STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 022-450
Drug Name:	Ofirmev (acetaminophen injection)
Indication(s):	Management of mild to moderate pain, management of moderate to severe pain with adjunctive opioid analgesics, and reduction of fever
Applicant:	Mallinckrodt Inc.
Date(s):	Received: April 29, 2016
	PDUFA: January 27, 2017
Review Priority:	Priority
Biometrics Division:	II
Statistical Reviewer:	Feng Li, Ph.D.
Concurring Reviewers:	David Petullo, M.S.
Medical Division:	Division of Anesthesia, Analgesia, and Addiction Products
Clinical Team:	Medical Officer: Christina Fang, M.D.
	Medical Team Leader: Joshua Lloyd, M.D.
Project Manager:	Shelly Kapoor
Keywords:	NDA review, Pediatric Written Request (PWR), Clinical Studies

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

Mallinckrodt Inc. submitted the results from a pediatric study (CPI-APA-353) that evaluated Ofirmev for the treatment of acute pain. This study was conducted to fulfill the Pediatric Written Request (PWR) #3 that was issued on March 30, 2015. Ofirmev is an intravenous (IV) formulation of acetaminophen (APAP) that is currently approved for the management of mild to moderate pain, the management of moderate to severe pain with adjunctive opioid analgesics, and for the reduction of fever in both adults and children 2 years of age and older.

Study CPI-APA-353 was a randomized, double-blind, placebo controlled, parallel, multiple-dose, multicenter study that was conducted to show that IV APAP was superior to placebo for the treatment of acute pain in inpatient pediatric patients less than 2 years old. The primary efficacy variable was the total rescue opioid consumption during the 24 hours of treatment.

In my opinion, the study conducted by the applicant has met the requirement of the PWR from statistical perspective. However, based on the primary efficacy endpoint, Study CPI-APA-353 failed to demonstrate that IV APAP was superior to placebo in treating acute pain in pediatric patients less than 2 years old.

2. INTRODUCTION

2.1 Overview

In addition to efficacy, the PWR required the applicant assess the pharmacokinetic (PK) and pharmacodynamics (PD) relationship of IV APAP.

The PWR required that the study must include a sufficient number of patients to produce a sample size adequately powered (at least 80%) for detecting differences based on estimates of the effect size of the primary efficacy endpoint. The number of patients must be approximately evenly distributed between genders, approximately evenly distributed across the age ranges, and reasonably distributed within the age ranges. These groups have been determined by assessment of differences in developmental physiology.

The PWR also specified that analgesic effects (including analgesic duration) must be studied. It is essential to identify a single primary efficacy outcome reflecting adequacy of analgesia. Clinical assessments will be made using validated, age-appropriate instruments. Inter-rater variability will be evaluated. Evaluation will include assessment by blinded caretakers and assessors. Rationale for choice of scale will be provided in the protocol and must be agreed upon by the Division of Anesthesia, Analgesia and Addiction Products.

My statistical review will focus on whether data from Study CPI-APA-353 demonstrated the efficacy of IV APAP in [redacted] (b) (4) pediatrics.

2.2 Data Sources

All data were supplied electronically by the applicant as SAS transport files and can be found at the following location in the CDER electronic document room (EDR):\\Cdsesub1\evsprod\NDA022450\0091\m5\datasets\cpi-apa-353.

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The datasets and associated define files were of acceptable quality, and were sufficient for validating study results. However, the applicant did not submit subgroup efficacy results for race in the original submission. The applicant subsequently submitted these subgroup analyses in response to the division's information request.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Study CPI-APA-353 was a randomized, double-blind, placebo-controlled, parallel group, multiple-dose, multicenter study that evaluated the efficacy, PK and PD of IV APAP for the treatment of acute pain in inpatient pediatric patients less than 2 years old. The study enrolled patients who had undergone surgery or had a medical situation where an IV analgesic regimen would be needed for management of pain for at least 24 hours. There was an expectation that IV treatment would be required for the duration of the study.

