AcelRx Pharmaceuticals, Inc.

PROTOCOL No. IAP312

A Multicenter, Open-Label Trial to Evaluate the Overall Performance of the Zalviso System (sufentanil sublingual tablet system) 15 mcg

Statistical Analysis Plan for Final Data Analysis

(Final Version, 6/10/2017)

Prepared by: _____ _____
Yu-Kun Chiang, PhD **Date**
Statistical Consultant to AcelRx

AcelRx:
Approved by: _____ _____
Pamela Palmer, M.D., Ph.D. **Date**
Chief Medical Officer

TABLE OF CONTENTS

1 Introduction

This is a multicenter, open-label, single treatment clinical study designed to evaluate the overall functionality of the Zalviso System (sufentanil sublingual tablet system) 15 mcg for the treatment of acute moderate-to-severe postoperative pain.

The study was designed to demonstrate that the device dispense failure rate is non-inferior to a target of 2%. This is based on the criteria that the upper limit of a 90% confidence interval (CI) of the device dispense failure rate is not worse than 5%.

The statistical issues regarding the conduct and analysis of this study are described briefly in the following sections.

2 Statistical Methods

2.1 Determination of Sample Size

A sample size of approximately 330 patients is planned for this study in order to have at least 315 patients receive study drug and have available efficacy data for data analysis.

A sample size of 315 patients will have 90% power to demonstrate that the device dispense failure rate is non-inferior to a target of 2%. This failure rate is the proportion of patients who experienced at least one of following events: (1) a system-generated error, (2) a tablet dispensed but not requested, or (3) a tablet dispensed when the Zalviso System is in lockout. All events listed above will be obtained from the Controller patient usage data.

This is based on the criteria that the upper limit of a 90% confidence interval (CI) of the device dispense failure rate is not worse than 5%. This calculation is based on the exact binomial test of one-sample proportion for a one-sided case against a 2% target proportion with a delta of 3%, and a one-sided significance level of α=0.05. This sample size was generated from the SAS procedure PROC POWER for one-sample proportion. Exact power computations are based on the binomial distribution and computing formulas provided by Johnson and Kotz (1970). Assuming a 5% non-evaluable rate, a sample size of 330 patients is planned for this study.

2.2 Data Processing and Management

All data management activities including the Case Report Form (CRF) data entry, query generation, query resolution and data base maintenance, were performed by the Data Management Department of PRA Health Sciences.

2.3 Performance of Statistical Data Analysis

The statistical data analysis will be performed by Essence Sciences Inc.

2.4 Protocol Deviations and Special Cases

Enrolled patients who deviate from the study protocol will be listed in Appendix 16.2. Protocol deviations at entry and deviations related to the efficacy measurements during the study will be identified.

2.5 Test of Hypothesis and Significance Levels

Both continuous and categorical efficacy variables will be summarized descriptively. A point estimate of the device dispense failure rate and its 90% exact Clopper-Pearson confidence interval will be constructed. The statistical tests for the analysis of other efficacy variables and safety data will be performed at the $\alpha = 0.05$ significance level. All tests will be two-sided.

2.6 Randomization and Blinding of the Treatment Assignment

This is an open-label single treatment study. There is no randomization of patients or blinding of treatment. Patients who meet the eligibility requirements will be enrolled into this study sequentially within each study center to receive study treatment.

2.7 Patient Enrollment and Disposition

Post-text Table 14.1.1 summarizes the number of patients enrolled in this study by surgery type within each investigator's site. Post-text Table 14.1.2 defines the patient population used for the efficacy and safety data analysis.

2.8 Premature Termination from the Study

Patients who prematurely terminate from the study will be summarized according to the reason for termination. Each unique reason will be determined by the investigator.

Patients who prematurely terminated from the study and reasons for not entering next time period will be summarized in Post-text Tables 14.1.3 to 14.1.7.

