CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

209511Orig1s000

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 209-511
Drug Name:	Xaracoll (bupivacaine HCI collagen-matrix implants)
Indication(s):	Post-surgical analgesia
Applicant:	Innocoll Inc.
Date(s):	Initial Submission Date: 10/31/2016 Resubmission Date: 2/2/2018 PDUFA Due Date: 11/30/2018
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:	Renee Petit-Scott, MD Martha Van Clief, MD
Project Manager:	Allison Meyer

Keywords: NDA review, clinical trials, missing data, imputation

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Innocoll Inc. submitted a new drug application (NDA) for INL-001 for the treatment of postsurgical pain following [redacted] (b) (4) This NDA was initially submitted on October 31, 2016. However, FDA issued a refusal to file (RTF) letter on December 23, 2016. Issues identified included but not limited to lack of pharmacokinetic, nonclinical, and biocompatibility data. Clinical concerns regarding safety database, study design, and labeling were also addressed in the RTF letter. Innocoll Inc. resubmitted the application on February 2, 2018.

This review focuses on the results of two Phase 3, multicenter, randomized, double-blind, and placebo-controlled studies (INN-CB-014 and INN-CB-016). Enrolled patients were healthy male or female who aged 18 years or older with a planned unilateral inguinal hernioplasty under general anesthesia. In each of the two studies, approximately 300 patients were stratified by gender and history of previous hernia repair using mesh and randomly assigned to receive either INL-001 or placebo at a 2:1 ratio within each stratum. The primary efficacy endpoint was time-weighted sum of pain intensity from 0 to 24 hours (SPI24). The key secondary efficacy endpoints were total use of opioid analgesia from 0 to 24 hours (TOpA24), sum of pain intensity from 0 to 48 hours (SPI48), total use of opioid analgesia from 0 to 48 hours (TOpA48), sum of pain intensity from 0 to 72 hours (SPI72), and total use of opioid analgesia from 0 to 72 hours (TOpA72).

In this review, although not explicitly stated, the primary estimand of interest was the difference in mean time-weighted sum of pain intensity from 0 to 24 hours comparing patients with postsurgical pain assigned to INL-001 versus those assigned to placebo regardless of treatment adherence. If a subject used rescue medication, the pain score recorded immediately prior to use of rescue medication was used for any scheduled pain assessment that occurred during the next 30 minutes. This is an acceptable estimand and the methods used to estimate it were appropriate.

The efficacy results in Table 1 have shown that INL-001 was statistically significantly (p-value <0.05) superior to placebo with respect to SPI24 and TOpA24 in the first study, and with respect to SPI24, TOpA24, SPI48, TOpA48 in the second study.

Based on the sequential testing strategy, the treatment effects in the rest of the endpoints (not in bold font) were not considered significantly different. Also, patients who received INL-001 had a longer time to first use of rescue opioid analgesia with a less total use of analgesia compared to patients who received placebo. I therefore conclude that there is sufficient evidence to support the efficacy of INL-001 in providing postsurgical analgesia following open laparotomy inguinal hernioplasty.

Table 1. Efficacy Results for Primary and Key Secondary Endpoints

	Study INN-CB-014				*Study INN-CB-016*			
	LS Mean*		**Difference (95% CI)**	**P-value**	**LS Mean***		**Difference (95% CI)**	**P-value**
	INL-001	**Placebo**			**INL-001**	**Placebo**		
SPI24	68.8	90.3	-21.5 (-33.1, -9.9)	**<0.001**	72.3	101.5	-29.1 (-39.8, -18.4)	**<0.001**
TOpA24	5	10	-4.0 (-6.5, -2.0)	**<0.001**	5	14	-6.0 (-9.0, -4.0)	**<0.001**
SPI48	143.8	166.0	-22.2 (-45.3, 0.9)	0.060	158.4	185.2	-26.8 (-49.0, -4.6)	**0.018**
TOpA48	5	14	-2 (-6.0, 0)	0.025	10	20	-6.0 (-10.0, -2.0)	**<0.001**
SPI72	211.3	232.3	-21 (-54.9, 12.9)	0.224	232.9	255.9	-23 (-55.8, 9.8)	0.170
TOpA72	5	14	-2.0 (-5.0, 0)	0.066	10	20	-6.0 (-11.0, -2.0)	0.002

Source: Reviewer

*For all TOpA endpoints, observed median was reported instead of least squares mean

2 INTRODUCTION

2.1 Overview

Innocoll has submitted a new drug application (NDA) for INL-001 for the treatment of postsurgical pain following [redacted] (b) (4) in adult patients. INL-001 is a single-use surgical implant consisting of three resorbable Type I bovine collagen matrices (3 x 100 mg collagen-matrix implants). Each 5 × 5 cm matrix contains 100 mg of bupivacaine hydrochloride (HCl), a locally acting amide anesthetic that according to the applicant, will produce postsurgical analgesia.

This 505(b)(2) application relies on the previous findings of safety and efficacy for Marcaine® (bupivacaine HCl injection; NDA 016964) approved in the U.S. in 1972. Marcaine® has been used to provide local analgesia for surgical procedures.

All relevant communications with the FDA are summarized below:

At the Type B meeting held on December 5, 2011, FDA provided the following comments:

1. The use of an integrated assessment of pain and opioid use is statistically valid and is acceptable as a primary efficacy endpoint.
2. The analysis must account for multiplicity if secondary endpoints will be included in the label.
3. Intent-to-treat population should not depend on post-treatment efficacy assessments.
4. Imputing pain scores for 1 and 72 hours post-dose only is not appropriate. The analysis should account for all time points. In single-dose acute pain setting, missing data should not be an issue and therefore single imputation methods may be acceptable. All early discontinuations should be thoroughly documented and included as a negative finding.

Innocoll initially submitted this application on October 31, 2016. A refusal to file letter was issued on December 23, 2016 after a preliminary review by FDA. Although there were no statistical issues, the clinical team addressed review issues on study design and proposed labeling. Specifically, given these were identical Phase 3 studies in the same surgical population, the results may not support a broad labeling indication. These trials may not address safety and efficacy questions adequately. Upon resubmission the indication was revised to be treatment of postsurgical pain following [redacted] (b) (4)

On January 30, 2018, Innocoll requested a deferral of submitting the ongoing pediatric study INN-CB-020 in children aged 2 to 17 years, as well as a deferral of initiating the planned pediatric study INN-CB-021 in children aged 0 to 1 year.

To support efficacy, two identical Phase 3, multicenter, randomized, double-blinded, placebo-controlled studies (INN-CB-014 and INN-CB-016) were conducted in adult patients undergoing abdominal surgical procedures.

2.2 Data Sources

All documentation including the study protocol, statistical analysis plan (SAP), clinical study report, and literature referenced, as well as the SDTM and ADaM datasets were submitted under the network path \\CDSESUB1\evsprod\NDA209511\0000. Datasets were submitted by the applicant to the CDER electronic data room in SAS transport format.

Note that the regular folder in Module 5.3.5.3 only contained integrated summary tables from the efficacy analyses. The formal integrated summary of efficacy (ISE) was submitted in Module 2.7 (Summary Clinical of Efficacy).

In response to the information request (IR) sent on July 23, 2018, the applicant discovered that the previously submitted SAP for Study INN-CB-014 was outdated. The final version of SAP including the programming specifications was submitted under the network path \\CDSESUB1\evsprod\NDA209511\0028.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submitted datasets were of acceptable quality and were adequately documented. I was able to reproduce the results of all primary and secondary analyses using the ADaM datasets. However, it was noted that some patients had multiple pre-rescue or pre-acetaminophen pain intensity (PI) assessments at the same analysis time point. In response to an IR sent on July 23, 2018, the applicant explained why this occurred.

