions

I was not able to confirm the applicant's efficacy results for the primary and key secondary analyses due to lack of acceptable quality and adequate documentation of the efficacy datasets. I also was not able to examine the impact of missing data on the primary and secondary efficacy analyses. This information was conveyed to the applicant on June 7 and July 12. The applicant confirmed this finding and stated that revised datasets would be submitted by October or November. However, since these datasets were submitted on November 22, 2017, I did not have sufficient time to review the datasets during this cycle.

3.2.2 Study AFT-MX-6E

Study AFT-MX-6E is a multicenter, randomized, prospective, double-blind, placebo-controlled, parallel-design phase 3 study. The efficacy of Maxigesic was compared to each individual component and placebo. Treatment groups were stratified based on baseline pain level (moderate pain with baseline pain score 40-69 mm, and severe pain with baseline pain score ≥70 mm). All the surgeries were conducted under general anesthesia. The study consisted of patient screening, day of surgery, randomization, 24-hour double-blind treatment period, a site visit follow-up at 8-10 days and a telephone follow-up at 30 days. The primary efficacy endpoint was SPID24. The key secondary efficacy endpoints were time to onset of perceptible and meaningful pain relief, respectively.

Similarly as Study AFT-MX-6, primary efficacy analysis for Study AFT-MX-6E adopted a general linear model including randomized treatment and stratification factors with pairwise comparisons between Maxigesic and each of the three other study treatments were used to compare the efficacy of the four treatments. In order for Maxigesic to be considered a superior treatment it needs to be statistically superior to all three other study treatments. The applicant did not adjust for baseline pain score in the general linear model for the primary analysis.For secondary efficacy analysis, Kaplan-Meier curves and stratified log-rank tests were used to compare the times to onset of perceptible and meaningful pain relief, respectively, in the listing order. The statistical significance of key secondary endpoints will only be assessed if the primary endpoint shows a statistically significant advantage to Maxigesic over the other three treatments. There was no pre-specified testing hierarchy on secondary endpoints.

Using the applicant's analysis datasets, the results from my analysis were not consistent with the applicant's. Maxigesic was not significantly different from placebo, acetaminophen, or ibuprofen. Therefore, the results from this study do not support efficacy.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

AFT submitted the subgroup analyses by gender, age, ethinicity, and geographic regions subgroups for the primary efficacy endpoint in response to our first information request. However, I was not able to confim these analyses.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

I was not able to conclude whether or not the results from study AFT-MX-6 provided substaintial evidence of efficacy as the datasets submitted were not of sufficient quality to allow a thorough review of the data. Major statistical issues identified that affected my review were:

- Unable to reproduce the primary analysis results or generate and derive the analysis datasets due to poor quality of data. After sending an IR to the applicant stating such, they discovered that the datasets were generated using Excel and were not able to reproduce the required datasets.
- Missing data was not imputed as pre-specified in the SAP.

AFT is committed to clarifying the existing issues. This includes the following:

- AFT will contract a third party to produce SAS programs to reproduce the primary and secondary analyses from the submitted ADaM domains.
- As an extra level of confirmation, for complete traceability, AFT will contract a third party to produce SAS programs for the derivation of new ADaM datasets as well as primary and secondary endpoint analyses from the submitted SDTM domains.

5.2 Conclusions and Recommendations

Since the datasets submitted by the applicant in the original submission were not of sufficient quality to confirm their primary and secondary analyses, I was not able to confirm that the results from study AFT-MX-6 provided substantial evidence of efficacy. I was also not able to examine the impact of missing data on the primary and secondary analyses. When I requested the applicant submit all programs used to derive the primary and secondary analyses, they were unable to do so. In light of this, the applicant will contract a third party to produce SAS programs to reproduce the primary and secondary analyses from the submitted ADaM domains. The applicant will also contract a third party to produce SAS programs for the derivation of new ADaM datasets as well as primary and secondary endpoint analyses from the submitted SDTM domains. The applicant stated that revised datasets would be submitted by October or November. Since these datasets were submitted on November 22, 2017, I did not have sufficient time to review the data during this cycle.

5.3 Labeling Recommendations

The label will not be reviewed during this cycle.

6 Appendix

Table 6. Examples of data discrepancies between variables QSTPT and QSDTC (Study AFT-MX-6)

USUBJID	QSSEQ	QSTPT	QSDTC	Planned Time from Baseline	Actual Time from Baseline
(b) (6)	1	BASELINE	(b) (6)	0	0.0
	20	IMMEDIATELY BEFORE DOSE 6		30	30.1
	21	APPROXIMATELY 2 HOURS AFTER DOSE 6		32	56.1
	22	IMMEDIATELY BEFORE DOSE 7		36	36.1
	1	BASELINE		0	0.0
	13	APROXIMATELY 2 HOURS AFTER DOSE 2		8	7.0
	14	IMMEDIATELY BEFORE DOSE 3		12	36.0
	15	APPROXIMATELY 2 HOURS AFTER DOSE 3		14	14.0
	1	BASELINE		0	0.0
	21	APPROXIMATELY 2 HOURS AFTER DOSE 6		32	32.0
	22	IMMEDIATELY BEFORE DOSE 7		36	47.9
	23	APPROXIMATELY 2 HOURS AFTER DOSE 7		38	.
	1	BASELINE		0	0.0
	18	IMMEDIATELY BEFORE DOSE 5		24	23.9
	19	APPROXIMATELY 2 HOURS AFTER DOSE 5		26	30.0
	20	IMMEDIATELY BEFORE DOSE 6		30	36.0
	21	APPROXIMATELY 2 HOURS AFTER DOSE 6		32	.

Source: Reviewer

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

/s/

YI N REN
11/28/2017

DAVID M PETULLO
11/28/2017
I concur.