Table II.2.8 Time Period Analysis of VAS

Factors	Hours 0-3	Hours 4-6	Hours 7-9	Hours 10-12
Among Treatments	0.163*	0.011	0.628	0.354
400 µg vs. placebo	NA	0.310	NA	NA
800 µg vs. placebo	NA	0.003	NA	NA
800 vs 400 µg	NA	0.054	NA	NA
Group by Time	0.007	0.198	0.741	0.970
400 µg vs. placebo	0.729	NA	NA	NA
800 µg vs. placebo	0.004	NA	NA	NA
800 vs 400 µg	0.021	NA	NA	NA
Treatment by Center Interaction	0.288	0.151	0.939	0.821
Time by Treatment by Center Interaction	0.930	0.994	0.225	0.341

*: P-value of ANOVA

Global Pain Evaluation -

Patients received 800 µg OTFC had statistically significant lower global pain score than the other two groups prior to the second administration period. Though the same pattern showed in administration #3 and #4, the differences among the three groups were not significant in either of the two administrations.

Table II.2.9, Analysis of Global Pain Evaluation

	Evaluation Score No.(%)						
	1 - None	2 - Mild	3 - Discomforting	4 - Distressing	5 - Horrible	6 - Excruciating	Among Trt p-value
Prior to Adm 2							0.037[1]
Placebo (33)	0 (0)	8 (24)	9 (27)	12 (36)	2 (6)	2 (6)	
400µg(34)	0 (0)	9 (26)	9 (26)	9 (26)	6 (18)	1 (3)	0.985[2]
800µg(34)	2 (6)	13 (36)	12 (35)	2 (6)	5 (15)	0 (0)	0.042[2]
Prior to Adm 3							0.228
Placebo (31)	1 (3)	10 (32)	12 (39)	6 (19)	2 (6)	0 (0)	
400µg(30)	1 (3)	12 (40)	8 (27)	5 (17)	4 (13)	0 (0)	
800µg(31)	3 (10)	13 (42)	9 (29)	2 (6)	2 (6)	0 (0)	

Prior to Adm 4							0.326
Placebo (26)	1 (4)	10 (38)	9 (35)	5 (19)	1(4)	0 (0)	
400µg(29)	3 (10)	12 (41)	6 (21)	5 (17)	3 (10)	0 (0)	
800µg(22)	3(14)	9 (41)	7 (32)	2 (0	1 (5)	0 (0)	

1: Mantel-Haenszel dose response test Among all three groups; 2: Comparing 400 µg with placebo; 3: Comparing 800µg with placebo

Sedation -
There was no significant difference among the treatment groups.

Analysis of Safety Data
Respiratory depression was defined in this study as a respiratory rate of 8 breaths per minute or an oxygen saturation of 85% or less. Respiratory depression showed a significant dose response relationship when all three treatment groups were analyzed. Significantly more patients who received 800 µg OTFC had respiratory depression than patients who received placebo (p=0.002, Extended Mantel-Haenszel test) or 400 µg OTFC (p=0.023, Extended Mantel-Haenszel test). There was no significant difference between patients treated with 400 µg OTFC and placebo.

Respiratory rate decreases that met the criteria of being potentially clinically significant showed a significant dose response relationship. The placebo group produced fewer decreases than either 400 µg OTFC group (p=0.014, Extended Mantel-Haenszel test) or 800 µg OTFC group (p=0.018, Extended Mantel-Haenszel test). The difference between the two OTFC groups was not significant.

Twenty-five of the twenty-seven withdrawals in this study were due to an adverse event. Six patients in the placebo group, nine in the 400 µg OTFC group and ten in the 800 µg OTFC group were withdrawn because of adverse events.

II.2.5 Reviewer's Evaluation
In this controlled clinical trial, the sample size was not statistically determined for a pre-specified target efficacy size. However, this study shows the treatment effect of 800 µg OTFC in reduction of morphine usage (with p=0.02 in repeated measurement analysis, Table II.2.6), and the effect is significant within 6 hours after administration (p=0.011 at the first administration and p=0.004 at the 2nd administration, Table II.2.5). Patients treated with 400 µg OTFC did not have a significant reduction in morphine usage as compared with placebo patients. Patients in Duke medical center experienced more reduction than Utah medical center as shown in Figure 2 (pp. 10-602 of NDA submission) of sponsor's submission.

Both respiratory depression and respiratory rate decrease showed a significant dose response relationship such that the rates increased with the dose.

Twenty-five of the twenty-seven withdrawals in this study were due to an adverse event. Six

patients in placebo group, nine in 400 μg OTFC group and ten in 800 μg OTFC group were withdrawn because of adverse events.

III. Clinical Trial of Patients with Chronic Pain AC200/013

AC200/013 is a placebo controlled, double-blind, multicenter (23 centers), 2-phase crossover study. It consists of an open label dose titration phase and a randomized double-blind, crossover phase.

III.1 Study population

Patients with pain related to cancer or cancer treatment were eligible to participate in the study if they were regularly having at least one but no more than four episodes of breakthrough pain per day while taking 60-1000 mg oral morphine per day or 50-300 μg/hr transdermal fentanyl to treat pain associated with their disease. If patients had more than one type of breakthrough pain or had breakthrough pain in more than one location, they were asked to identify one of the pains and consider it their "target" breakthrough pain. OTFC was used to treat the patient's target breakthrough pain and efficacy evaluations were made for treatment of only the target breakthrough pains.

Table III.1 Number of patients Enrolled by Center

Center	# Pts Enrolled	Center	# Pts enrolled
Robert Berris, MD	16	James Cleary, MD	6
Allen Cohen, MD	2	Robert Ellis, MD	8
John Farrar, MD	13	Janet Gargiulo, MD	2
Stuart Grossman, MD	2	Lowell Hart, MD	17
Laurel Herbst	4	Howard Homesley, MD	9
Laura Hutchins, MD	1	K.S. Kumar, MD	2
Michael Levy, MD	1	John Marshall, MD	2
Timothy Ness, MD	1	Kelly Pendergrass, MD	10
Richard Rauck, MD	11	Lee Schwartzberg, MD	5
Mark Seligman, MD	3	Gregory Smith, MD	4
Charles von Guntman, MD	3	William Whaley, MD	6
Donna Zhukovsky, MD	7	Total	130

III.2 Study Design

AC200/013 consists of two trial phases, the open label, dose titration phase, and the double-blind crossover phase.

In the open label, dose titration phase, the patients were titrated to the dose of OTFC at which the patient received adequate pain relief for an episode of breakthrough pain using a single dose of OTFC. Patients started on a 200 μg dose of OTFC and patients self-administered a complete OTFC dose for each episode of breakthrough pain. Patients were encouraged to used their regular rescue medication if additional pain relief was needed. Patients were contacted daily by the study staff to assess the dose titration and any problem they might be experiencing. If patients reported excessive or intolerable adverse events, the OTFC dose was decreased. Patients who consistently achieved effective relief using a single OTFC unit for each episode of breakthrough pain were eligible for the crossover phase. If a patient was not able to achieve effective pain relief from the highest tolerated OTFC dose, or if dose titration continued for more than a month, the patient would be discontinued from the study.