- Since the category of pain assessment (scheduled, pre-rescue, pre-acetaminophen) was chosen by the subject in the eDiary at the time the PI score was entered, it is possible that the subject chose the wrong category when entering their PI score into the eDiary. It is also possible that subjects may have had duplicate entries if they did not realize that they had successfully entered their PI assessment in the eDiary.
- ALL available PI scores from time 0 up to 24 hours, were used in the calculation of the AUC (SPI24).

When there were multiple pre-rescue or pre-acetaminophen PI assessments within 30 minutes (or analysis time window) even if due to random data input error, using all PI scores may not be appropriate for calculating the time-weighted sum of pain intensity (SPI) endpoints. In my analyses, if there were multiple PI assessments for any category above, the last PI assessment was used in the calculation of SPI and all SPI analyses were performed based on this calculation using the data derived from the SDTM datasets.

Since there was no baseline pain score available to calculate a SPID, the final version of SAP revised the primary efficacy endpoint, time-weight sum of pain intensity difference (SPID), to the summed pain intensity (SPI). This did not change any results because the baseline PI was set to 0 for all patients as prespecified in the previous version of SAP.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Phase 3 studies INN-CB-014 and INN-CB-016 were multicenter, randomized, double-blind, placebo-controlled studies in adult patients with acute postsurgical pain. The studies were identically designed to evaluate the efficacy, safety, and tolerability of INL-001 after its implantation into surgical wounds. The surgical model in both studies was open laparotomy inguinal hernioplasty using mesh.

Based on the applicant's response to refusal to file letter, although the studies were identical in design and surgical population, they were independent studies for a total of 39 unique sites in the two studies (20 sites in INN-CB-014 and 19 sites in INN-CB-016) and should support an indication to produce postsurgical analgesia following soft tissue surgeries. Randomization was stratified by gender and history of previous ipsilateral hernia repair using mesh. Patients were randomly assigned to treatment with 300 mg INL-001 implants or placebo implants at a 2:1 ratio within each stratum.

Following surgery, patients were transferred to a post-anesthesia care unit (PACU) or other recovery area where they may have received rescue medication of parenteral morphine as needed for breakthrough pain. Once they could tolerate oral medication, they started on a standardized oral analgesic of acetaminophen 650 mg 3 times daily, as well as received oral morphine (15 mg) to manage breakthrough pain only as needed.

Pain intensity at time of implantation (Time 0 or baseline) was set to 0 for all patients. After Time 0, PI was assessed by the patient at 1, 2, 3, 5, 8±1, 12±2, 24±3, 48±3, and 72±4 hours and immediately before taking any parenteral or oral rescue opioid analgesia using an 11-point numerical rating scale (NRS). In addition, patients recorded PI right before taking scheduled acetaminophen until the 72-hour visit.

The primary efficacy endpoint was time-weighted sum of pain intensity from 0 to 24 hours (SPI24) calculated as the area under the PI curve from 0 to 24 hours. The key secondary efficacy endpoints were:

- Total use of opioid analgesia from 0 to 24 hours (TOpA24)
- SPI48
- TOpA48
- SPI72
- TOpA72

I also reviewed other endpoints including TOpA from 0 to 1, 2, 3, 5, 8, and 12 hours, time to first use of rescue opioid analgesia (FOpA), number of patients who did not use any rescue opioid analgesia through 24, 48, and 72 hours, and time to discharge from PACU. Exploratory efficacy endpoints included the integrated time-weighted sum of PI and total use of opioid analgesia from

0 to 24 hours (I-SPI-TOpA24), I-SPI-TOpA48, and I-SPI-TOpA72, etc. The I-SPI-TOpA was calculated based on Silverman's method[1].

Although not explicitly stated, the estimand of primary interest was the difference in mean time-weighted sum of pain intensity from 0 to 24 hours comparing patients with postsurgical pain assigned to INL-001 versus those assigned to placebo regardless of treatment adherence. If a subject used rescue medication, the pain score recorded immediately prior to use of rescue was used for any scheduled pain assessment that occurred during the next 30 minutes.

3.2.2 Statistical Methodologies

The intent-to-treat (ITT) population was used for summarizing patient disposition. It included all randomized patients. All the efficacy analyses were performed using the mITT population which included all randomized patients who received study drug and who have at least one PI score prior to hospital discharge.

Patients were randomly assigned to receive either INL-001 or placebo at a 2:1 ratio within each randomization stratum. A total sample size of 300 patients was planned to provide 90% power to detect a mean difference of 0.4 in SPI24 between treatment groups.

The primary efficacy endpoint was analyzed using an ANOVA model with treatment and randomization strata as factors. Per the SAP, "the normality of the residuals from the ANOVA model will be assessed using the Shapiro-Wilk statistic and graphic examination. If the residuals are not normally distributed, then the primary efficacy analysis will be carried out using the Wilcoxon rank sum test. Analyses of the key secondary efficacy variables will be carried out using the same ANOVA model as for the primary efficacy variable." The SPI data satisfied the normality assumption (diagnostic results not shown) and therefore SPI at other time periods were analyzed using the same ANOVA model as the primary analysis. The TOpA data failed the normality test (diagnostic results not shown), therefore Wilcoxon rank sum test was used for all TOpA analyses. Median treatment difference and its 95% confidence interval (CI) were computed using Hodges-Lehmann estimator and Moses confidence limits. A sequential testing strategy was prespecified to account for multiplicity. If the primary endpoint was significant, p-value < 0.05, the key secondary endpoints were tested in the following order at the level of 0.05:

1. TOpA24
2. SPI48
3. TOpA48
4. SPI72
5. TOpA72

There was no multiplicity adjustment for any other efficacy endpoints.

Time to first use of rescue opioid analgesia (FOpA) was analyzed using Kaplan-Meier curves and stratified log-rank test.

[1] Silverman DG, O'Connor TZ, Brull SJ. Integrated assessment of pain scores and rescue morphine use during studies of analgesic efficacy. Anesth Analg. 1993;77(1):168–170.

For exploratory efficacy endpoints, the applicant adopted Silverman's integrated assessment variable and derived from the time-weighted sum of PI and total use of opioid analgesia from 0 to 24 hours (I-SPI-TOpA24), I-SPI-TOpA48, and I-SPI-TOpA72. The I-SPI-TOpA was calculated as follows:

1. Rank all patients based on their individual endpoint (SPI or TOpA) from the smallest to the largest, when ties occur, assign the average of the ranks.
2. Compute the percentage difference of each patient's rank from the mean rank for each individual endpoint. The mean rank is calculated as (N+1)/2 where N is the total sample size.
3. Add the percentage differences for two individual endpoints to create the subject-level integrated endpoint I-SPI-TOpA.

The integrated endpoint approximately ranges from -200% to 200%. The highest score indicates the least comfortable or the most pain despite the greatest use of opioid analgesia. This method may account for a negative correlation between pain intensity and use of opioid rescue analgesia and therefore be more sensitive to evaluate the treatment effect.

The number and percentage of patients who did not use any rescue opioid analgesia through 24, 48, and 72 hours was tabulated by treatment group. Time to discharge from PACU was analyzed using Kaplan-Meier curves and stratified log-rank test.