In line with the PWR, patients in the following age ranges were enrolled:

- Neonates: birth (≥ 28 weeks to ≤ 40 weeks gestational age at birth) to ≤ 28 days chronological age at randomization
- Younger infants: ≥ 29 days to < 6 months of age at randomization
- Intermediate age infants: ≥ 6 months to < 12 months of age at randomization
- Older infants: ≥ 12 months to < 24 months of age at randomization.

Eligible patients were randomly assigned to one of the two IV APAP dosing groups (treatment groups A or B) or the two matched placebo (treatment groups C or D) in a 2:2:1:1 ratio. All patients received their assigned treatment plus standard of care (SOC) rescue opioids. Randomization was stratified by age group and the total opioid consumption over a 6-hour pre-

randomization qualification period. A patient's pre-randomization consumption was classified as either low or high based upon a threshold average of 30 μg/kg/h IV morphine equivalent. Patients were scheduled to receive four doses of study drug over a 24-hour period (a single dose every 6 hours) by continuous IV infusion over a period of 15 minutes. Details of the dosing regimen can be found in the Appendix.

Following randomization, all patients had a pain assessment and immediately thereafter received a single standard bolus dose of protocol defined opioid (Baseline Opioid Dose) within 30 minutes of the infusion start time (T0) and within one hour after randomization. It was up to investigators to decide what opioid dose (full loading dose or lesser amount) was appropriate for any given subject based upon the subject's pain score. Pain intensity was additionally assessed at 0.5, 1, 2, 3, 4, 6, 12, 18, 24 hours after T0 and prior to each dose of rescue medication as well as at early termination. The Leuven Neonatal Pain Scale (LNPS) with a maximum score of 14 was used for assessing pain intensity in neonates and younger infants. The Face, Leg, Activity, Cry, and Consolability (FLACC) scale with a maximum score of 10 was used for assessing pain intensity in intermediate age and older infants. Rescue medication was mandatory any time the assessed pain intensity was severe (LNPS or FLACC score ≥ 6).

Sedation was assessed using the University of Michigan Sedation Scale at T0, 3, 6, 9, 12 hours after T0, and prior to administration of the first dose of rescue opioid. Global evaluation of satisfaction with study treatments was assessed at the end of the sturdy treatment period.

The primary efficacy variable was the total rescue opioid consumption during the 24 hours after T0.

3.2.2 Statistical Methodologies

The primary efficacy variable was analyzed using an analysis of variance (ANOVA) model with treatment and age group as factors. The primary analysis compared each IV APAP group (A or B) to the combined placebo control groups (C and D) using a two-sided test at 2.5% level of significance to control the overall type-I error at the two-sided level of 0.05. Efficacy analyses was carried out using the modified intent-to-treat (mITT) population, defined as all randomized patients who received at least one completed dose of study medication and had at least one pain assessment after T0.

The total rescue opioid consumption was calculated using available data regardless of study treatment discontinuation. Sensitivity analyses for the primary efficacy variable were conducted by excluding early withdrawn patients and patients who had no rescue medication after the first dose of study treatment.

Pain intensity scores following administration of rescue medication were replaced with the most recent non-missing pain score in the analysis of summed pain intensity difference (SPID) variables.

Time to first rescue medication was analyzed using a stratified logrank test. Categorical variables such as global evaluation of satisfaction were assessed using the Cochran-Mantel-Haenszel (CMH) test stratified by age group.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 215 patients were randomized, 138 to the IV APAP groups (A and B), and 77 to the control treatment groups (C and D). There were 16 (7%) patients who discontinued before receiving randomized treatments. In addition, one placebo patient completed the study but did not provide any pain assessment after T0. The applicant-defined mITT population included 197 (92%) patients. Overall, IV APAP groups had a higher rate of study completion than the placebo control groups (Table 1).

The demographic and baseline characteristics were comparable across treatment groups (Table 2). The majority of the patients were male and white. The number of patients was approximately evenly distributed across the age ranges. The median weight was 7 kg for all treatment groups.