2.9 Baseline Comparability

Demographics and baseline characteristics will be summarized for all enrolled patients in Post-text Table 14.1.8. Data will be pooled for all study centers for the descriptive summary of baseline data. No formal hypothesis testing will be performed for baseline data. Similar summaries will be performed for the safety population separately in Post-text Table 14.1.9.

2.10 Analysis of Efficacy Data

Sections related to the analysis of efficacy data are specified below by topic.

2.10.1 Efficacy Variables

The efficacy variables and parameters are:

(1) Proportion of patients who experienced at least one system-generated error based on the Controller data while using the Zalviso System

(2) Proportion of patients, if any, with tablet dispensed but not requested

(3) Proportion of patients, if any, with tablet dispensed when the Zalviso System is in lockout

(4) Proportion of patients with misplaced tablet(s)

(5) Number of misplaced tablets (i.e., tablet found outside the patient's mouth)

(6) Proportion of patients who experienced either a system-generated error or a misplaced tablet (i.e. a dispense failure)

(7) Assessment of the number of Zalviso System notifications to the nurse to retrain patient to not pull down on Controller while dosing (i.e., mitigation of Error 302)

(8) Proportion of patients who rate the Patient Global Assessment (PGA) of method of pain control over 24, 48 and 72 hours as "good" or "excellent"

(9) Proportion of patients who responded in each category of the PGA

(10) Proportion of Healthcare Professionals who rate the Global Assessment (HPGA) of method of pain control over 24, 48 and 72 hours as "good" or "excellent"

(11) Proportion of Healthcare Professionals who responded in each category of the HPGA

(12) Proportion of patients who terminate from the study due to inadequate analgesia over the 24-hour, 48-hour and 72-hour study period

(13) Time to terminate from the study due to inadequate analgesia

(14) Time-weighted summed pain intensity difference (SPID) over the 24-hour study period (SPID24), over the 48-hour study period (SPID48), and over the 72-hour study period (SPID72)

(15) SPID up to each evaluation time point

(16) Total pain relief (TOTPAR) over the 24-hour study period (TOTPAR24), over the 48-hour study period (TOTPAR48) and over the 72-hour study period (TOTPAR72)

(17) TOTPAR up to each evaluation time point

(18) Pain intensity (PI) at each evaluation time point

(19) Pain intensity difference (PID) at each evaluation time point

(20) Pain relief (PR) at each evaluation time point

(21) Patient Usability Questionnaire (PUQ)

(22) Nurse Usability Questionnaire (NUQ)

(23) Total number of study drug doses used over 24, 48, and 72-hour study period, average hourly use, and average inter-dosing interval.

(24) Total amount of supplemental morphine utilized

2.10.2 Analysis Population and Handling of Dropouts

All enrolled patients are those who are enrolled and receive the study treatment. The main analysis of efficacy endpoints includes all enrolled patients.

For patients missing PI or PR data, the following methods will be applied to impute the missing data at evaluation time points for the duration of study period:

(1) Missing data will be first imputed on a patient-by-patient basis by linear interpolation method between two observed pain scale values.

(2) Data occurring after a patient terminated from study or did not provide any follow-up data after last available data prior to the end of study period, the pain scale values at follow-up time points post-termination up to the end of the study period will be imputed on a patient-by-patient basis as described below.

The last observation carried forward (LOCF) method will be used to impute any remaining missing data points after termination due to reasons other than adverse event up to the end of the study period. For patients who prematurely terminate from the study due to adverse event, the worst observation carried forward (WOCF) method will be used to impute the remaining missing data points up to the end of the study period. The worst PID is the smaller number between number zero and the last PID obtained prior to termination. The worst PR is number zero.

For patients who use any supplemental opioid during the study period, the last observed pain intensity prior to using each dose of supplemental opioid will be carried throughout a follow-up 1-hour time interval. Any pain intensity collected within 1 hour after the start of any supplemental opioid will be excluded from the calculation of the time-weighted SPID. This same imputation method will also be used to calculate the efficacy endpoints derived from pain assessment data.