In Double-blind crossover phase, patient were given 10 randomized, pre-numbered oral transmucosal (OT) units. Seven of the units were OTFC dose found to be effective in the titration phase, 3 units were placebo. Patients were instructed to take a single OT units in the designated order (1 to 10), for each episode of breakthrough pain. If the patient did not experience significant pain relief 30 minutes after start of study drug administration, patients were instructed to take their normal rescue medication. Additional episodes of breakthrough pain could be treated with study medication after 2 hours had elapsed since taking the previous dose. Patients remained in the study until all units were taken or for fourteen days.

III.3 Efficacy Data

On each treatment day, patients reported on each episode of breakthrough pain treated with the study drug, noting pain intensity, pain relief and any adverse event. Patients completed a performance evaluation (see the following table) of study drug 60 minutes after starting the study drug or at the time of additional medication.

Table III.2 Efficacy Evaluation Schedule

Scheduled Time	Pain Intensity (PI)	Pain Relief (PR)	Global Performance Evaluation
Titration Phase:			
0 (prior to OTFC)	x		
15 minutes	x	x	
30 minutes	x	x	
45 minutes	x	x	
60 minutes	x	x	x

Double-Blind Phase:			
0 (Prior to OTFC)	x		
15 minutes	x	x	
30 minutes	x	x	
45 minutes	x	x	
60 minutes	x	x	x

For each patient, a maximum of 10 episodes were planned with the study drugs and each of the episode was determined as evaluable or unevaluable. All together, seventy four episodes were considered unevaluable, of which fifty two were active drug and twenty two were placebo drug. The reasons the seventy four episodes were considered unevaluable are given in the following table.

Table III.3 Summary of Reasons an Episode Was Considered Unevaluable

Reasons Episode Considered Unevaluable	Number of Episodes Unevaluable		
	Active	Placebo	Total
At least one observation time was outside the window allowed	27	13	40
ATC medication changed	1	1	2
<2 hrs since previous OT	13	1	14
Patient did not follow protocol in phase 1	6	3	9
Episode not target breakthrough pain	1	0	1
OT consumption not complete	12	7	19
Total[a]	52	22	74
Active to Placebo Ratio	52:22 = 2.36		
Ratio of Total Episodes	552:247 = 2.255		

a: Episode could be unevaluable for multiple reasons.

III.4 Completion Status

Of the 130 patients entering the titration phase, a total of 37 (29%) patients withdrew from the study before the end of the titration phase. Of them, 22 (17%) patients withdrew from the study due to adverse event, and 15 (12%) withdrew for other reasons including no effective dose. Ninety-three (72%) found an effective OTFC dose during the titration phase. One patient completed the titration phase but did not enter the double-blind phase.

Among the 92 patients entering the double-blind phase, a total of 20 (22%) patients withdrew from the study in the double-blind phase. Of them, 7 (8%) patients withdrew due to adverse

event and 13 (14%) patients withdrew in this phase for other reasons. A total of 72 (78%) patients completed the double-blind phase. Hence a total of 55.4% of patients entering the study found the effective OTFC dose and completed both phase of the study. The distribution of patients who withdrew from the study and who completed the study in each center is given in Table 11 (vol.1.25 p10-129b of NDA submission). The reasons of withdrawals other than adverse event were given in Table 12 (vol.1.25 p10-130 to 131 of NDA submission). In summary, during the titration phase, 6 patients withdrew from study because the breakthrough pain ceased or decreased, 4 patients requested to withdraw because of preference for rescue medication other than OTFC, 4 patients withdrew because either unable or unwilling to complete diaries, 3 patients requested withdrawal without providing reason. During the double-blind phase, 2 patients requested different therapy, 1 patient was scheduled to have radiation therapy, and the other 10 patients withdrew either before receiving any dose or due to study closure.

There was no obvious heterogeneity among the distribution of withdrawals among the centers or between OTFC and placebo episodes. Since the efficacy comparison between the study drug and placebo was carried out between the episodes within the same patient, the large percentage of withdrawals would not significantly impact the study result when the reasons of withdrawal were not treatment episode related.

Table III.4 Patient Flow Chart - All patient enrolled = 130

Patient Status	No (%)
Number of Patients Enrolled	130
Received Drug and Entered Titration Phase	130 (100)
Total withdrawals	37(29)
Withdrew due to adverse event in titration phase	22(17)
Withdrew due to other reasons in titration phase	15(12)
Completed titration phase	93(72)
Entered Double-Blind Phase	92(100)
Total withdrawals	20 (22)
Withdrew due to adverse event in double-blind phase	7(8)
Withdrew due to other reasons in double-blind phase	13(14)
Completed 10 episodes in double-blind phase	72(78)

III.5 Withdrawals

Twenty-two (18.3%) patients withdrew due to an adverse event during the titration phase. Sponsor claimed that only ten (8.3% of enrolled) were considered at least possibly related to study drug. In the double-blind phase, seven patients withdrew due to an adverse event. Sponsor claimed that six of the adverse events were unrelated or unlikely related to study drug.

III.6 Efficacy Endpoints

The primary efficacy endpoints are Summary Pain Intensity Difference (SPID) and Total Pain Relief (TOTPAR). SPID and TOTPAR were defined in the following way. Let P_i and PR_i represent the scores of pain intensity and pain relief respectively at the i-th scheduled time point averaged over all evaluable episodes within patient for OTFC and for placebo. P_0 represents the pain intensity prior to study drug.

$PID_i = P_0 - P_i$
$SPID_i = SPID_{i-1} + PID_i$
$TOTPAR_i = TOTPAR_{i-1} + PR_i$

with $PID_0 = SPID_0 = TOTPAR_0 = PR_0 = 0$.

Missing scores happened when additional rescue medication was taken within 60 minutes after the study drug in the crossover phase. When a score was missing in each episode, the most recent score was used in the calculation (last observation carried forward) of P_i and PR_i. Of the total of 804 episodes in crossover phase, additional rescue medication was given in 165 episodes (following table). Eighty-one of the 165 episodes were placebo episodes. The ratio of rescue medication is much higher in placebo episodes as compared with the ratio of all episodes.

Table III.6 Time Until Rescue Medications in Double-Blind Phase

Time Until Rescue Medication	OTFC Episode No. (%)	Placebo Episode No. (%)
0-29 min	0	0
30-44 min	43(8)	49(20)
45-59 min	11(2)	14(6)
≥ 60 min	30(5)	18(7)
Additional Meds Never Taken	473(85)	166(67)
Total Episodes	557	247
Ratio of Rescue Medication	84:81 = 1.04	
Ratio of Total Episodes	557:247 = 2.255	

For analytical purpose, the small centers were pooled; the centers with only one patient were pooled as center 9999, centers with two patients were pooled as center 8888 and center with 3 patients were pooled as center 7777.

Table III.7 Number of Patients Completed the Study

Center	Number of Patients			Center	Number of Patients		
	Phase 1	Phase 2	Complete		Phase 1	Phase 2	Complete
Berris, MD	16	10	9	Cleary, MD	6	5	3*
Cohen, MD	2	2	1*	Ellis, MD	8	6	6
Farrar, MD	13	10	6	Gargiulo, MD	2	2	2**
Grossman, MD	2	1	0	Hart, MD	17	12	10
Herbst	4	2	0	Homesley, MD	9	6	5
Hutchins, MD	1	1	1*	Kumar, MD	2	2	2**
Levy, MD	1	0	0	Marshall, MD	2	2	2**
Ness, MD	1	1	0	Pendergrass, MD	10	8	8
Rauck, MD	11	8	8	Schwartzberg, MD	5	3	3***
Seligman, MD	3	2	0	Smith, MD	4	3	2**
von Guntman, MD	3	1	1*	Whaley, MD	6	5	2**
Zhukovsky, MD	7	1	1*	Total	130	93	72

* Pooled as center 9999; ** Pooled as center 8888; Pooled as center 7777.