Table 1: Patient Disposition

	IV APAP			Control		
Population	**A**	**B**	**A+B**	**C**	**D**	**C+D**
All randomized	**N=66**	**N=72**	**N=138**	**N=35**	**N=42**	**N=77**
mITT population	61 (92%)	67 (93%)	128 (93%)	31 (89%)	38 (90%)	69 (90%)
Completed, n (%)*	52 (79%)	55 (76%)	107 (78%)	26 (74%)	26 (62%)	52 (68%)
Discontinued, n(%)*	14 (21%)	17 (24%)	31 (22%)	9 (26%)	16 (38%)	25 (32%)
Adverse event	0	2 (3%)	2 (1%)	3 (9%)	6 (14%)	9 (12%)
Non-compliance	1 (2%)	0	1 (1%)	0	0	0
Other	8 12%)	9 (13%)	17 (12%)	3 (9%)	6 14%)	9 (12%)
Physician decision	1 (2%)	2 (3%)	3 (2%)	2 (6%)	4 (10%)	6 (8%)
Withdrawal by subject	4 (6%)	4 (6%)	8 (6%)	1 (3%)	0	1 (1%)

Source: Clinical study report, Table 14.1.1.1 and Table 14.1.1.2

*: Percentages are based on the total number of randomized patients.

Table 2: Summary of Demographics and Baseline Characteristics (mITT population)

	IV APA			Control
	A (N=61)	**B (N=67)**	**A+B (N=128)**	**C+D (N=69)**
Age (days)				
Mean (SD)	223 (175)	246 (209)	235 (193)	220 (175)
Median	202	187	199	194
Min, Max	2, 643	3, 725	2, 725	1, 664
Age group, n (%)				
Neonates	13 (21%)	12 (18%)	25 (20%)	13 (19%)
Younger Infants	16 (26%)	20 (30%)	36 (28%)	18 (26%)
Intermediate Age Infants	18 (30%)	17 (25%)	35 (27%)	20 (29%)
Older Infants	14 (23%)	18 (27%)	32 (25%)	18 (26%)
Gender, n (%)				
Male	44 (72%)	40 (60%)	84 (66%)	43 (62%)
Female	17 (28%)	27 (40%)	44 (34%)	26 (38%)
Race, n (%)				
White	40 (66%)	49 (73%)	89 (70%)	46 (67%)
Black or African American	12 (20%)	9 (13%)	21 (16%)	9 (13%)
American Indian or Alaska Native	1 (2%)	0	1 (1%)	0
Asian	3 (5%)	4 (6%)	7 (5%)	6 (9%)
Other	4 (7%)	4 (6%)	8 (6%)	5 (7%)
Missing	1 (2%)	1 (2%)	2 (2%)	3 (4%)
Weight at screening (kg)				
Mean (SD)	7 (3)	7 (3)	7 (3)	7 (3)
Median	7	7	7	7
Min, Max	3, 14	1, 14	1, 14	1, 12
Baseline opioid dose (µg/kg)				
Mean (SD)	46 (36)	50 (36)	48 (36)	47 (32)
Median	48	50	50	50
Min, Max	0, 170	0, 205	0, 205	0, 116

Source: Clinical study report, Table 14.1.4; SD: standard deviation

3.2.4 Results and Conclusions

Results from the analysis of total rescue opioid consumption are presented in Table 3. The primary analysis failed to demonstrate a statistically significant reduction in total rescue opioid use in patients receiving IV APAP when compared to patients receiving placebo. There were no statistically significant differences in the total rescue opioid consumption between IV APAP groups A, B, or in combination (A+B), and the combined control treatment group (C+D). The total rescue opioid consumptions by dosing interval were also similar between IV APAP and the control groups (Appendix). These results are consistent with the applicant's study report.

The total rescue opioid consumption was calculated using available data regardless of study treatment discontinuation. There were no missing values in the mITT population. However, it is unclear whether the amount of rescue opioid use was still recorded after early treatment discontinuation. If not recorded, the calculated total amount would underestimate the actual

overall opioid consumption for patients who discontinued treatment early. A sensitivity analysis that excluded patients who discontinued treatment early was conducted to evaluate the potential impact of dropouts on comparison of the total opioid consumption. The results were in favor of IV APAP treatment numerically (Table 4). However, another sensitivity analysis excluding patients who requested no rescue after receiving IV APAP or placebo infusion was numerically in favor of placebo.