The applicant classified PI assessments as shown in

Table 2. In my analysis, if a patient had multiple PI assessments prior to using rescue medication, the last PI assessment that immediately precedes the administration of rescue medication was considered as the "pre-rescue" PI assessment. The applicant's analysis considered all PI assessments including all pre-rescue scores in the calculation of the SPI endpoints.

Table 2. Category of Pain Intensity Assessments

Source	Purpose	Description	Category
eCRF	Pre-rescue PI in the PACU	PI assessment does not immediately precede the administration of a rescue medication within 30 minutes	Unscheduled
eCRF	Pre-rescue PI in the PACU	PI assessment immediately precedes the administration of a rescue medication within 30 minutes	Pre-Rescue[a]
e-Diary	Scheduled PI		Scheduled
e-Diary	Pre-rescue PI after discharge from PACU		Pre-Rescue
e-Diary	Pre-acetaminophen PI after day 0		Pre-Acetaminophen

[a] In determination of the "Pre-Rescue" PI for rescue medications taken in the PACU, a 30 minutes window is used to associate a PI assessment in the PACU (recorded on the CRF) with the rescue medication. If a patient had multiple PI assessments precede a rescue medication within 30 minutes, then the last PI assessment that immediately precedes the administration of rescue medication will be considered as the "Pre-Rescue" PI assessment for that rescue medication.

Source: SAP Table 1

The applicant also defined several circumstances where the PI assessment was considered as missing (referred to Section 5.1.2 of the applicant's SAP (version 2.0 dated April 28, 2016)):

- PI at scheduled time point: If no observed PI assessment falls within the time window of a scheduled time point, the PI value at that time point will be considered as missing.
- Pre-rescue PI in the PACU: If a patient received a rescue medication in the PACU but there is no PI assessment recorded on the CRF that immediately precedes the administration of the rescue medication within 30 minutes, the "pre-rescue" PI score for this rescue medication will be considered as missing.
- Pre-rescue PI after discharge from PACU: If a patient recorded a rescue medication in the eDiary but there is no associated PI assessment recorded in the eDiary, the "pre-rescue" PI score for this rescue medication will be considered as missing.
- Pre-acetaminophen PI after Day 0: If a patient recorded an acetaminophen administration in the eDiary but there is no associated PI assessment recorded in the eDiary, the "pre-acetaminophen" PI score for this acetaminophen dose will be considered as missing.

To handle missing data during the study period up to 72 hours, the applicant used the following methods:

- Missing PI scores before the first observed PI score were imputed using the patient's worst PI score.
- Missing PI score between two observed PI scores, i.e. intermittent missing, was imputed using linear interpolation.
- Missing PI scores after the last observed PI score, i.e. monotone missing, were imputed using last observation carried forward (LOCF), except:

    - If a patient was early terminated and surgery was performed, worst observation carried forward (WOCF) method was applied from the time of early termination through 72 hours.
    - If a patient was early terminated having taken opioid rescue less than 4 hours before termination, LOCF method was applied on the last pre-rescue PI score, otherwise WOCF was applied through 72 hours.
- A pre-rescue or pre-acetaminophen PI score may replace a scheduled PI assessment if it fell within the time window of the scheduled assessment,

In addition, a sensitivity analysis that penalized early dropouts by replacing good pain scores with (worse) group medians was conducted by the applicant.

Based on the discussion of PI assessments (see Section 3.1), besides replicating the applicant's method using ADaM datasets I also used a multiple imputation (MI) method to handle missing pain assessments in my primary analysis of SPI24. As a sensitivity analysis, I used LOCF to impute missing pain assessments. The MI model I used was Markov Chain Monte Carlo (MCMC) option with predictors treatment, gender, and history of previous hernia repair. The data were imputed 20 times and were then analyzed using the same ANOVA model as the primary analysis. This method assumed that the data were missing at random. Given that there was only a small amount of missing data (3% at 24 hours) and the drug is administered as a single implantation, a subject cannot discontinue treatment unless it is surgically removed. Therefore, LOCF may be considered an acceptable approach. In the implementation, all missing PI scores were imputed using LOCF except that missing values before the first observed PI score were imputed using the patient's worst PI score.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A summary of patient disposition for each study is presented in
Table 3 and Table 4. In Study INN-CB-014, 305 patients (298 patients in mITT population) were randomly assigned to INL-001 (204 patients; 197 patients in mITT) or placebo (101 patients). In Study INN-CB-016, 319 patients (312 patients in mITT) were randomly assigned to INL-001 (213 patients; 207 patients in mITT) or placebo (106 patients; 105 patients in mITT). As expected, completion rates were greater than 95% in both studies. The major reason for early dropout in both studies was being lost to follow-up (approximately 2%). Other reasons included adverse event (2 patients), subject did not show up at Day 30 (2 patients), investigator decision (1 patient), randomized but not enrolled (3 patients), fail to meet randomization criteria (1 patient), subject withdrawal (1 patient), and death (1 patient).

Table 3. Patient Disposition, Study INN-CB-014

	INL-001	Placebo	Total
Randomized	204	101	305
Completed the study	196 (96.1%)	100 (99.0%)	296 (97.1%)
Discontinued the study	8 (3.9%)	1 (1.0%)	9 (2.9%)
Adverse event	1 (0.5%)	0	1 (0.3%)
Lost to follow-up	4 (2.0%)	1 (1.0%)	5 (1.6%)

Other	3 (1.5%)	0	3 (1.0%)

Source: Study INN-CB-014 CSR Table 4

Table 4. Patient Disposition, Study INN-CB-016

	INL-001	Placebo	Total
Randomized	213	106	319
Completed the study	203 (95.3%)	103 (97.2%)	306 (95.9%)
Discontinued the study	10 (4.7%)	3 (2.8%)	13 (4.1%)
Adverse event	0	1 (0.9%)	1 (0.3%)
Lost to follow-up	5 (2.3%)	1 (0.9%)	6 (1.9%)
Other	5 (2.3%)	1 (0.9%)	6 (1.9%)

Source: Study INN-CB-016 CSR Table 4

Demographics and baseline characteristics are presented in Table 5 and Table 6. There were no statistically significant differences between treatment groups for any of the demographic and baseline variables. Most patients were male (96.1% in Study 014 and 97.5% in Study 016), White (90.5% in Study 014 and 85.3% in Study 016), less than 65 years (84.3% in Study 014 and 85.3% in Study 016), and most did not have a history of hernia repair using mesh (89.5% in Study 014 and 89.3% in Study 016).