III.7 Patient Population (in Double-Blind phase)

Demographic statistics of patients in double-blind phase are given in the following table

Table III.8 Demographic Statistics of Patients Completed the Double-Blind Phase

Variable	Value	Variable	Value	Variable	Value
Age (yr)		Height (cm)		Weight (kg)	
Mean	54±12	Mean	169±10	Mean	70±20
SEM	1.3	SEM	1.1	SEM	2.1
Range	27-84	Range	142-193	Range	40-129
Gender		Race			
Female n(%)	51(55)	Black n(%)	5 (5)		
Male n(%)	41(45)	Asian n(%)	1 (1)		
		White n(%)	86 (93)		

Patient descriptions including cancer diagnosis, pain pathophysiologies, severity of pain at screening, opioid ATC medications, opioid rescue medications, medication levels (in morphine equivalence), medication levels for persistent and breakthrough pain, dose of rescue medication

in relation to ATC medication are given in Tables 16 to Table 22 and Figures 1 and 2 of NDA submission (vol 1.25 pp.10-135-10-141).

III.8 Sponsor's Efficacy and Safety analysis (Double-Blind Phase)

III.8.1 Evaluability Status of Patients in Double-blind Phase

The evaluability status of patients entering the double-blind phase are given in the following table.

Table III.9 Evaluability Status - All patients who entered double-blind phase (N=92)

Patient Status	No.
Entered Double-Blind Phase	92
No. Episodes Treated with Placebo only	2
No. Episodes Treated with Active only	1
Intent-to-Treat Analysis	89
Episodes Treated with Active and Placebo, but all Placebo Episodes Unevaluable	1
All Episodes Unevaluable	2
Evaluable for PI, PR. and Additional Rescue	86
No Performing Rating for Any Placebo Episodes	2
Evaluable for Global Performance Rating	84

III.8.2 Dosing in Double-Blind Phase

The mean dose taken during the double-blind phase was about 800µg. There was no difference in OTFC dose for the three completion status groups (p=0.57). All dose levels (200 µg-1600 µg) provided in the study were required by patients. The distribution of patients at each of the unit dose level was even (See Table III.9).

Table III.10 OTFC Dose During Double-Blind Phase by Patient Completion Status

OTFC Dose	Completed 10 Units No. (%) (n=73)	AE Withdrawal No. (%) (n=7)	Other Withdrawal No.(%) (n=13)	Total No.(%)(n=92)	p-value (Completion Status)
200µg	7(10)	1(14)	5(38)	13(14)	
400µg	16(22)	2(29)	1(8)	19(21)	
600µg	12(17)	0(0)	2(15)	14(15)	
800µg	14(19)	2(29)	2(15)	18(20)	
1200µg	12(17)	0(0)	1(8)	13(14)	
1600µg	11(15)	2(29)	2(15)	15(16)	
Mean±SD	808±452	829±571	662±519	789±468	0.57
Standard Error of Mean	53	216	144	49	

III.8.3 Pain Intensity Analysis

The analysis of pain intensity data are shown in the following table. The difference in mean PI score was not significant at baseline (p=0.06). OTFC produced lower PI scores than placebo, beginning at the 15 minute time point and at every time point to 60 minutes (all p<0.0001) (See the following table). OTFC reduced pain intensity from baseline better than placebo beginning at 15 minute after the study drug administration, as seen in the comparison of pain intensity difference (PID). The differences are statistically significant with p-values less than 0.0001 in three way ANOVA with treatment, center and center-by-treatment interaction as factors.

The efficacy in reducing pain intensity is also shown in the analysis of summed pain intensity difference as given in the following table. The differences between OTFC and placebo are statistically significant with p-values less than 0.0001 at every time point.

The treatment differences among centers were statistically significant at 45 minute and 60 minute time point for PI and at the 60 minute time point for PID and SPID. The interactions were quantitative rather than qualitative.

Table III.11 Mean values of Pain Intensity, Pain Intensity Difference, and Summed Pain Intensity Difference

	Time in Minutes				
	0	15	30	45	60
PAIN INTENSITY SCORE					
OTFC					
Mean	5.87	4.25	3.46	2.99	2.68
SD	1.93	1.95	1.82	1.76	1.74
Standard Error	0.21	0.12	0.30	0.19	0.19
Minimum	2.43	0.71	0.57	0.14	0.00
Maximum	10.00	48.07	8.00	8.00	8.00
Placebo					
Mean	6.01	4.99	4.50	4.10	3.89
SD	1.98	2.06	1.90	1.95	2.03
Standard Error	0.21	0.22	0.21	0.21	0.22
Minimum	2.00	0.00	0.00	0.00	0.00
Maximum	10.00	9.67	9.00	9.00	9.00
p-values					
Treatment	0.06	<0.0001	<0.0001	<0.0001	<0.0001
Center	0.03	0.005	0.008	0.02	0.03
Treatment *Center	0.14	0.13	0.13	0.07	0.03
PAIN INTENSITY DIFFERENCE					
OTFC					
Mean		1.62	2.41	2.88	3.19
SD		1.16	1.37	1.46	1.60
Standard Error		0.12	0.15	0.16	0.17
Minimum		0.00	0.00	0.00	0.00
Maximum		4.43	5.80	7.00	7.75
Placebo					
Mean		1.02	1.51	1.91	2.13
SD		1.10	1.27	1.58	1.78
Standard Error		0.12	0.14	0.17	0.19
Minimum		-1.33	-1.00	-1.33	-1.33
Maximum		5.0	5.50	5.50	7.50

p-values				
Treatment	<0.0001	<0.0001	<0.0001	<0.0001
Center	0.02	0.20	0.27	0.12
Treatment*Center	0.11	0.17	0.07	0.03
SUMMED PAIN INTENSITY DIFFERENCE				
OTFC				
Mean	1.62	4.03	6.92	10.11
SD	1.16	2.44	3.79	5.20
Standard Error	0.12	0.26	0.41	0.56
Minimum	0.00	0.40	0.60	0.60
Maximum	4.43	9.60	16.33	23.83
Placebo				
Mean	1.02	2.53	4.44	6.56
SD	1.10	2.29	3.77	5.41
standard Error	0.12	0.25	0.41	0.58
Minimum	-1.33	-2.33	-3.67	-5.00
Maximum	5.0	10.5	16.00	21.50
p-values				
Treatment	<0.0001	<0.0001	<0.0001	<0.0001
Center	0.02	0.07	0.15	0.20
Treatment*Center	0.11	0.14	0.11	0.08

III.8.4 Pain Relief Analysis

Mean pain relief. PR, scores and mean total pain relief, TOTPAR, scores are given in the following Table III.11. OTFC provided significantly more pain relief than placebo beginning at 15 minute time point and continuing through 60 minute time point (all p-values <0.0001 in a three way ANOVA).

It was also shown that there was no significant difference in pain relief between those who completed the double-blind phase and those who did not (NDA SUPPLMT Table S16).