2.10.3 Definition of Baseline Measurements

Baseline measurements are those taken prior to the start of study drug administration.

2.10.4 Methods for the Analysis of Device Dispense Failure Rate

For the analysis of the device dispense failure rate, a point estimate of the device dispense failure rate and its 90% exact Clopper-Pearson confidence interval will be constructed. Patients from all study centers will be pooled for the descriptive summary of these efficacy data.

Results generated from the analysis of these dichotomous efficacy endpoints derived from the device performance data and misplaced tablets will be presented in Post-text Tables 14.2.1 to 14.2.7.

2.10.5 Pooling of Investigators

The data collected from all study centers will be pooled for the analysis of efficacy data.

2.10.6 Methods for the Analysis of Categorical Efficacy Variables

For the analysis of the dichotomous outcome data, a point estimate and its 95% confidence interval will be constructed. Patients from all study centers will be pooled for the descriptive summary of the efficacy data.

Results generated from the analysis of these categorical efficacy variables will be presented in Post-text Tables 14.2.8 to 14.2.18.

2.10.7 Methods for the Analysis of Continuous Efficacy Variables

The method used to derive the efficacy endpoint time-weighted summed pain intensity difference (SPID) over the study period is described in this section. Pain intensity will be measured using an 11-point NRS with 0 (no pain) and 10 (worst possible pain).

Patients will provide the pain intensity (PI) and pain relief (PR) measures at baseline and at each evaluation time point after the initiation of the first on-demand dose of study drug: ¼ [15 min], ½ [30 min], ¾ [45 min], 1, 2, 4, 6, 8, 10, and 12 hours, then every 4 hours at 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64, 68, and 72 hours.

The PID at each evaluation time point after the initiation of the first dose is the difference in pain intensity at the specific evaluation time point and baseline pain intensity [PID (evaluation time after the first dose) = pain intensity (baseline) – pain intensity (evaluation time after the first dose)]. The time-weighted SPID24 is the time-weighted summed PID over the 24-hour study period.

Time-weighted SPID24 = $\sum$ [T(i) – T(i-1)] x PID(i),

Where : T(0) = Time 0 (baseline), T(i) is the scheduled or unscheduled assessment time, and PID(i) is the PID score at time i for i=0 to 24 hours

Other time-weighted SPID, SPID48 and SPID72 will be derived similarly. TOTPAR over follow-up evaluation time periods (e.g., TOTPAR24, TOTPAR48, and TOTPAR72, etc.) will be derived from the PR data collected during the study using similar formula described above.

For the analysis of the continuous efficacy data, a point estimate and its 95% confidence interval will be constructed. Patients from all study centers will be pooled for the descriptive summary of the efficacy data

2.11 Analysis of Safety Data

The safety data collected in this study include:

(1) Adverse events(AEs)
(2) Pregnancy test results
(3) Vital signs
(4) Physical examination data
(5) Suspected technical failures
(6) Concomitant medications

All enrolled patients who receive at least one dose of study drug will be included in the analyses and summaries of safety data.

2.11.1 Extent of Exposure to Study Drug

The length of exposure to study drug will be summarized in Post-text Table 14.3.1.

2.11.2 Adverse Events

Adverse events occurring while patients were on study medication during the treatment period or within 12 hours after the discontinuation of study medication will be summarized in Post-text Table 14.3.2. All incidences of AEs will be summarized in Post-text Table 14.3.3. The MedDRA® (Medical Dictionary for Regulatory Activities) V11.0 thesaurus will be used to map the AE verbatim to lowest level term (LLT), preferred term (PT) and System Organ Class (SOC) for summary purposes. Adverse event data including the AE verbatim term and the associated AE thesaurus preferred term will be provided in the patient data listings.

The most frequent AEs (e.g., those in at least 1% of all treated patients) will be summarized in Post-text Table 14.3.4.