Table 5. Demographics and Baseline Characteristics, Study INN-CB-014

	INL-001 (N=204)	Placebo (N=101)	Total (N=305)
Age (years): n (%)			
< 65	174 (85.3)	83 (82.2)	257 (84.3)
≥ 65	30 (14.7)	18 (17.8)	48 (15.7)
Mean (SD)	53.1 (12.8)	53.3 (14.0)	53.1 (13.2)
Min, max	19, 83	21, 86	19, 86
Gender: n (%)			
Male	196 (96.1)	97 (96.0)	293 (96.1)
Female	8 (3.9)	4 (4.0)	12 (3.9)
Ethnicity: n (%)			
Hispanic or Latino	77 (37.7)	38 (37.6)	115 (37.7)
Not Hispanic or Latino	127 (62.3)	62 (61.4)	189 (62.0)
Missing	0 (0)	1 (1.0)	1 (0.3)
Race: n (%)			
White	185 (90.7)	91 (90.1)	276 (90.5)
Black or African American	15 (7.4)	7 (6.9)	22 (7.2)
Asian	2 (1.0)	2 (2.0)	4 (1.3)
Other	2 (1.0)	1 (1.0)	3 (1.0)

Body Mass Index (kg/m2)			
Mean (SD)	27.1 (3.9)	27.3 (4.6)	27.1 (4.1)
Min, max	18.7, 39.6	19.2, 42.1	18.7, 42.1
Missing, n (%)	2 (1.0)	1 (1.0)	3 (1.0)
Previous Hernia Repair using Mesh: n (%)			
Yes	20 (9.8)	12 (11.9)	32 (10.5)
No	184 (90.2)	89 (88.1)	273 (89.5)

Source: Modified CSR Table 7, Study INN-CB-014
SD: standard deviation

Table 6. Demographics and Baseline Characteristics, Study INN-CB-016

	INL-001 (N=213)	**Placebo (N=106)**	**Total (N=319)**
Age (years): n (%)			
< 65	180 (84.5)	92 (86.8)	272 (85.3)
≥ 65	33 (15.5)	14 (13.2)	47 (14.7)
Mean (SD)	50.7 (13.7)	48.5 (13.9)	50.0 (13.8)
Min, max	18, 85	19, 75	18, 85
Gender: n (%)			
Male	208 (97.7)	103 (97.2)	311 (97.5)
Female	5 (2.3)	3 (2.8)	8 (2.5)
Ethnicity: n (%)			
Hispanic or Latino	44 (20.7)	23 (21.7)	67 (21.0)
Not Hispanic or Latino	169 (79.3)	83 (78.3)	252 (79.0)
Race: n (%)			
White	182 (85.4)	90 (84.9)	272 (85.3)
Black or African American	23 (10.8)	12 (11.3)	35 (11.0)
Asian	4 (1.9)	3 (2.8)	7 (2.2)
Other	3 (1.4)	1 (0.9)	4 (1.2)
Missing	1 (0.5)	0 (0)	1 (0.3)

Body Mass Index (kg/m2)			
Mean (SD)	26.8 (4.0)	27.2 (5.1)	27.0 (4.4)
Min, max	17.8, 40.8	17.4, 45.9	17.4, 45.9
Missing, n (%)	4 (1.9)	0 (0)	4 (1.2)
Previous Hernia Repair using Mesh: n (%)			
Yes	22 (10.3)	10 (9.4)	32 (10.0)
No	189 (88.7)	96 (90.6)	285 (89.3)
Missing	2 (0.9)	0 (0)	2 (0.6)

Source: Modified CSR Table 7, Study INN-CB-016
SD: standard deviation

3.2.4 Results and Conclusions

The primary efficacy analysis results for both studies are shown in Table 7 and Table 8. The results from the MI method ("Reviewer's Results") were consistent with the replicated results from the applicant's method ("Applicant's Results") in both studies. INL-001 was statistically significantly superior (p-values < 0.001) to placebo with respect to SPI24. Patients who received INL-001 had less pain over the 24-hour postsurgical period compared with patients who received placebo.

Table 7. Primary Efficacy Analysis Results for SPI24, Study INN-CB-014

	INL-001	Placebo
Applicant's Results		
LS Mean	71.0	91.7
Diff (95% CI)	-20.8 (-32.2, -9.4)	
P-value	<0.001	
Reviewer's Results		
LS Mean	68.8	90.3
Diff (95% CI)	-21.5 (-33.1, -9.9)	
P-value	<0.001	

Source: Reviewer

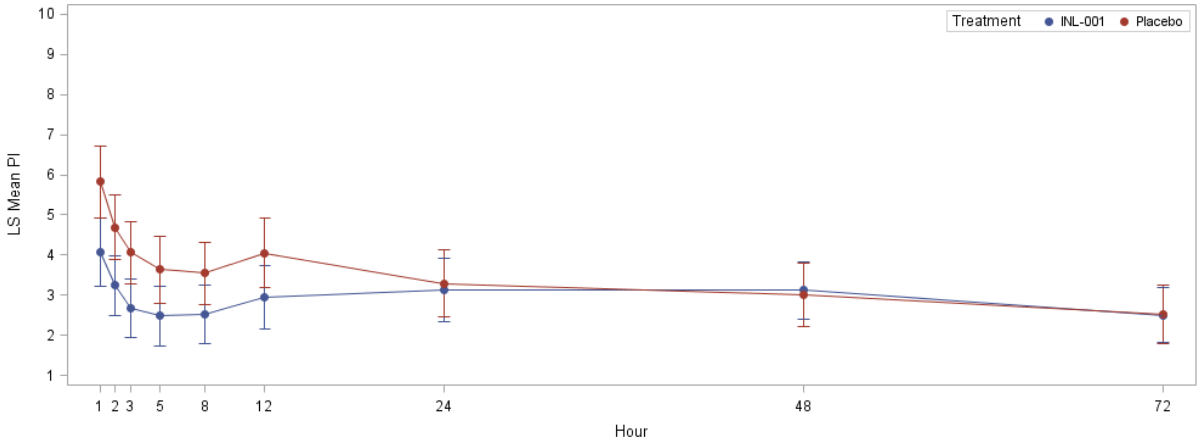
Table 8. Primary Efficacy Analysis Results for SPI24, Study INN-CB-016

	INL-001	Placebo
Applicant's Results		
LS Mean	68.6	96.5
Diff (95% CI)	-27.8 (-38.6, -17.1)	
P-value	<0.001	
Reviewer's Results		
LS Mean	72.3	101.5
Diff (95% CI)	-29.1 (-39.8, -18.4)	
P-value	<0.001	

Source: Reviewer

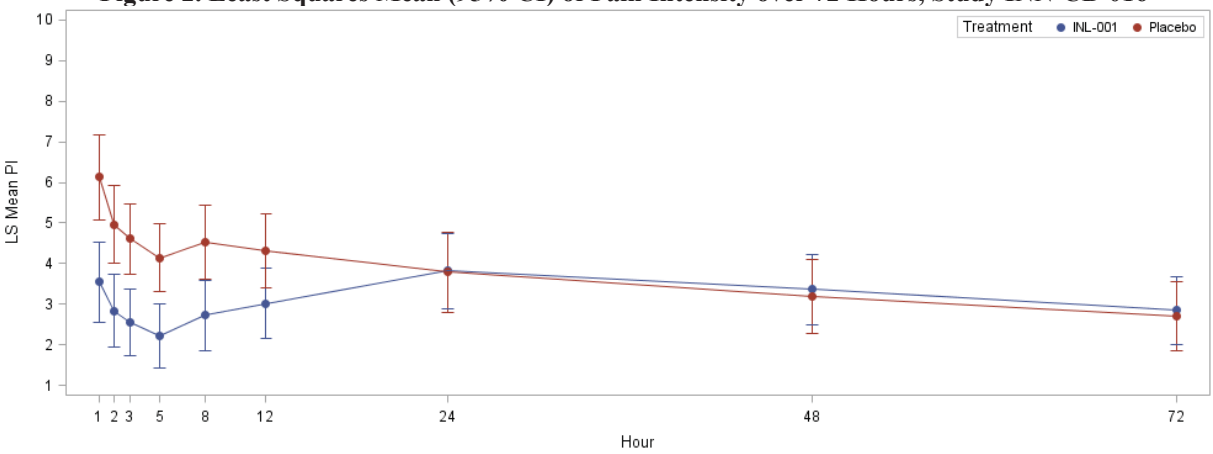
Since differenece in summed pain intensity (SPI) is not easily interepretted clinicially, I also examined mean PI scores over the 72 hours using the same ANOVA model as the primary analysis. Starting from 24 hours, there was no difference in pain intensity between treatment groups and the mean PI scores at 24 and 48 hours were greater than 3 for both studies (Figure 1 and Figure 2). The results from these analyses, although not pre-specified, support the primary efficacy endpoint.