Table III.12 Mean values of Pain Relief (PR), Total Pain Relief (TPR)

	Time in Minutes			
	15	30	45	60
PAIN RELIEF				
OTFC				
Mean	1.42	1.80	2.00	2.14
SD	0.76	0.78	0.78	0.84
SEM	0.08	0.08	0.08	0.09
Minimum	0.00	0.00	0.57	0.57
Maximum	3.29	3.43	3.86	4.00
Placebo				
Mean	0.93	1.11	1.30	1.33
SD	0.81	0.82	0.90	0.91
SEM	0.09	0.09	0.10	0.10
Minimum	0.00	0.00	0.00	0.00
Maximum	4.00	4.00	4.00	4.00
p-values				
Treatment	<0.0001	<0.0001	<0.0001	<0.0001
Center	0.09	0.02	0.012	0.05
Treatment*Center	0.26	0.17	0.08	0.14
TOTAL PAIN RELIEF				
OTFC*				
Mean	1.42	3.23	5.23	7.37
SD	0.76	1.46	2.16	2.89
SEM	0.08	0.16	0.23	0.31
Minimum	0.00	0.20	0.80	1.40
Maximum	3.29	6.71	10.14	13.71
Placebo				
Mean	0.93	2.04	3.34	4.67
SD	0.81	1.55	2.38	3.23
SEM	0.09	0.17	0.26	0.35
Minimum	0.00	0.00	0.00	0.00
Maximum	4.00	8.00	12.00	16.00

p-values				
Treatment	<0.0001	<0.0001	<0.0001	<0.0001
Center	0.09	0.04	0.03	0.03
Treatment*Center	0.26	0.16	0.12	0.12

III.8.5 Patients' Global Performance Evaluation Analysis

As shown in the following table, the mean global performance ratings for OTFC is significantly higher than the placebo group.

Table III.13 Mean Global Performance Rating During the Double-Blind Phase

Variable	Group		P-values (ANOVA)		
	OTFC	Placebo	Treatment	Center	Trt*center
Mean Global Performance Rating (N=84)	1.98	1.19	<0.0001	0.05	0.06

III.8.6 Additional Rescue Medication Analysis

Additional rescue medication was taken for thirty-five percent of the placebo episodes. In contrast, rescue medication was taken in fifteen percent of the OTFC episodes. The difference is statistically significant (p-value<0.0001, three-way ANOVA).

III.8.7 Alternative Intent to Treat Analysis

In order to determine if the exclusion of unevaluable episodes and score imputation might affect the analysis, then sponsor also performed an intent to treat analysis. In intent to treat analysis, all data were assumed to be collected at the scheduled time and no 'last observation carried forward' was used. The p-values of the analysis of PI and PR are shown in the following table. A comparison of the results in Table III.14 with Tables III.12 and III.13, indicates that the effect of using last observation carried forward is very minor.

Table III.14 Intent to Treat Analysis of PI and PR

	Time in Minutes				
	0	15	30	45	60
PAIN INTENSITY SCORE					
OTFC					
Mean	5.84	4.18	3.37	2.60	2.26
SD	1.90	1.96	1.82	1.70	1.61
Placebo					
Mean	5.94	4.86	4.34	3.43	3.07
SD	2.00	2.09	1.99	2.00	1.91

p-values					
Treatment	0.16	<0.0001	<0.0001	0.0003	0.0007
Center	0.04	0.003	0.002	0.004	0.01
Treatment *Center	0.13	0.08	0.18	0.15	0.07
PAIN RELIEF					
OTFC					
Mean		1.45	1.85	2.21	2.37
SD		0.81	0.81	0.90	0.92
Placebo					
Mean		0.98	1.19	1.64	1.67
SD		0.88	0.92	1.08	1.06
p-values					
Treatment		<0.0001	<0.0001	0.0004	<0.0001
Center		0.03	0.002	<0.0001	0.0003
Treatment*Center		0.11	0.38	0.03	0.13

III.9 Reviewer's Evaluation

The design of this trial is different from the double-blind, randomized, parallel trial used in AC400/001 and AC200/006 of patients with postoperative pain. This crossover trial design is efficient for patients with chronic pain. But the sample size of the study was not statistically determined for a pre-specified target efficacy size. In this study, there was a rather high percentage of withdrawals (44% withdrawals of all patients). In titration phase, there was 18 percent of withdrawals due to adverse events. Eight percent of patients withdrew in double blind phase due to adverse effects. Dizziness, nausea, and somnolence were the most common adverse events associated with OTFC. Dizziness, nausea and somnolence were also the most common adverse events during the OTFC phases in the two titration studies (AC200/011 and AC200/012). In the two titration studies the percentages of the three adverse events were 15% (in either AC200/011 or AC200/012) with dizziness, 20% (in AC200/011) and 21% (in AC200/012) with nausea, and 28% (in AC200/011) and 18% (in AC200/012) with somnolence.

This study supports the treatment efficacy of OTFC in pain relieving among cancer patients with chronic pain. It shows that patients had significantly lower pain intensity during their OTFC episodes than in their placebo episodes after 15 minutes of treatment ($P<0.0001$)(See Table III.10). The efficacy is also supported by comparing pain intensity difference and summed pain intensity and pain relief (Tables III.10 and III.11). All the comparison have p-values less than 0.0001 without multiple comparison adjustment. The study also shows that 35% of the placebo episodes required additional rescue medication in comparison to 15% of all the OTFC episodes. The difference is statistically significant. The analysis was also carried out using both 'last observation carried forward' for the unevaluable episodes and using intent-to-treat procedure.

Results of the two procedures are consistent.

Treatment dosage in this study is not uniformly defined for all patients. The data shows that the effective dosage size of OTFC unit used in the OTFC episodes ranged from 200 μg to 1600 μg.

IV. Dose Equivalence Trial AC2000/010

The Anesta AC2000/010 is a randomized, multicenter, double-blind parallel clinical trial that is designed for the following two objectives:

1. Establish the dose equivalency of OTFC 200 μg (changed from 300 μg in the first amendment) and 800 μg in treating moderate-to-severe postsurgical pain relative to morphine 2 mg and 10 mg (changed from 6 mg in the first amendment).

2. Establish the safety and tolerability of OTFC in patients with moderate-to-severe pain following lower abdominal surgery.

IV.1. Study Design

The study was designed as a five center, double blind, randomized block, four-arm, parallel-group clinical trial. It was conducted on 120 postsurgical patients with moderate-to-severe pain resulting from lower abdominal surgery. Patients were randomly assigned to either one of the two doses of OTFC (200 or 800 μg) combined with a placebo 10 mL IV injection or one of two doses of IV morphine (2 or 10 mg in 10 mL) injection combined with an OT placebo. Patients were encouraged to remain in the study for at least one hour. They were followed until they requested additional analgesia, for up to six hours after study drug administration. Adverse events were collected for 24 hours following study drug administration.

There are six efficacy endpoint parameters. They are defined as follow (in the order of importance)

1. PCA (patient controlled analgesia) Morphine Use - Measured cumulatively from midnight to 4 am and then from 4 am to PCA discontinued;
2. Time In Study - the time patients were in the study before requesting rescue analgesia;
3. Pain Intensity (at each scheduled time point) - the observed pain score or the most recent observed score when a score is missing. It is a 100 mm visual Analog Scale (VAS) (0=no pain,...., 100=worst imaginable);
4. Pain relief (at each time point) - defined the same way as Pain Intensity;
5. Time of Onset of Meaningful Pain Relief - Stopwatch
6. Global Assessment of Pain Relief - 5-pt ordinal scale (1=excellent,...., 5=poor),

The four safety outcome parameters are periodic vital signs, oxygen saturation, adverse events, respiratory depression.