For all treatment groups, the average pain intensity score was maintained approximately at 3 after administration of the study treatment. The observed average pain scores of patients receiving IV APAP were similar to those of patients receiving placebo during the 24 hours treatment period, as depicted in Figure 1 (neonates and younger infants) and Figure 2 (intermediate age and older infants). There were some notable differences in average pain intensity among treatment groups prior to the administration of the randomized treatments, which was not concerning here as the differences became much smaller after all patients were administered the standard bolus does of opioid before receiving the randomized treatments and the primary endpoint counted the total opioid consumption during the 24 hours after receiving the study infusion.

There were also no statistically significant differences between IV APAP groups and the combined placebo control group in secondary endpoints such as global evaluation and time to first rescue medication.

Table 3: Summary and Analysis of Total Rescue Opioid Consumption

	IV APAP			**Control**
	A (N=61)	**B (N=67)**	**A+B (N=128)**	**C+D (N=69)**
Total Opioid consumption (µg/kg)				
Mean	167	180	174	180
SD	225	193	208	185
Median	91	127	106	132
ANOVA analysis				
LS means	163	177	170	175
Difference vs combined control	-12	2	-5	
97 5% CI	(-91, 67)	(-76, 79)	(-72, 62)	
P-value#	0.74	0.97	0.87	

Source: Clinical Study Report, Table 14.2.1.1.1;
#P-values are nominal without multiplicity adjustment.

Table 4: Sensitivity Analyses of Total Rescue Opioid Consumption

	IV APAP			Control
ANOVA	**A (N=61)**	**B (N=67)**	**A+B (N=128)**	**C+D (N=69)**
Excluding patients who discontinued treatment early				
n	52	55	107	51
LS means	183	191	187	209
Difference vs combined control	-25	-18	-22	
97 5% CI	(-118, 67)	(-110,73)	(-101, 58)	
P-value#	0.54	0.65	0.54	
Excluding patients who received no rescue after first dose				
n	47	53	100	58
LS means	214	224	219	208
Difference vs combined control	6	16	11	
97 5% CI	(-84, 96)	(-71, 103)	(-64, 87)	
P-value	0.89	0.68	0.74	

Source: Clinical Study Report, Table 14.2.1.1.5;
#P-values are nominal without multiplicity adjustment.