Figure 1. Least Squares Mean (95% CI) of Pain Intensity over 72 Hours, Study INN-CB-014



Source: Reviewer

Figure 2. Least Squares Mean (95% CI) of Pain Intensity over 72 Hours, Study INN-CB-016



Source: Reviewer

My analyses of the key secondary efficacy endpoints were also consistent with the applicant's analyses in both studies. In Study INN-CB-014, patients who received INL-001 used statistically significantly less (median) total opioid rescue analgesic compared with patients who received placebo over the first 24 hours (TOpA24). There was no significant mean difference with respect to SPI48. Based on the sequential testing strategy, the rest of key secondary endpoints (TOpA48, SPI72, and TOpA72) were not significant.

Table 9. Key Secondary Efficacy Results for Study INN-CB-014

	Applicant's Results	*Reviewer's Results*

	LS Mean*		Diff (95% CI)	P-value	LS Mean*		Diff (95% CI)	P-value
	INL-001	Placebo			INL-001	Placebo		
TOpA24	5.0	10.0	-4.0 (-6.5, -2.0)	**<0.001**	5.0	10.0	-4.0 (-6.5, -2.0)	**<0.001**
SPI48	147.6	169.5	-22.0 (-44.6, 0.6)	0.057	143.8	166.0	-22.2 (-45.3, 0.9)	0.060
TOpA48	5.0	14.0	-2.0 (-6.0, 0)	0.025	5.0	14.0	-2.0 (-6.0, 0)	0.025
SPI72	215.4	238.5	-23.1 (-56.5, 10.3)	0.174	211.3	232.3	-21.0 (-54.9, 12.9)	0.224
TOpA72	5.0	14.0	-2.0 (-5.0, 0)	0.066	5.0	14.0	-2.0 (-5.0, 0)	0.066

Source: Reviewer
*For all TOpA endpoints, observed median was reported instead of least squares mean

In Study INN-CB-016, INL-001 was statistically significantly superior to placebo with respect to TOpA24, SPI48, and TOpA48. Based on the sequential testing strategy, the rest of key secondary endpoints SPI72 and TOpA72 were not significant. Patients who received INL-001 had both less pain and used less opioid rescue analgesic through 48 hours compared with patients who received placebo.

Table 10. Key Secondary Efficacy Results for Study INN-CB-016

	Applicant's Results				*Reviewer's Results*			
	LS Mean*		Diff (95% CI)	P-value	LS Mean*		Diff (95% CI)	P-value
	INL-001	Placebo			INL-001	Placebo		
TOpA24	5.0	14.0	-6.0 (-9.0, -4.0)	**<0.001**	5.0	14.0	-6.0 (-9.0, -4.0)	**<0.001**
SPI48	154.8	179.0	-24.2 (-45.7, -2.8)	**0.027**	158.4	185.2	-26.8 (-49.0, -4.6)	**0.018**
TOpA48	10.0	20.0	-6.0 (-10.0, -2.0)	**<0.001**	10.0	20.0	-6.0 (-10.0, -2.0)	**<0.001**
SPI72	223.5	247.1	-23.6 (-55.7, 8.5)	0.149	232.9	255.9	-23.0 (-55.8, 9.8)	0.170
TOpA72	10.0	20.0	-6.0 (-11.0, -2.0)	0.002	10.0	20.0	-6.0 (-11.0, -2.0)	0.002

Source: Reviewer
*For all TOpA endpoints, observed median was reported instead of least squares mean

Table 11

For time to first use of rescue opioid analgesia, patients who received INL-001 had a statistically significantly longer time to first use of opioid analgesia compared with patients who received placebo (p-values < 0.001) in both studies (Figure 3 and Figure 4). There was no significant difference between treatment groups with respect to time to discharge from PACU in both studies (Figure 5 and Figure 6).

APPEARS THIS WAY ON ORIGINAL

Table 11. Total Use of Opioid Analgesia From 0 up to 72 Hours shows the observed mean and median total use of opioid analgesia from 0 up to 72 hours. In both studies, patients who received INL-001 required statistically significantly less (median) opioid analgesic at all time periods through 48 hours compared with patients who received placebo. The median difference in TOpA72 was also statistically significant in Study INN-CB-016.

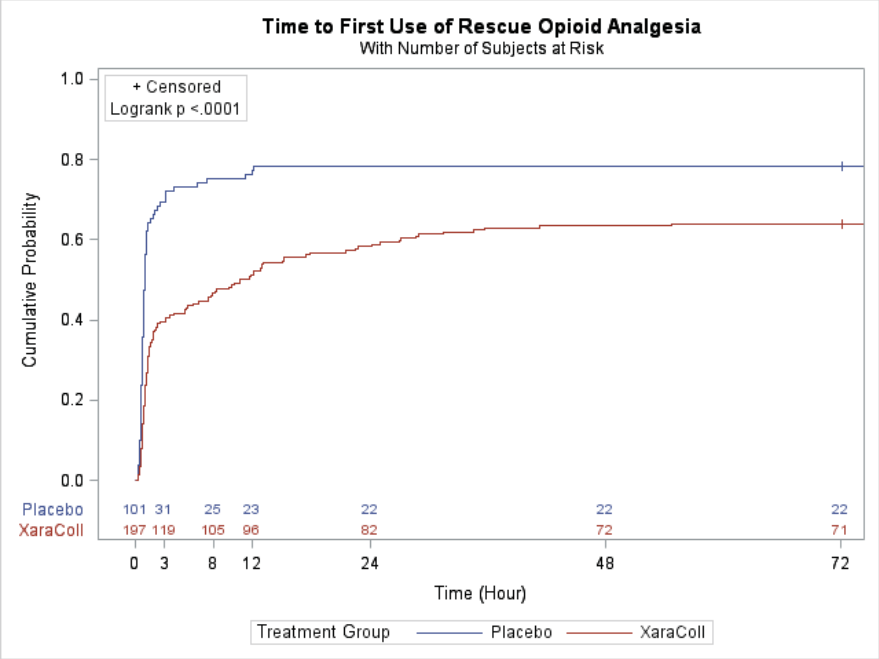
For time to first use of rescue opioid analgesia, patients who received INL-001 had a statistically significantly longer time to first use of opioid analgesia compared with patients who received placebo (p-values < 0.001) in both studies (Figure 3 and Figure 4). There was no significant difference between treatment groups with respect to time to discharge from PACU in both studies (Figure 5 and Figure 6).