IV.2 Patient Screening and Clinical Population

American Society of Anesthesiology (ASA) physical status I-II patients undergoing lower abdominal surgery were recruited and enrolled in five hospitals. PCA morphine was discontinued on the morning following surgery. When patients reported moderate-to-severe pain and requested analgesia, they were randomly assigned to one of the four treatments.

IV.3 Protocol violations

There were sixty-eight protocol violations in sixty-eight patients during the study. All violations are minor.

IV.4 Patient Evaluability

Patients were classified as fully evaluable, partially evaluable or unevaluable for efficacy. Patient evaluability was assessed after the completion of the study and before the study blind was broken.

Patients were considered partially valuable for efficacy for one of three reasons: 1) patient received concomitant analgesic after receiving study drug, in which case efficacy data after the concomitant medication were excluded; 2) patient had an OT consumption time greater than 20 minutes, in which case data from some efficacy parameters were excluded (time until meaningful relief, proportion experiencing meaningful relief, global assessment, and pain intensity and pain relief scores prior to 60 minutes); 3) patients had assessment(s) made more than 10% or 5 minutes off of the scheduled time, in which case the assessment was excluded from the efficacy analysis.

Patients were considered unevaluable for efficacy for the following reasons: 1) patient received concomitant analgesic after midnight and prior to study start; 2) patient did not consume 90% of study drug, 3) patient was enrolled at a center that did not have at least one evaluable patients in every treatment group.

All patients who received drug were considered evaluable for safety.

Dispositions of patients by center were given by sponsor in Tables 7a and 7b. One center, Brigham & Women's Hospital, had only one patient in each treatment group receiving study treatment.

The total numbers of patients evaluable, partially evaluable and not evaluable were appeared in sponsor's Table 7a. The missing pattern is random among the treatments with p-value =0.88 for Mantel-Haenszel test for equal propostions of evaluable patients of the randomized among the four treatment groups. The Texas-SW center had a much smaller sample size (30-40% of the size of the Chicago center).

Table IV.1 (Sponsor's Table 7a) Disposition of patients (135 patient randomized)

	OTFC		Morphine		p-value Trt
	200 µg	800 µg	2 mg	10 mg	
# Pts Randomized	34	33	34	34	
# Pts Received Drug	33	32	34	34	
# Pts Evaluable for Efficacy	30	30	31	32	0.88[a]
# Pts Fully Evaluable	27	26	28	29	
# Pts Partially Evaluable	3	4	3	3	
# Pts Unevaluable for Efficacy	3	2	3	2	
# Pts Evaluable for Safety	33	32	34	34	

a: Mantel-Haenszel test between treatments

IV.5 Patient characteristics

Weight, gender, ASA class, age, height, and race of patients on the four treatment are not different significantly (See Sponsor's Table 8), though the mean age of the 10mg morphine patients was 6 years older than the patients treated with 800µg OTFC (47 versus 41). None of the characteristics has significant treatment-by-center interaction.

IV.6 Surgical procedures

Surgical procedures were grouped into five categories - hysterectomy (no cancer), hysterectomy (cancer), other gynecological, colorectal and others. The majority of patients in all treatment groups were undergoing hysterectomy or related procedures (Sponsor's Table 9 of NDA submission). One patient was withdrawn from the study prior to requesting additional analgesia.

Table IV.2 (Sponsor's Table 8) Patient Characteristics

	OTFC		Morphine		P-value		
	200 µg	800 µg	2 mg	10 mg	Trt	Cntr	TrtxCntr
# of Pts	33	32	34	34	-	-	-
Gender M/F	3/30	1/31	1/33	1/33	0.34[a]	-	-
ASA Class I/II	8/25	10/22	10/24	7/27	0.77[a]		
Age (yr)							
Mean (SD)	42 (10)	41 (8)	43 (10)	47 (9)	0.07[b]	0.28	0.29
SEM	1.7	1.5	1.7	1.5			
Range	21-60	28-61	21-65	26-63			

Height (cm)							
Mean (SD)	161 (9)	167 (7)	164 (7)	163 (8)	0.19[b]	0.055	0.32
SEM	1.5	1.3	1.2	1.4			
Range	135-183	153-181	152-180	142-178			
Weight (kg)							
Mean (SD)	71 (15)	71 (13)	71 (17)	71 (13)	0.18[b]	0.44	0.13
SEM	2.6	2.2	2.9	2.2			
Range	45-100	51-96	51-120	51-92			
Race							
Black	15	17	14	13	0.20[a,c]		
White	14	11	20	20			
Other	4	4	0	1			

a: Mantel-Haenszel test, b: Two-way ANOVA, c: Only Black and White

Table IV.3 (Sponsor's Table 9) Surgical Procedures

	OTFC		Morphine	
	200 µg	800 µg	2 mg	10 mg
Hysterectomy (Noncancer)	16	19	18	18
Hysterectomy (Cancer)	5	4	7	9
Other Gynecological	9	8	7	5
Colorectal	2	1	1	1
Other	2	1	1	2

IV.7 Sponsor's Analysis

IV.7.1 Prestudy and Baseline Comparisons

Cumulative PCA morphine use during 0-4 a.m, cumulative PCA morphine during 4 a.m. -PCA off, total PCA morphine use, time from 4 a.m, to PCA off, pain intensity score when PCA off, pain intensity score at baseline, and increase in pain intensity from PCA off to baseline were compared among treatments. Only the difference in cumulative PCA morphine use during 4 a.m. to PCA off among the treatment groups is statistically significant. There were significantly higher levels of PCA morphine usage in the low dose groups than the high dose groups. Center differences were also identified in time from a.m. to PCA off, in time from PCA off to baseline and in pain intensity at baseline. Moderate center-by-treatment interactions were also detected in some measurements (See sponsor's Table 11). Although the PCA morphine levels were higher in low dose groups than the higher dose group, it was comparable between OTFC and IV

morphine patients in either high or low dose groups.

IV.7.2 Efficacy Analysis

Sponsor provided six statistical analyses of the efficacy parameters.

Time to rescue analysis -

Sponsor performed two separate analyses on this parameter. The first is a life table analysis that is used to test if the survival rates of OTFC patients were significantly different than IV morphine patients. The second analysis is to compare the mean time to rescue of the OTFC groups and the corresponding IV morphine groups. The p-values of Wilcoxon pairwise comparison are 0.69 (OTFC 800 µg vs. IV morphine 10 mg), 0.04 (OTFC 800 µg vs. OTPC 200 µg), and 0.03 (OTFC 800 µg vs. IV morphine 2 mg). The second is ANOVA of means of time to rescue with treatment, center and center-by-treatment interaction. The ANOVA gave a p-value of 0.01 for difference among the four treatments. The F-test of center effect is also significant with a p-value of 0.01. It is also shown that patients in Duke (168 min) and Yale (152 min) had significantly (p-value <.02) shorter time than Chicago (214 min) and Texas SW (218 min). Almost all patients remained in the study for at least 1 hr with no difference among the treatment groups (See Sponsor's Table 12 (below) and Figure 1).