2.11.3 Death Reports and Serious Adverse Events

Death that occurred during the study will be listed and discussed. Serious adverse events will be summarized in Post-text Table 14.3.5.

2.11.4 Adverse Events Causing Premature Discontinuation of Study Drug

Adverse events causing premature discontinuation of study drug will be summarized in Post-text Table 14.3.6.

2.11.5 Summary of Adverse Events by Severity and Relationship to Study Drug

Adverse events will be summarized by severity and relationship to study drug (related or not related) in Post-text Table 14.3.7. Related AEs are those AEs possibly or probably related to study drug. Adverse events will be summarized by maximum severity in Post-text Table 14.3.8 and summarized by closest relationship to the study drug in Post-text Table 14.3.9. Adverse events related to the study drug, severe adverse events, and severe adverse events related to the study drug will be summarized by treatment groups in Post-text Table 14.3.10 and 14.3.12 respectively. 3.6.

2.11.6 Subgroup Summary of Adverse Events

Adverse events will be summarized by age group (< 65 years and ≥ 65 years), sex group (male and female), race group (Caucasian and Non-Caucasian), and body mass index group (< 30 kg/m^2 and ≥ 30 kg/m^2) in Post-text Table 14.3.13 and 14.3.16 respectively. A two-sided Fisher's exact test will be used for the comparison among subgroups.

2.11.7 Pregnancy Test Results

Pregnancy test results will be listed in the Appendix.

2.11.8 Vital Signs

Vital signs, including blood pressure, heart rate, respiratory rates, and oxygen saturation data, taken at baseline and follow-up time points will be summarized in Post-text Tables 14.3.17 to 14.3.21. Patients who had baseline data and at least one follow-up data will be included in this analysis. The descriptive summary statistics will be presented. A paired t-test will be used for the test of mean change from baseline to follow-up time points within each group.

The lowest oxygen saturation value will be identified for each patient. Two dichotomous variables will be defined based on this lowest observed value at two cut-off levels of 93% and 95% of oxygen saturation. These lowest oxygen saturation variables will be summarized. The results of these summaries will be presented in Post-text Table 14.3.22.

2.11.9 Physical Examination Data

The physical examination data will be listed in the Appendix.

2.11.10 Concomitant Medications

Concomitant medications used during the study will be summarized in Post-text Table 14.3.23. Medications discontinued before the enrollment date will be

excluded from the summary. The World Health Organization (WHO) Drug Dictionary (WHO-DD, 01DEC2015 version) will be used to map the medication name to the Anatomical Therapeutic Chemical (ATC) classification codes for summary purpose.

If a medication is mapped to more than one ATC-code, a unique ATC-code will be selected based on the condition treated. The concomitant medications will be summarized by the medication category, medication class, and medication subgroup based on the ATC-codes. The medication category will be determined by the first character of the ATC-code. The medication class will be determined by the first three characters of the ATC-code. The medication subgroup will be determined by all the characters (maximum five characters) of the ATC-code.

2.12 Handling of Partial Dates and Partial Times

Complete date and time values were required for certain date fields on the CRFs (e.g., assessment date and time, dosing date and time, etc.). However, partial date values may have been acceptable for selected date fields on the CRFs (e.g., start dates of concomitant medications or adverse events prior to entry). In order to calculate the duration of events, the following assumptions were applied to estimate the complete date and time values for the partial event start date and time values:

(1) If only the "time" value was missing, the first minute (00:01) of the day was used.

(2) If only the "day" value was missing, the first day of the month was used.

(3) If only the "month" value was missing, January was used.

(4) If both "day" and "month" values were missing, January first was used.

The "year" value of the date field was required for all cases. The estimated event date/time should not be later than the available event stop date/time for a given event.

For those adverse events with partial AE stop date/time and stopped prior to the study termination, the following assumptions were applied to estimate the complete date values:

(1) If only the "time" value was missing, the last minute (23:59) of the day was used.