Figure 1: **Average Pain Intensity over Time (Neonates and Younger Infants)**

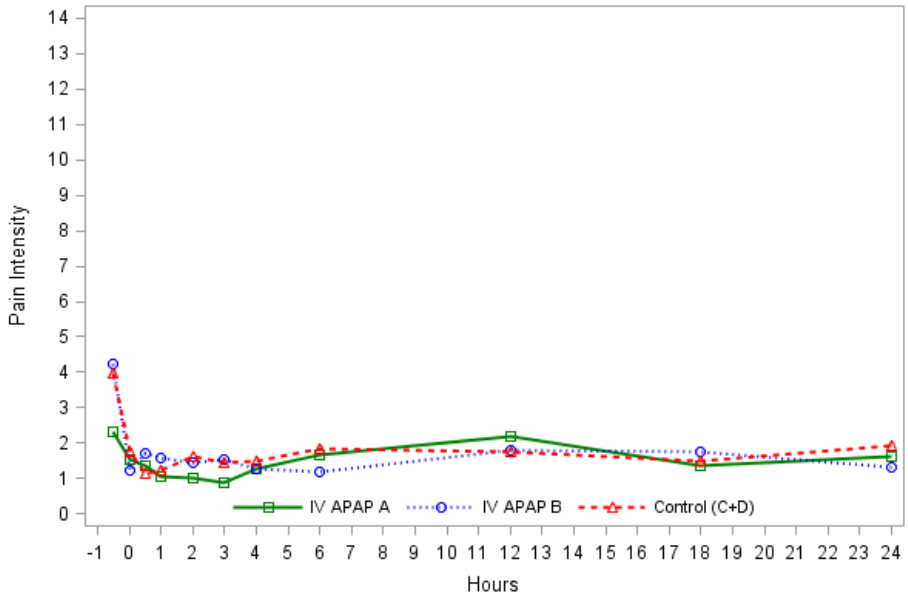
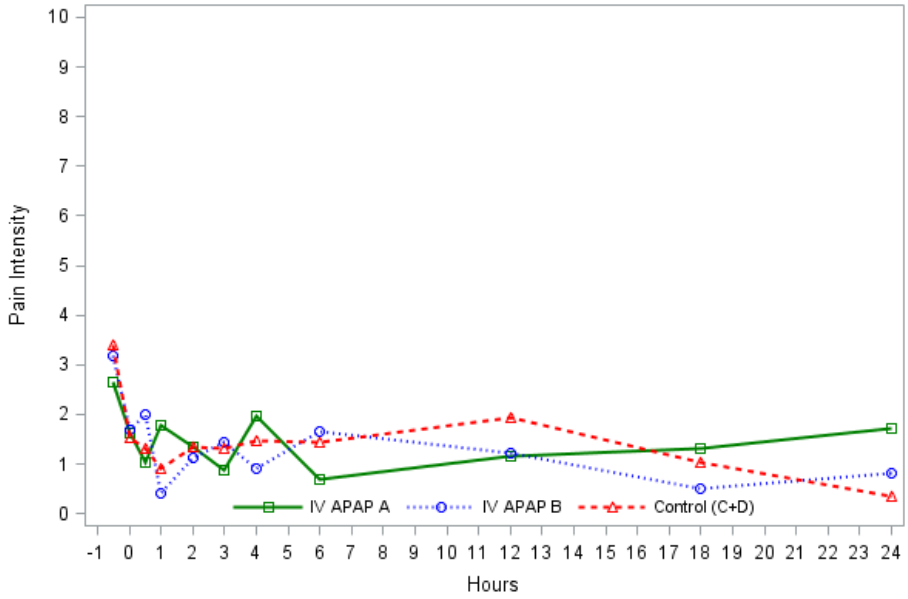


Figure 2: Average Pain Intensity over Time (Intermediate and older infants)



3.3 Evaluation of Safety

The evaluation of the safety data was conducted by the clinical reviewer, Dr. Christina Fang. There were no major safety findings. Please refer to Dr. Fang's review for detailed information regarding the adverse event profile.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Age and Race

Subgroups analyses by gender and race are presented in Table 5. Female patients receiving IV APAP numerically consumed less rescue opioid than those receiving placebo. In contrast, male patients receiving IV APAP used more rescue opioid than those receiving placebo. Neither the treatment effect for female patients nor the interaction of treatment by gender was statistically significant.

The applicant submitted the subgroup analyses for race in response to the division's information request. The "other" race category included the races other than white and black. Regardless of race, there was no evidence that IV APAP was superior to placebo (Table 5).

Analyses of opioid consumption by age group revealed that neonates and intermediate age infants treated with IV APAP requested less rescue opioid than those treated with placebo (Table 6). However, for the younger and older infants age groups, patients treated with IV APAP

requested more rescue medication. However, none of the treatment differences were statistically significant.

Table 5: Summary of Total Opioid Consumption by Sex and Race

Subgroup			A (N=61)	B (N=67)	A+B (N=128)	C+D (N=69)
Gender	Male	n	44	40	84	43
		LS means*	182	195	188	172
		Difference vs control	10	23	16	
	Female	n	17	27	44	26
		LS means	121	154	141	178
		Difference vs control	-56	-25	-37	
Race	White	n	40	49	89	46
		LS means	145	185	167	183
		Difference vs control	-38	2	-16	
	Black	n	12	9	21	9
		LS means	254	189	227	215
		Difference vs control	39	-26	12	
	Other	n	9	9	18	14
		LS means	177	152	167	170
		Difference vs control	7	-18	-3	

Source: Clinical Study Report, Table 14.2.1.1.2 and Efficacy Information Amendment, Table 14.1.2.1.1.7
*LS means are based on ANOVA analysis with treatment and age group as factor.