Table 11. Total Use of Opioid Analgesia From 0 up to 72 Hours

Analysis Visit	Treatment	Study INN-CB-014		Study INN-CB-016	
		Mean	Median	Mean	Median
1 Hour	Placebo	1.8	1	2.1	2
	INL-001	0.7	0	0.7	0
2 Hour	Placebo	4.1	3	5.6	5
	INL-001	1.8	0	2.3	0
3 Hour	Placebo	4.8	4	6.3	6
	INL-001	2.1	0	2.6	0
5 Hour	Placebo	5.7	5	7.2	6
	INL-001	2.6	0	3.4	0
8 Hour	Placebo	7.1	6	9	8
	INL-001	3.6	0	4.1	1
12 Hour	Placebo	8.9	8	12	11
	INL-001	4.6	1	5.9	2
24 Hour	Placebo	12.3	10	16.6	14
	INL-001	7.6	5	9.4	5
48 Hour	Placebo	16.4	14	23	20
	INL-001	13.5	5	15.6	10
72 Hour	Placebo	18.6	14	27.4	20
	INL-001	16.6	5	19.7	10

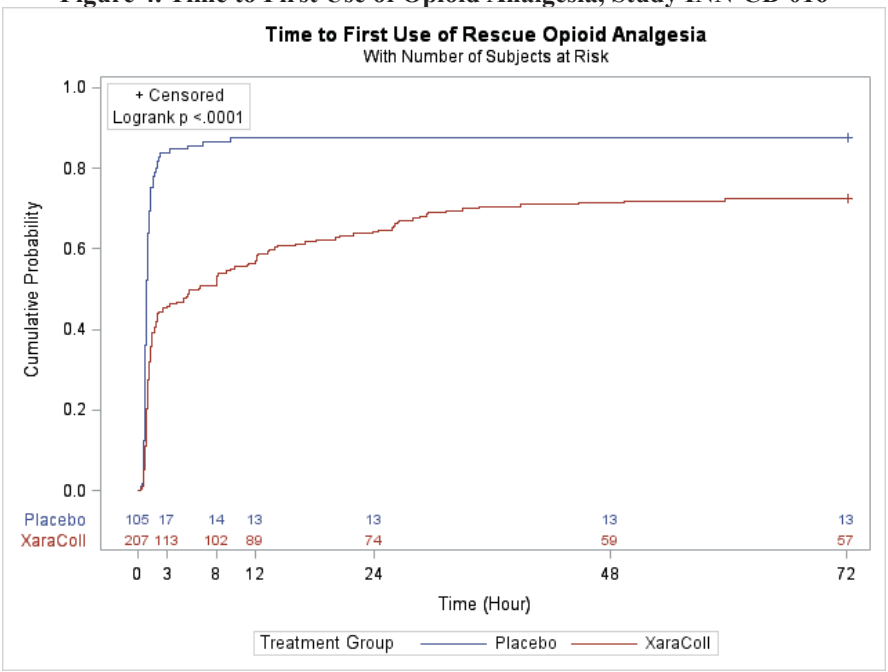
Source: Reviewer

Figure 3. Time to First Use of Opioid Analgesia, Study INN-CB-014



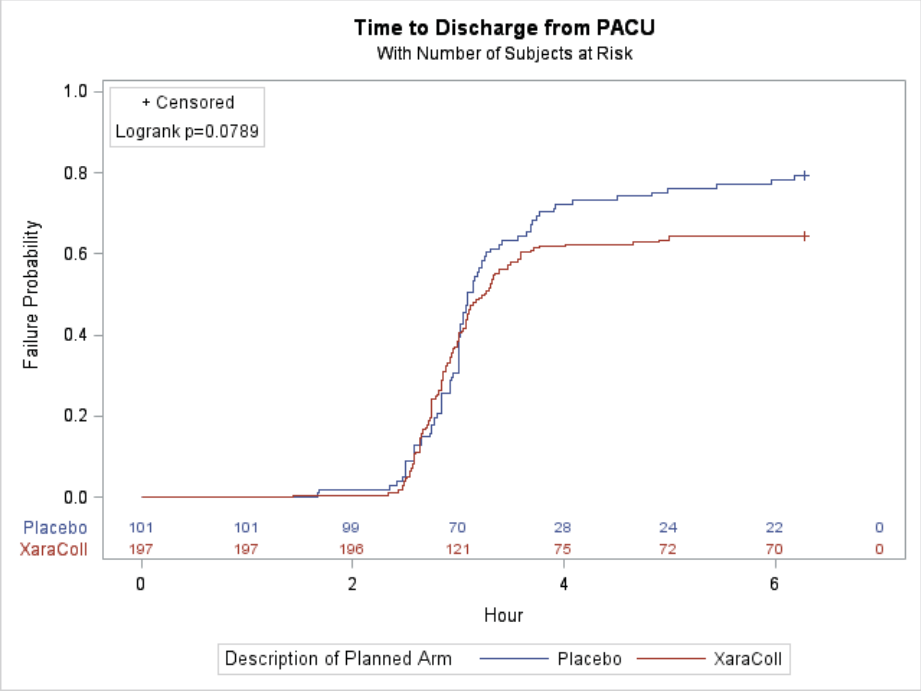
Source: Reviewer

Figure 4. Time to First Use of Opioid Analgesia, Study INN-CB-016



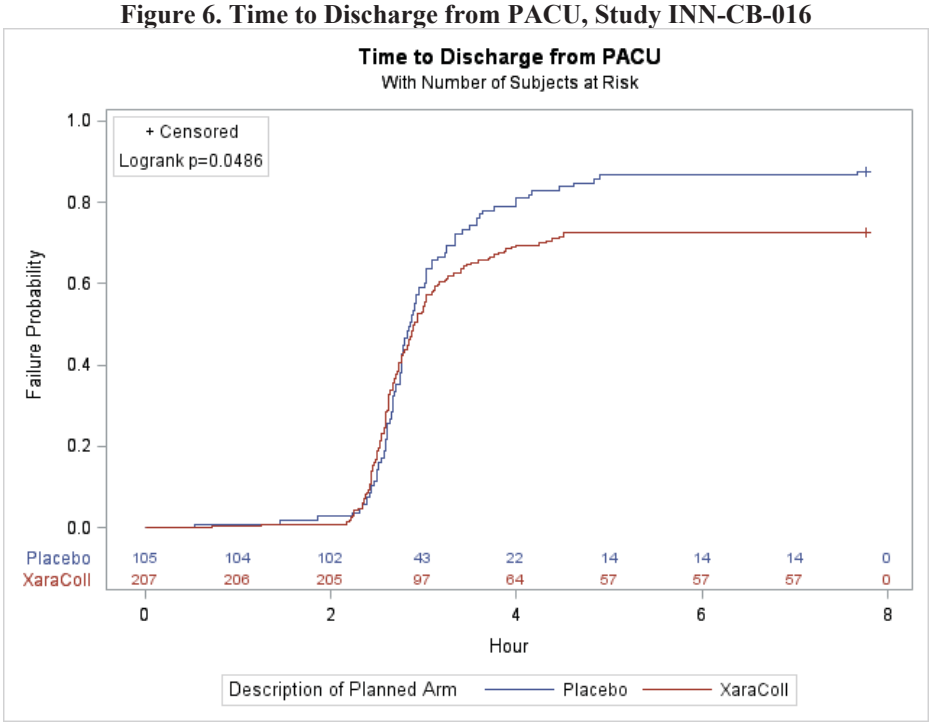
Source: Reviewer

Figure 5. Time to Discharge from PACU, Study INN-CB-014



Source: Reviewer

Figure 6. Time to Discharge from PACU, Study INN-CB-016



Source: Reviewer

The number and percentage of subjects who did not use any rescue opioid analgesia are listed in Table 12. Across the two studies, more patients who received INL-001 did not use any rescue

opioid medication at all time periods through 72 hours compared with patients who received placebo. In Study INN-CB-014 (Study INN-CB-016), 36% (27.5%) patients in INL-001 group did not use rescue opioid medication throughout 72 hours compared with 21.8% (12.4%) patients in placebo group.