(2) If only the "day" value was missing, the last day of the month was used.

(3) If only the "month" value was missing, December was used.

(4) If both "day" and "month" values are missing, December 31st was used.

If the estimated AE stop date/time based on the assumptions listed above was later than the study termination date, the partial AE stop date/time was estimated from the study termination date/time.

2.13 The Presentation of Data Listing

In addition to the summary Tables presenting the summarized data, individual

patient response data for the efficacy analysis will be presented in Appendix 16.2. These by-patient data listings present the measurements that will be used in supporting the summary Tables. Other by-patient data listings of baseline characteristics and safety measurements will also be presented in Appendix 16.2.

2.14 Computer Software

All data analyses will be performed using SAS® for Windows, Release 9.1[2].

References

1. Johnson, N. L. and Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions — 1, New York: John Wiley & Sons.
2. SAS Release 9.1. SAS Institute Inc., Cary, North Carolina, 2004.

Table A-1 Evaluation Time Points with Evaluation Windows Used for the Summary of Global Assessment Data by Evaluation Time Point

Evaluation Time Point	Evaluation Time Point Number	Minutes from Baseline to Target Time	Evaluation Window (Minute)		
Baseline	0	1	-14	-	1
2 hours	1	121	2	-	181
4 hours	2	241	182	-	361
8 hours	3	481	362	-	721
16 hours	4	961	722	-	1201
24 hours	5	1441	1202	-	1681
32 hours	6	1921	1682	-	2161
40 hours	7	2401	2162	-	2641
48 hours	8	2881	2642	-	3121
56 hours	9	3361	3122	-	3601
64 hours	10	3841	3602	-	4081
72 hours	11	4321	4082	-	4561
80 hours	12	4801	4562	-	5041
Interim	x.x			---	

Note: Unscheduled evaluation time points and time points for premature termination are considered as regularly scheduled time points. The time point closest to the target time will be used to determine the evaluation time-specific measurement. If there are two measures that obtained at the same time distance away from the target time for the same evaluation time point number, the one prior to the target time will be used in the summary. Measures that do not fall within the evaluation windows will be considered as interim data and assigned a decimal evaluation time point number between two evaluation time point numbers in integers (e.g., 2.1 or 3.9).

Table A-2 Scheduled Evaluation Time Points with Evaluation Windows Used for the Summary of Pain Assessment Data, Blood Pressures, and Heart Rates by Evaluation Time Point

Evaluation Time Point	Evaluation Time Point Number	Minutes from Baseline to Target Time	Evaluation Window (Minute)		
Screening	-1	-15	≤ -15		
Baseline	0	1	-14	-	1
15 minutes	1	16	2	-	23
30 minutes	2	31	24	-	38
45 minutes	3	46	39	-	53
1 hour	4	61	54	-	91
2 hours	5	121	92	-	181
4 hours	6	241	182	-	301
6 hours	7	361	302	-	421
8 hours	8	481	422	-	541
10 hours	9	601	542	-	661
12 hours	10	721	662	-	841
16 hours	11	961	842	-	1081
20 hours	12	1201	1082	-	1321
24 hours	13	1441	1322	-	1561
28 hours	14	1681	1562	-	1801
32 hours	15	1921	1802	-	2041
36 hours	16	2161	2042	-	2281
40 hours	17	2401	2282	-	2521
44 hours	18	2641	2522	-	2761
48 hours	19	2881	2762	-	3001
52 hours	20	3121	3002	-	3241
56 hours	21	3361	3242	-	3481
60 hours	22	3601	3482	-	3721
64 hours	23	3841	3722	-	3961

68 hours	24	4081	3962	-	4201
72 hours	25	4321	4202	-	4441
76 hours	26	4561	4442	-	4681
Interim	x.x			---	

Note: Unscheduled evaluation time points and time points for premature termination are considered as regularly scheduled time points. The time point closest to the target time will be used to determine the evaluation time-specific measurement. If there are two measures that obtained at the same time distance away from the target time for the same evaluation time point number, the one prior to the target time will be used in the summary. Measures that do not fall within the evaluation windows will be considered as interim data and assigned a decimal evaluation time point number between two evaluation time point numbers in integers (e.g., 2.1 or 3.9).