Table 6: Summary of Total Opioid Consumption by Age Group

Age group		A (N=61)	B (N=67)	A+B (N=128)	C+D (N=69)
Neonates	n	13	12	25	12
	means	60	144	101	165
	Difference vs control	-105	-21	-64	
	97.5% CI	(-253, 43)	(-172, 130)	(-195, 66)	
	P-value#	0.11	0.75	0.25	
Younger infants	n	16	20	36	18
	means	193	169	180	110
	Difference vs control	83	59	70	
	97.5% CI	(-40, 205)	(-57, 175)	(-33, 172)	
	P-value	0.13	0.24	0.12	
Intermediate age infants	n	18	17	35	20
	means	211	177	195	265
	Difference vs control	-54	-88	-70	
	97.5% CI	(-244, 135)	(-281, 105)	(-233, 92)	
	P-value	0.51	0.30	0.32	
Older infants	n	14	18	32	18
	means	180	218	201	167
	Difference vs control	13	51	34	
	97.5% CI	(-153, 179)	(-104, 206)	(-101,171)	
	P-value	0.85	0.45	0.56	

Source: Clinical Study Report, Table 14.2.1.1.3; #: P-values are nominal without multiplicity adjustment.

4.2 Other Special/Subgroup Populations

It was found that about 58% of the patients in the combined IV APAP group and 65% of the patients in the combined placebo group (Table 7) had quantifiable concentrations of APAP prior to the initial dose of study drug. The APAP concentration was either zero or missing in other patients. To examine the potential impact of this finding, I examined the primary efficacy endpoint, amount of rescue opioid use, in patients with detectable APAP concentrations at baseline and those without. For those patients without a detectable APAP concentration at baseline, the low dose IV APAP group (A) consumed numerically less rescue opioid than the combined placebo group. In contrast, the high dose IV APAP group (B) consumed more rescue than the combined placebo group. For patients who had non-zero baseline APAP concentrations, the findings were opposite: the high dose IV APAP group requested less rescue opioids whereas the low dose IV APAP group requested more on average in comparison with the combined the placebo group. These findings did not indicate that the non-zero baseline APAP concentration in some patients was the primary cause of the failure of the primary efficacy analysis.

Table 7: Summary of Total Opioid Consumption by Baseline APAP

		IV APAP			Control
Baseline APAP		**A (N=61)**	**B (N=67)**	**A+B (N=128)**	**C+D (N=69)**
Missing	n	9 (15%)	3 (5%)	12 (9%)	2 (3%)
0	n	15 (25%)	27 (40%)	42 (33%)	22 (32%)
	LS Mean*	166	196	186	191
	Difference	-25	5	-5	
	P-value#	0.7	0.9	0.9	
>0	N	37 (61%)	37 (55%)	74(58%)	45 (65%)
	LS Mean	185	163	174	173
	Difference	12	-10	1	
	P-value	0.8	0.8	0.97	

*LS means are based on ANOVA analysis with treatment and age group as factor.
#P-values are nominal without multiplicity adjustment.
Source: Reviewer

5. SUMMARY AND CONCLUSIONS

Statistical Issues

Some statistical issues were identified in the applicant's efficacy analyses.

First, it is unclear whether the amount of rescue opioid use was recorded for a patient that discontinued study treatment prior to the end of the study. The total opioid consumption for these patients may have been underestimated. A sensitivity analysis using completers only yielded results that were numerically in favor of IV APAP.

Second, the bolus opioid administration within 30 minutes of the start of the study infusion treatment might have confounded the treatment effect. After the administration of the bolus opioid, patients might not have sufficient pain to demonstrate the efficacy IV APAP.

5.2 Collective Evidence

Overall, the observed data failed to demonstrate the efficacy of IV APAP in management of acute pain in pediatric patients under 2 years old. Neither the primary efficacy endpoint nor the secondary endpoints achieved statistical significance.

5.3 Conclusions and Recommendations

Study CPI-APA-353 was an efficacy study conducted to fulfill the requirements specified in the PWR. (b) (4) I think the applicant has met the requirements specified in the PWR from statistical perspective.