Table 12. Number of Patients Who Did Not Use Rescue Opioid Analgesia

Study	Analysis Visit	Treatment	N	Percentage (%)	Total
014	24 Hours	Placebo	22	21.8	101
		INL-001	82	41.6	197
	48 Hours	Placebo	22	21.8	101
		INL-001	72	36.5	197
	72 Hours	Placebo	22	21.8	101
		INL-001	71	36.0	197
016	24 Hours	Placebo	13	12.4	105
		INL-001	74	35.7	207
	48 Hours	Placebo	13	12.4	105
		INL-001	59	28.5	207
	72 Hours	Placebo	13	12.4	105
		INL-001	57	27.5	207

Source: Reviewer

The exploratory analyses for the integrated pain and total use of opioid analgesia, I-SPI-TOpA, showed that INL-001 was statistically significantly superior (p-values < 0.05) to placebo from Time 0 through 24 and 48 hours in both studies, and up to 72 hours in Study INN-CB-016 (p = 0.008).

Table 13. Integrated Sum of Pain Intensity and Total Use of Opioid Analgesia

	Study INN-CB-014				***Study INN-CB-016***			
	INL-001	**Placebo**	**Diff (95% CI)**	**P-value**	**INL-001**	**Placebo**	**Diff (95% CI)**	**P-value**
ISPITOpA24	-49.5	2.1	-51.5 (-74.9, -28.2)	**<0.001**	-51.0	18.6	-69.6 (-92.0, -47.2)	**<0.001**
ISPITOpA48	-43.4	-15.8	-27.6 (-51.5, -3.7)	**0.024**	-42.6	-2.1	-40.5 (-63.8, -17.3)	**0.001**
ISPITOpA72	-39.5	-18.7	-20.8 (-44.7, 3.1)	0.088	-40.3	-8.1	-32.2 (-55.7, -8.7)	**0.008**

Source: CSR Table 14.2.6.2

The sensitivity analyses confirmed the primary findings in both studies and concluded that the model assumptions were met. There was no issue with missing data. INL-001 was statistically significantly superior (p-values < 0.001) to placebo with respect to SPI24 (Table 14 and Table 15).

Table 14. Sensitivity Analyses, Study INN-CB-014

Study INN-CB-014	***INL-001***	***Placebo***	***Diff (95% CI)***	***P-value***
Applicant's Results	71.1	91.2	-20.1	<0.001

			(-31.5, -8.7)	
Complete Case	60.6	79.8	-19.2 (-29.9, -8.4)	<0.001
LOCF	68.9	91.6	-22.7 (-34.3, -11.0)	<0.001

Source: Reviewer

Table 15. Sensitivity Analyses, Study INN-CB-016

Study INN-CB-016	***INL-001***	***Placebo***	***Diff (95% CI)***	***P-value***
Applicant's Results	67.4	94.7	-27.4 (-38.2, -16.6)	<0.001
Complete Case	70	93.5	-23.5 (-34.5, -12.5)	<0.001
LOCF	71.7	102.9	-31.2 (-42.0, -20.4)	<0.001

Source: Reviewer

4 FINDINGS IN SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Surgical History

I conducted subgroup analyses for the primary efficacy endpoint SPI24 by age (<65 years and ≥65 years), gender (male and female), race (White and non-White), and history of previous hernia surgery (yes or no).

Some subgroups such as female, non-Caucasian, and patients who previously had hernia surgery had limited number of patients and resulted in large variability and unreliable results (Figure 7 and Figure 8). The direction of treatment effect was consistent across subgroups although not statistically significant in gender and surgical history subgroups. Numerically lower mean sum of pain intensity from 0 through 24 hours for the INL-001 group were observed in all subgroups. There were statistically significant interactions between treatment and age in both studies. The results showed that patients who aged ≥65 years had a lower mean SPI24 than patients who aged <65 years but the studies were not designed to look for differences due to age.