Table IV.4 (Sponsor's Table 11) Time to Rescue Analysis

Time in Study	OTFC		Morphine		p-value		
	200 µg (n=30)	800 µg (n=30)	2 mg (n=31)	10 mg (n=32)	Trt	Cntr	TrtxCntr
Median (min)	145	215	130	188	0.03[A]	0.01	0.14
MEAN(SEM)	159(17)	220(20)	153(15)	210(18)	0.01[B]		
Range	40-360	30-362	25-360	61-360			
< 1 hr (n%)	1 (3%)	2 (7%)	3 (10%)	0(0%)	0.27[C]		
1-6 hrs (n %)	26(87%)	19(63%)	26(84%)	27(84%)			
Completed 6 hrs	3 (10%)	9 (30%)	2 (6%)	5 (16%)			

Pain intensity scores -

Sponsor performed comparisons (including among all treatments, among all centers, treatment-by-center interaction, OTFC vs. Morphine, high dose vs. low dose) of mean imputed PI score at each single time point using two-way ANOVA. It is shown that the difference between OTFC and morphine is not statistically significant when pooled or in dose stratified setting (See Attached Abbreviated Table of Sponsor's Figure 2). The dose response is shown at 120 minutes and later with p-values of F-test being less than .05 at 180, 240 and 300 minutes (See Sponsor's Figure 2).

Table IV.5 (Abbreviated Table of Sponsor's Figure 2) P-Values of the Comparisons of Mean Pain Intensity Score - Evaluable Patients =123

Test[a]	Time in minutes										
	PCAoff	0	15	30	45	60	120	180	240	300	360
Complete ANOVA[A]											
Treatment (T)	0.20	0.92	0.16	0.13	0.41	0.17	0.11	0.01	0.01	0.12	0.12
Center ©	0.16	.005	0.13	0.39	0.18	0.10	0.12	0.16	0.30	0.21	0.26
T*C Interaction	0.45	0.88	0.80	0.87	0.99	0.91	0.83	0.32	0.24	0.60	0.37
High vs Low[B]			0.77	0.61	0.55	0.09	0.09	.002	.004	0.07	0.046
Parallel[C]			0.04	0.03	0.12	0.16	0.08	0.39	0.14	0.12	0.19
OTFC vs Morphine[D]			0.43	0.50	0.88	0.58	0.92	0.85	0.81	0.84	0.96

A: Two way ANOVA at each time point with 4 treatment, 2 centers and treatment-by-center interaction as factors
B: F test of testing for equality of high dose group vs low dose group
C: F test of testing for equality of difference in high dose and difference in low dose treatments
D: F test for equality of OTFC and morphine groups

Pain intensity difference (PID) -

PID is calculated as difference in pain intensity score between the time scheduled and baseline. Sponsor performed comparisons (including among all treatments, among all centers, treatment-by-center interaction, OTFC vs. Morphine, high dose vs. Low dose) of mean PID at each single time point using two-way ANOVA. It is shown that the difference between OTFC and morphine is not statistically significant when in pooled or in dose stratified setting. The dose response is shown at 120 minutes and later with p-values of F-test being less than .05 at 180 and 240 minutes (See Attached Abbreviated Table of Sponsor's Figure 3).

Table IV.6 (Abbreviated Table of Sponsor's Figure 3) P-Values of the Comparisons of Mean Pain Intensity Difference - Evaluable Patients = 123

Test[a]	Time in minutes								
	15	30	45	60	120	180	240	300	360
Complete 2-way ANOVA At Every Time Point									
Treatment (T)	0.54	0.52	0.76	0.24	0.27	0.056	0.04	0.17	0.17
Center ©	0.07	0.17	0.34	0.07	0.28	0.30	0.50	0.81	0.72
T*C Interaction	0.27	0.78	0.93	0.68	0.88	0.70	0.53	0.68	0.46
High vs Low	0.71	0.33	0.35	0.11	0.13	0.007	0.009	0.06	.046
Parallel	0.25	0.33	0.65	0.35	0.23	0.60	0.29	0.23	0.33
OTFC vs Morphine	.45	0.63	0.85	0.33	0.63	0.80	0.52	0.76	0.66

Normalized weighted summed pain intensity difference (NWSPID) -
NWSPID is a normalized cumulative score of pain intensity that is calculated as

$NWSPID_i = 100xWSPID_{i-1}/MAX_i$

Where
t_i = i-th scheduled time in minutes;
PID_i = pain intensity difference at the i-th scheduled time;
$WSPID_i = WSPID_{i-1} + (PID_{i-1}+PID_i)(t_i - t_{i-1})/(2x60)$
$WSPID_0 = MAX_0 = 0$
$MAX_i = MAX_{i-1} + P_0(t_i - t_{i-1})/(2x60)$
P_i = pain intensity score at i-th scheduled time.

Sponsor performed comparisons (including among all treatments, among all centers, treatment-by-center interaction, OTFC vs. Morphine, high dose vs. low dose) of mean NWSPID at each single scheduled time using two-way ANOVA. It is shown that the difference between OTFC and morphine is not statistically significant when in pooled or in dose stratified setting. The dose response is shown at 120 minutes and later with p-values of F-test being less than .05 at 120, 180, 240, 300 and 360 minutes (Figure 4)(by center analyses were also given in Supplementary Figure 1).

Table IV.7 (Abbreviated Table of Sponsor's Figure 4) P-Values of the Comparisons of Mean Normalized Weighted Summed Pain Intensity Difference - Evaluable Patients = 123

Test[a]	Time in minutes								
	15	30	45	60	120	180	240	300	360
Complete Two-way ANOVA at each time point									
Treatment (T)	0.58	0.34	0.37	0.35	0127	0.11	0.09	0.09	0.11
Center ©	0.24	0.31	0.59	0.62	0.57	0.63	0.63	0.66	0.69
T*C Interaction	0.72	0.89	0.93	0.98	0.97	0.96	0.93	0.91	0.90
High vs Low	0.60	0.29	0.22	0.22	0.04	0.03	0.02	0.02	0.02
Parallel	0.41	0.38	0.37	0.19	0.20	0.33	0.45	0.48	0.51
OTFC vs Morphine	0.34	0.25	0.43	0.94	0.98	0.83	0.79	0.79	0.78

Pain relief score (PR Score) -
PR Score is a parameter that is not calculated from pain intensity scores. Sponsor performed comparisons (including among all treatments, among all centers, treatment-by-center interaction, OTFC vs. Morphine, high dose vs. low dose) of mean PR Score at each single scheduled time using two-way ANOVA. It is shown that there were overall difference among all treatments at

some scheduled time (p =.008 at 15 min, p=0.02 at 180 min). The difference was attributed to difference between high and low dose treatment at 180 and 240 minutes. It also showed that mean PR Score was greater in 10 mg morphine than in OTFC 800 µg but greater in OTFC 200 µg than in morphine 20 mg up to 60 minutes. The effects were significantly unparallel at all scheduled time before 60 minutes (Figure 5).