5.4 Labeling Recommendations

The applicant did not include any efficacy results from Study CPI-APA-353. I think the applicant should include some study information in Section 8.4 and state that effectiveness of IV APAP in treating acute pain has not been established in pediatric patients under 2 years old.

Appendix

Table 8: IV APAP Dosing Group A

Age Category (Age at Randomization)	Volume of IV APAP (10 mg/mL) to be Dispensed	IV APAP Dose Regimen
Extreme pre-term neonates (≥ 28 to < 32 weeks gestational age) or any neonate with bilirubin > 1.5 X ULN	0.75 mL/kg	7.5 mg/kg IV over 15 minutes, given Q6h X 4 doses
Pre-term neonates (≥ 32 to < 37 weeks gestational age) with normal bilirubin levels	1.0 mL/kg	10 mg/kg IV over 15 minutes, given Q6h X 4 doses
Full-term neonates(≥ 37 weeks to ≤ 40 weeks gestational age) with normal bilirubin levels	1.0 mL/kg	10 mg/kg IV over 15 minutes, given Q6h X 4 doses
All infants (≥ 29 days to < 24 months of age)	1.25 mL/kg	12.5 mg/kg IV over 15 minutes, given Q6h X 4 doses

Q6h = Every 6 hours.

Table 9: IV APAP Dosing Group B

Age Category (Age at Randomization)	Volume of IV APAP (10 mg/mL) to be Dispensed	IV APAP Dose Regimen
Extreme pre-term neonates(≥ 28 to < 32 weeks gestational age) or any neonate with bilirubin > 1.5 X ULN	1.0 mL/kg	10 mg/kg IV over 15 minutes, given Q6h X 4 doses
Pre-term neonates (≥ 32 to < 37 weeks gestational age) with normal bilirubin levels	1.25 mL/kg	12.5 mg/kg IV over 15 minutes, given Q6h X 4 doses
Full-term neonates (≥ 37 weeks to ≤ 40 weeks gestational age) with normal bilirubin levels	1.25 mL/kg	12.5 mg/kg IV over 15 minutes, given Q6h X 4 doses
All infants (≥ 29 days to < 24 months of age)	1.5 mL/kg	15 mg/kg IV over 15 minutes, given Q6h X 4 doses

Q6h = Every 6 hours.

Table 10: Matching Placebo Groups C and D

Age Category (Age at Randomization)	Volume of Normal Saline for Control Group C (Matching Placebo for IV APAP Treatment Group A)	Volume of Normal Saline for Control Group D (Matching Placebo for IV APAP Treatment Group B)
Extreme pre-term neonates (≥ 28 to < 32 weeks gestational age) or any neonate with bilirubin > 1.5 X ULN	0.75 mL/kg	1.0 mL/kg
Pre-term neonates (≥ 32 to < 37 weeks gestational age) with normal bilirubin levels	1.0 mL/kg	1.25 mL/kg
Full-term neonates (≥ 37 weeks to ≤ 40 weeks gestational age) with normal bilirubin levels	1.0 mL/kg	1.25 mL/kg
All infants (≥ 29 days to < 24 months of age)	1.25 mL/kg	1.5 mL/kg

Table 11: Summary and Analysis of Total Rescue Opioid Consumption by Dosing Interval

		IV APAP			Control
Dosing Interval	**Statistics**	**A (N=61)**	**B (N=67)**	**A+B (N=128)**	**C+D (N=69)**
After T0 and before T6	LS means	56	**53**	**54**	**53**
	P-value#	0.86	0.99	0.93	
After T6 and before T12	LS means	43	53	49	46
	P-value	0.81	0.59	0.85	
After T12 and before T18	LS means	33	33	33	37
	P-value	0.67	0.65	0.61	
After T18 and Before T24	LS means	31	34	33	36
	P-value	0.62	0.81	0.68	

Source: Clinical Study Report, Table 14.2.1.2;
#P-values are from ANOVA without multiplicity adjustment.

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

/s/

FENG LI
11/30/2016

DAVID M PETULLO
11/30/2016
I concur.