APPEARS THIS WAY ON ORIGINAL

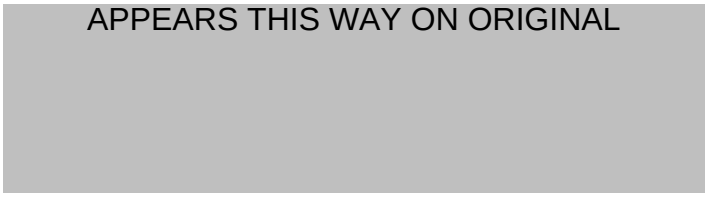
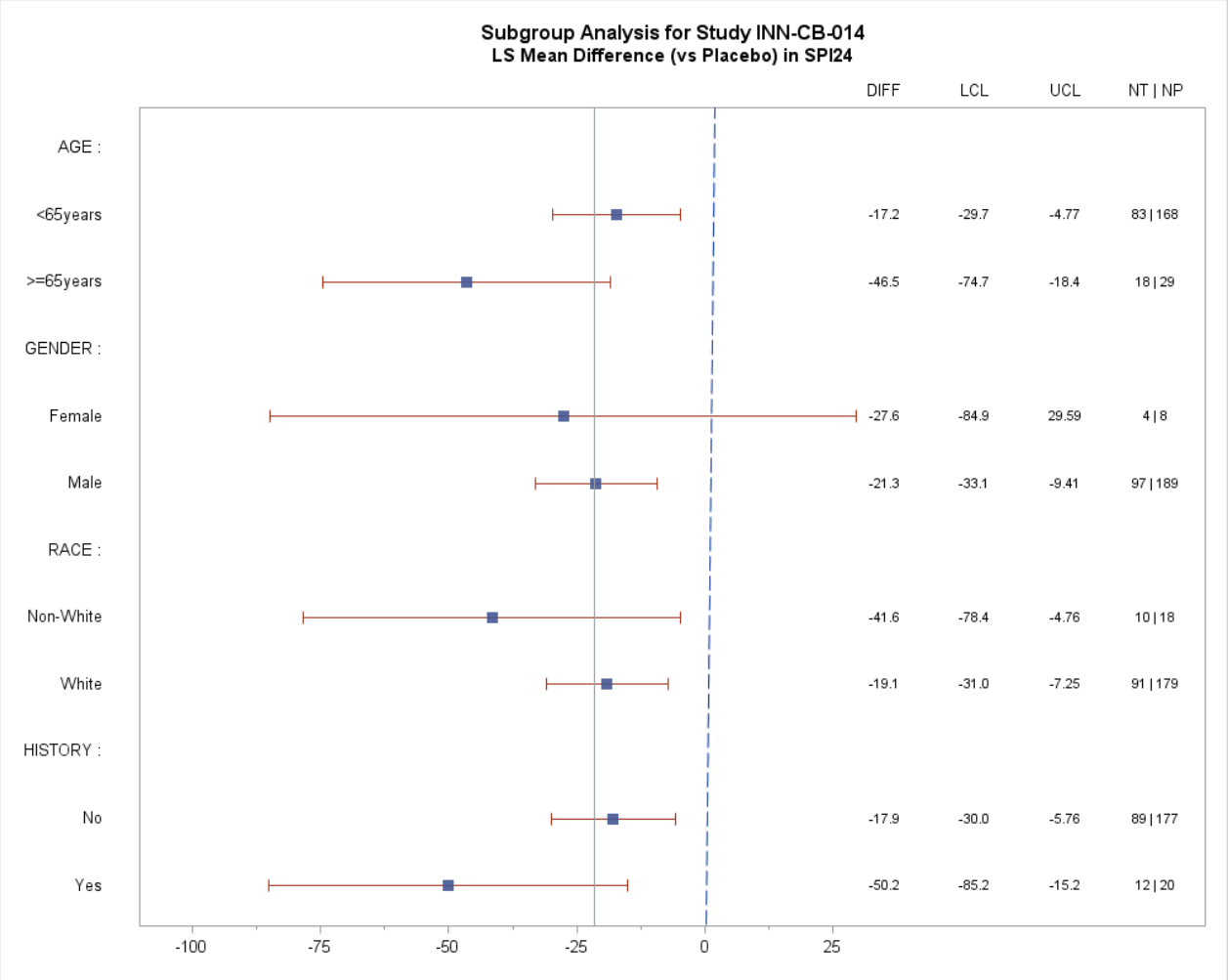
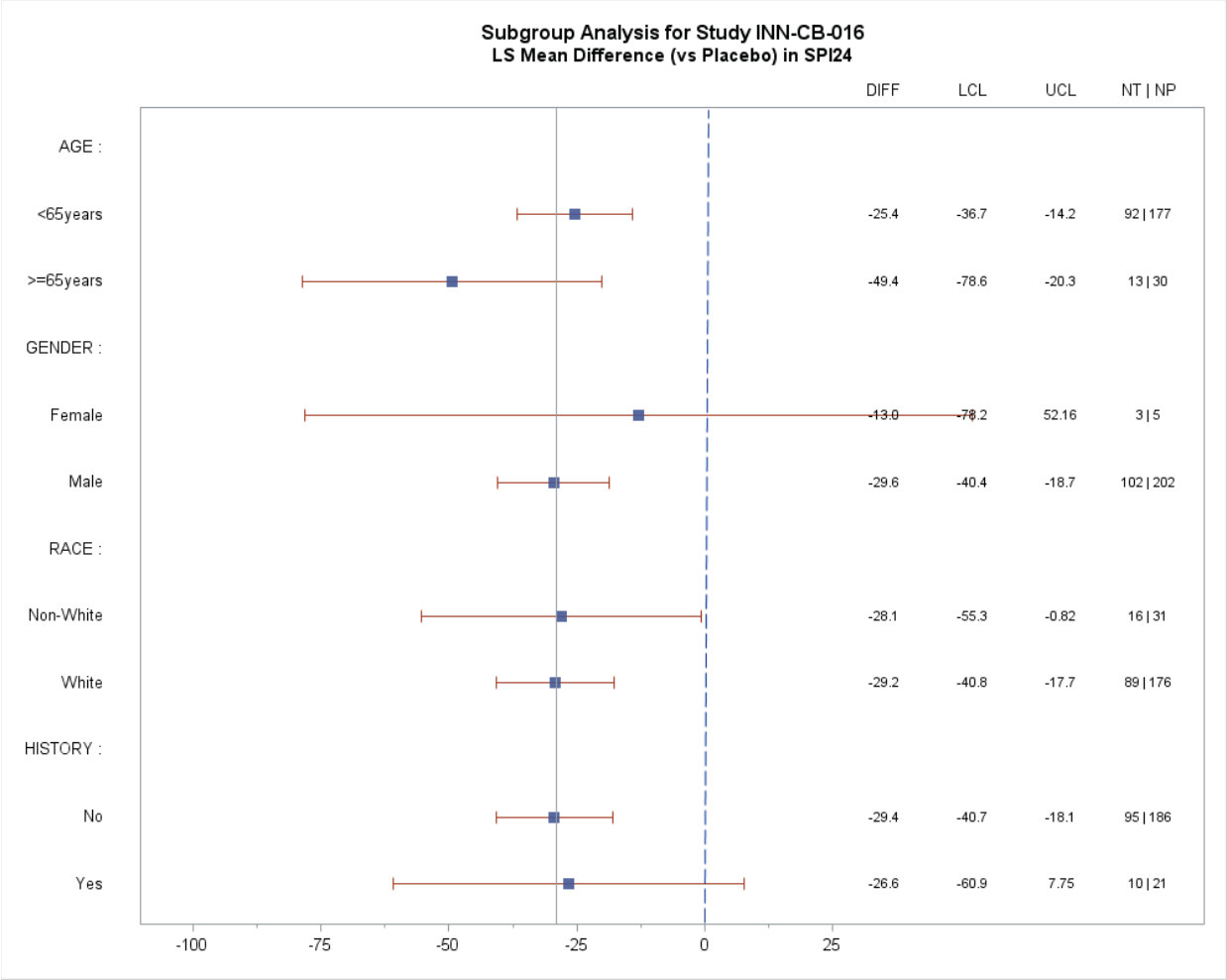


Figure 7. Primary Efficacy Analysis by Subgroup, Study INN-CB-014



Source: Reviewer
The least squares mean difference of INL-001 over placebo with respect to SPI24 was analyzed using ANCOVA with treatment, gender, history of previous hernia surgery, and interaction of treatment and the subgroup of interest.
DIFF: LS mean difference
LCL: lower 95% confidence limit
UCL: upper 95% confidence limit
NT: number of subjects in the active treatment group
NP: number of subjects in the placebo group
Solid vertical line represents mean overall estimated effect size.
Dashed vertical line represents no effect.

Figure 8. Primary Efficacy Analysis by Subgroup, Study INN-CB-016



Source: Reviewer
The least squares mean difference of INL-001 over placebo with respect to SPI24 was analyzed using ANCOVA with treatment, gender, history of previous hernia surgery, and interaction of treatment and the subgroup of interest.
DIFF: LS mean difference
LCL: lower 95% confidence limit
UCL: upper 95% confidence limit
NT: number of subjects in the active treatment group
NP: number of subjects in the placebo group
Solid vertical line represents mean overall estimated effect size.
Dashed vertical line represents no effect.

5 SUMMARY AND CONCLUSIONS

5.1 Conclusions and Recommendations

Based on the analysis of the primary efficacy endpoint, SPI24, there is sufficient evidence to support the efficacy of INL-001 in providing postsurgical analgesia following open laparotomy inguinal hernioplasty. Patients who received INL-001 had less pain over the 24-hour postsurgical period compared with patients who received placebo. The primary efficacy findings were supported by total use of opioid analgesia from 0 to 24 hours (TOpA24) and time to first use of rescue opioid medication (FOpA). Patients who received INL-001 had a longer time to first use of rescue opioid analgesia and on average, used less rescue opioid analgesia compared to patients who received placebo. When both pain and the use of rescue opioid analgesia (I-SPI-TOpA) were considered, INL-001 showed a significant treatment effect over placebo over the 48-hour postsurgical period.

Results from sensitivity analyses were consistent with the primary efficacy analysis regardless of imputation methods on missing data.

5.2 Labeling Recommendations

Adverse reactions reported in Section 6 of the draft labeling were based on combined Phase 3 studies. It is acceptable given the fact that the two studies were identically designed under the same surgical model. I agree with most of the clinical studies results reported in Section 14. However, I recommend removing the following (b) (4)

(b) (4)

Table 16. Efficacy Results for Primary and Key Secondary Endpoints from the Draft Labeling

(b) (4)

	Study 1		Study 2	
	XARACOLL N=197	Placebo[3] N=101	XARACOLL N=207	Placebo[3] N=105
SPI24[1] Mean (SD)	85.9 (47.2)	106.8 (48.2)	88.3 (47.0)	116.2 (44.0)

(b) (4)

Source: Draft Labeling Table 3

(b) (4)

References

Silverman DG, O'Connor TZ, Brull SJ. Integrated assessment of pain scores and rescue morphine use during studies of analgesic efficacy. Anesth Analg. 1993;77(1):168–170.

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/

YI N REN
10/24/2018

DAVID M PETULLO
10/24/2018
I concur.