Table IV.8 (Abbreviated Table of Sponsor's Figure 5) P-Values of the Comparisons of Mean Pain Relief Score - Evaluable Patients = 123

Test*	Time in minutes								
	15	30	45	60	120	180	240	300	360
Complete Two-way ANOVA at each time point									
Treatment (T)	0.008	0.07	0.17	0.24	0.21	0.02	0.049	0.16	0.22
Center ©	0.34	0.66	0.46	0.32	0.39	0.11	0.55	0.67	0.67
T*C Interaction	0.87	0.94	0.85	0.73	0.88	0.46	0.40	0.27	0.20
High vs Low	0.86	0.94	0.76	0.38	0.07	0.003	0.01	0.08	0.08
Parallel	<0.001	0.01	0.04	0.06	0.30	0.53	0.21	0.16	0.24
OTFC vs Morphine	0.47	0.37	0.60	0.95	0.79	0.64	0.87	0.86	0.98

Weighted Total Pain Relief (WTOTPAR) -
WTOTPAR is a parameter that is calculated from PRS as follow

$WTOTPAR_i = WTOTPAR_{i-1} + (PRS_{i-1} + PRS_i)(t_i - t_{i-1})/(2x60)$
where,

$WTOTPAR_0 = PRS_0 = 0$

Sponsor performed comparisons (including among all treatments, among all centers, treatment-by-center interaction, OTFC vs. Morphine, high dose vs. low dose) of mean PRS at each single scheduled time using two-way ANOVA. It is shown that there were overall differences among all treatments at almost all scheduled times (all p-values <.05 expect at 120, 180, 240 and 360 minutes). The difference was attributed to differences between high and low dose treatment at 240, 300 and 360 minutes. It also showed that mean WTOTPAR was greater in 10 mg morphine than in OTFC 800 µg but greater in OTFC 200 µg than in morphine 20 mg. The effects were significantly unparallel at all scheduled time before 180 minutes (Figure 6).

Table IV.9 (Abbreviated Table of Sponsor's Figure 6) P-Values of the Comparisons of Mean Weighted Total Pain Relief - Evaluable Patients = 123

Test*	Time in minutes								
	15	30	45	60	120	180	240	300	360
Complete Two-way ANOVA at each time point									
Treatment (T)	0.008	0.011	0.02	0.02	0.08	0082	0.054	0.049	0.06
Center ©	0.34	0.50	0.60	0.32	0.34	0.27	0.26	0.34	0.42
T*C Interaction	0.87	0.92	0.95	0.90	0.89	0.84	0.75	0.65	0.55
High vs Low	0.86	0.88	0.86	0.85	0.30	0.083	0.03	0.03	0.03
Parallel	<0.001	0.002	0.003	0.002	0.02	0.06	0.30	0.10	0.11
OTFC vs Morphine	0.47	0.41	0.41	0.66	0.78	0.74		0.81	0.83

Table IV.10 (Abbreviated Table of Sponsor's Figure 7) P-Values of the Comparisons of Mean Weighted Summed Pain Intensity Difference - Evaluable Patients = 123

Test*	Time in minutes								
	15	30	45	60	120	180	240	300	360
Complete Two-way ANOVA at each time point									
Treatment (T)	0.54	0.49	0.56	0.34	0.26	019	0.12	0.10	0.10
Center ©	0.07	0.07	0.12	0.15	0.12	0.16	0.20	0.27	0.34
T*C Interaction	0.27	0.53	0.74	0.82	0.80	0.85	0.79	0.79	0.77
High vs Low	0.71	0.51	0.40	0.51	0.19	0.04	0.03	0.03	0.03
Parallel	0.25	0.24	0.32	0.10	0.16	0.26	0.25	0.25	0.24
OTFC vs Morphine	0.45	0.49	0.65	0.65	0.51	0.59	0.59	0.59	0.60

IV.7.3 Safety Analysis

Respiratory Effects and Vital signs -

There was no apparent difference in the number of ptients with events of respiratory depression, of clinical significant changes in vital signs or significant drop in oxygen saturation level among the four treatments.

Adverse Events -

The most common adverse events were nausea, pruritus, and dizziness. The percentage of OTFC patients with the adverse events were not higher the paired morphine groups.

IV.8 Reviewer's evaluation

The reviewer's comments will be focussed on the issues on the goal of study, the primary and multiple endpoints. the repeated measurements of each endpoint, and the sample size and power of the study.

a. There are six highly correlated efficacy endpoints presented in the report with no indication which one is the primary endpoints that will be use for a 'make or break' decision and for primary labeling. With multiple endpoints, to say that two treatments are equivalent often requires that they are equivalent with respect to each of the endpoints in order to maintain the overall type I error for equivalency. Obviously, the power of the equivalence test will be low when equivalence criteria at multiple endpoints is required. Since most of the endpoints were calculated from pain intensity score and pain relief score, one would choose the first two endpoints as primary endpoints without adjustment for multiple endpoints and consider the rest as secondary.

b. Appropriateness of the study design for dose equivalence - Clinical trial with active control is often used for one of the three types of objectives. For the 'superiority' claim, the the trial is designed to test

H_{o1}: Exp(treatment-control) $\leq$ 0 versus H_{a1}: Exp(treatment-control) > 0.

The 'superiority' can be established if H_{01} is rejected. For the 'not inferior to' claim, the experimenter needs to define an equivalence limit $-\Delta$, which is often a constant determined by the variation of efficacy of active control. Sample size and power of the trial is determined for the testing of

H_{02}: Exp(treatment-control) $\leq -\Delta$ versus H_{a2}: Exp(treatment-control) > $-\Delta$.

The 'not inferior to' claim can be established if H_{02} is rejected. For the 'straight equivalence' claim, the trial is design to test for

H_{03}: |Exp(treatment) - Exp(control)| $\geq \Delta$ versus H_{a3}: |Exp(treatment) - Exp(control)| < Δ.

The 'straight equivalence' claim can be established when H_{03} is rejected. Hence the determinations of Δ and sample size are crucial for either the 'not inferior to' or the 'straight equivalence' claims. The claims can only be made when the data are collected and an appropriate null hypothesis is rejected, but can not be made directly by not able to reject the null hypothesis (ref - Blackwelder, CW, Proving the null hypothesis in clinical trial, ***Controlled Clinical Trials***. 3:345-353, 1982) . An indirect way to support the 'not inferior to' or the 'straight equivalence' claim can be make by not rejecting H_{o1} with the sample size appropriately determined to have enough power to detect a difference of Δ, the pre-specified equivalence limit. Conclusion of 'not inferior to' or 'straight equivalence' based on not rejecting H_{o1} is biased

toward the claim, when Δ is not specified before the trial or sample size is not determined appropriately (Ng, TH. Conventional null hypothesis testing in active control equivalence studies. ***Controlled Clinical Trials***, 16:356-358, 1995).

As stated in the protocol, the goal of the study is to establish the relative potency by showing the therapeutical equivalence of the treatment and active control paired at both high and low dose strength. However, the study was not designed with a pre-specified Δ and the sample size was not determined accordingly for the 'straight equivalence' testing.

c. As pointed out in the comments on appropriateness of study design for equivalence testing, the analysis performed by the sponsor and the conclusion of dose equivalence are therefore biased toward the 'equivalence' claim. The reviewer re-analyzed the data by calculating the estimate of the difference in pain intensity score and pain relief score, the standard deviation of the difference and the approximate 95% confidence limits of the difference between OTFC and Morphine in both 100 μg OTFC versus 2 mg morphine and 800 μg OTFC versus 10 mg morphine (Bauer, P, Kieser, M, A unifying approach for confidence intervals and testing of equivalence and difference, ***Biometrics***, 83:4:914-937, 1996) . The limits are to be compared with the equivalence limit Δ, for assessing 'straight equivalence'. The equivalence can be claimed if the lower limit is larger than -Δ and the upper limits is smaller than Δ at all the crucial time points. A 'not inferior to' claim can be made when the lower limit is larger than -Δ. Since Δ was not pre-specified by sponsor and can neither be specified by statistical reviewer, the reviewer will use the very liberal 20% of the mean measurement of control group as the equivalence limit in an arbitrary way for reference purpose only. The 30% of mean measurements of placebo, a limit which is not acceptable by FDA, are also calculated and given as additional reference points.

It is clear from this re-analysis that the standard errors of the difference are often large and that the differences of either one of the two endpoints have magnitude of either the lower or upper 95% confidence limits of the difference be larger than 20% of the mean of the morphine responses at most time points. For example, for pain intensity score, straight equivalence claim of 200 μg OTFC and 2 mg morphine or 800 μg OTFC and 10mg morphine fails at every time point when an equivalence limit of 20% of mean measurements of morphine group is used. Instead, if a 30% of the mean measurement value of morphine group is used as the equivalence limit, 200 μg and 2 mg morphine are 'straightly equivalent' at 300 and 360 minute time points only. For pain relief score, neither 200 μg OTFC and 2mg morphine nor 800 μg OTFC and 10 mg morphine are 'straightly equivalent' at any time point, using the 20% limit.

On the other hand, 200 μg OTFC is 'not inferior to' 2mg morphine (which means reject H_{03} : Exp(OTFC-morphine) ≤ 20% Exp(morphine)) in pain intensity score at all time points except at 180, 300 and 360 minute time points when the 20% limit applied. It is 'not inferior to' 2 mg morphine at all time points except 180 minute when the 30% limit applied. For the high dose comparison, 800 μg is 'not superior to' 10mg morphine (which mean reject H_{03}: Exp(response of 800 μg - response of 10 mg morphine) ≥ 20% Exp(response of 10mg morphine)), in pain

intensity score at 15, 30, 120, 300 and 360 minute time points when the 20% limit applied. Or the claim holds at all time points except at 60 minute time point when 30% limit applied. For pain relief score, the claim that 200 μg OTFC is 'no inferior to' 2mg morphine does not hold at any time point when a 20% limit applied. Or, the claim holds at 30, 45, 240 and 300 minute time points when a 20% limit applied.

Hence based on reviewer's analysis, this study does not have adequate support for the 'straight equivalence' claim in either efficacy endpoint using the 20% reference limits.

d. It seems that there is an alternative procedure for the dose ratio calibration under the same assumption of dose response relationship in morphine and in OTFC. For example, if there is a linear dose response relationship in morphine dose and in OTFC dose with PIS. Let (L_{low}, U_{Low}) and (L_{High}, U_{High}) be the 95% confidence interval of PIS for the low and high dose of OTFC respectively. One may estimate the dose response line and its 95% confidence bands of morphine. Then find the dose level from the confidence bands of the morphine dose response line corresponding to L_{low}, U_{Low}, L_{High} and U_{High}. Say, the four levels are L'_{Low}, L'_{High}, U'_{Low} and U'_{High}. One may claim that the equivalent dose for low OTFC dose is between L'_{Low} and L'_{High} and for high OTFC dose is between U'_{Low} and U'_{High}. However, it requires more than two doses of morphine to establish a dose response relationship.

Table VI.11. Treatment difference, Standard Errors of Difference and the 95% Confidence Limits of the Difference, and 20% of control response - Evaluable Patients =123

Comparison	Time in minutes								
	15	30	45	60	120	180	240	300	360
Pain Intensity Score									
200 µg OTFC - 2 mg Morphine									
Difference	6	7	8	10	8	-1	5	1	1
Standard Error	5	5.8	5.7	5.7	7.1	7.1	7.1	7.1	7.1
Lower 95% limit	-3.8[a]	-2.8	3.2	-1.2	-5.9	-14.9	-8.9	-12.9	-12.9
Upper 95% Limit	17.8[b]	16.8	19.2	21.2	21.9	12.9	18.9	14.9	14.9
20% of Morphine Mean	4	5.4	5.4	5.8	8	9.2	10.8	10.8	10.8
'Straight Equivalence'	no	no	no	no	no	no	no	no	no
'No Inferior to Control'	yes	yes	yes	yes	yes	no	yes	no	no
30% of Morphine Mean	6	8.1	8.1	8.4	12	13.8	16.2	16.2	16.2
'Straight Equivalence'	no	no	no	n0	no	no	no	yes	yes
'No Inferior to Control'	yes	yes	yes	yes	yes	no	yes	yes	yes
800 µg OTFC - 10 mg Morphine									
Difference	-13	-10	-6	-4	-7	-4	-4	-8	-7
Standard Error	5.7	5	5	4.2	9.8	5.7	6.4	6.4	5.7
Lower 95% limit	-23.1	-19.8	-15.8	-15.3	-18.4	-15.4	-16.6	-20.6	-18.1
Upper 95% Limit	-2.1	-0.2	3.8	4.3	4.4	7.4	8.6	4.6	4.1
20% of Morphine Mean	3.6	3.0	3.2	2.8	4.4	6.0	7.0	7.8	7.8
'Straight Equivalence'	no	no	no	no	no	no	no	no	no
'no superior to'	yes	yes	no	no	yes	no	no	yes	yes
30% of Morphine Mean	5.4	4.5	4.8	4.2	6.6	9.0	10.5	11.7	11.7
'Straight Equivalence'	no	no	no	no	no	no	no	no	no
'no superior to'	yes	yes	yes	no	yes	yes	yes	yes	yes

Pain Relief Score 200 µg OTFC - 2 mg Morphine									
Difference	-11	-9	-8	-11	-3	4	-2	-1	0
Standard Error	6.4	5.7	5.7	6.4	7.8	7.8	7.4	7.1	7.1
Lower 95% limit	-23.6	-20.1	-19.1	-23.6	-18.3	-11.3	-15.9	-14.9	-13.9
Upper 95% Limit	1.6	2.1	3.1	1.6	12.3	19.3	11.9	12.9	13.9
20% of Morphine Mean	12.4	13.4	13.8	13.2	11.6	9.8	8.4	11.4	8.2
'Straight Equivalence'	no	no	no	no	no	no	no	no	no
'no inferior to'	no	no	no	no	no	no	no	no	no
30% of Morphine Mean	18.6	20.1	20.7	19.8	17.4	14.7	16.1	17.1	12.3
'Straight Equivalence'	no	yes	yes	no	no	no	yes	yes	no
'no inferior to'	no	yes	yes	no	no	no	yes	yes	no
800 µg OTFC - 10 mg Morphine									
Difference	21	16	13	8	7	7	7	10	7
Standard Error	6.4	6.4	6.7	6.4	6.4	7.2	7.2	6.4	6.4
Lower 95% limit	8.4	3.4	-0.1	-4.5	-5.5	-7.1	-7.1	2.6	5.6
Upper 95% Limit	33.6	28.6	26.1	20.5	19.5	21.1	21.1	22.6	19.6
20% of Morphine Mean	15.4	16.0	16.4	16.2	14.4	13.2	12.0	11.0	11.0
'Straight Equivalence'	no	no	no	no	no	no	no	no	no
'no superior to'	no	no	no	no	no	no	no	no	no
30% of Morphine Mean	23.1	24	24.6	24.3	21.6	19.8	18.0	16.5	16.5
'Straight Equivalence'	no	no	no	yes	yes	no	no	no	no
'no superior to'	no	no	no	yes	yes	no	no	no	no

a: Lower 95% Confidence Limit of the Difference based on Normal Approximation
b: Upper 95% Confidence Limit of the Difference Based on Normal Approximation.
*: Unacceptable by FDA

The treatment difference, standard errors of difference, the 95% confidence limits and 20% limits are also calculated for the other endpoints and given in Table IV.12.