Table A-3 Scheduled Evaluation Time Points with Evaluation Windows Used for the Summary of Respiratory Rates and Oxygen Saturation Data by Evaluation Time Point

Evaluation Time Point	Evaluation Time Point Number	Minutes from Baseline to Target Time	Evaluation Window (Minute)		
Screening	-1	-15	≤ -15		
Baseline	0	1	-14	-	1
15 minutes	1	16	2	-	23
30 minutes	2	31	24	-	38
45 minutes	3	46	39	-	53
1 hour	4	61	54	-	76
1.5 hours	5	91	77	-	106
2 hours	6	121	107	-	136
2.5 hours	7	151	137	-	166
3 hours	8	181	167	-	196
3.5 hours	9	211	197	-	226
4 hours	10	241	227	-	256
4.5 hours	11	271	257	-	286
5 hours	12	301	287	-	316
5.5 hours	13	331	317	-	346
6 hours	14	361	347	-	376
6.5 hours	15	391	377	-	406
7 hours	16	421	407	-	436
7.5 hours	17	451	437	-	466
8 hours	18	481	467	-	496
8.5 hours	19	511	497	-	526
9 hours	20	541	527	-	556
9.5 hours	21	571	557	-	586
10 hours	22	601	587	-	616
10.5 hours	23	631	617	-	646

11 hours	24	661	647	-	676
11.5 hours	25	691	677	-	706
12 hours	26	721	707	-	736
12.5 hours	27	751	737	-	766
13 hours	28	781	767	-	796
13.5 hours	29	811	797	-	826
14 hours	30	841	827	-	856
14.5 hours	31	871	857	-	886
15 hours	32	901	887	-	916
15.5 hours	33	931	917	-	946
16 hours	34	961	947	-	976
16.5 hours	35	991	977	-	1006
17 hours	36	1021	1007	-	1036
17.5 hours	37	1051	1037	-	1066
18 hours	38	1081	1067	-	1096
18.5 hours	39	1111	1097	-	1126
19 hours	40	1141	1127	-	1156
19.5 hours	41	1171	1157	-	1186
20 hours	42	1201	1187	-	1216
20.5 hours	43	1231	1217	-	1246
21 hours	44	1261	1247	-	1276
21.5 hours	45	1291	1277	-	1306
22 hours	46	1321	1307	-	1336
22.5 hours	47	1351	1337	-	1366
23 hours	48	1381	1367	-	1396
23.5 hours	49	1411	1397	-	1426
24 hours	50	1441	1427	-	1501
26 hours	51	1561	1502	-	1621
28 hours	52	1681	1622	-	1741
30 hours	53	1801	1742	-	1861

32 hours	54	1921	1862	-	1981
34 hours	55	2041	1982	-	2101
36 hours	56	2161	2102	-	2221
38 hours	57	2281	2222	-	2341
40 hours	58	2401	2342	-	2461
42 hours	59	2521	2462	-	2581
44 hours	60	2641	2582	-	2701
46 hours	61	2761	2702	-	2821
48 hours	62	2881	2822	-	2941
50 hours	63	3001	2942	-	3061
52 hours	64	3121	3062	-	3181
54 hours	65	3241	3182	-	3301
56 hours	66	3361	3302	-	3421
58 hours	67	3481	3422	-	3541
60 hours	68	3601	3542	-	3661
62 hours	69	3721	3662	-	3781
64 hours	70	3841	3782	-	3901
66 hours	71	3961	3902	-	4021
68 hours	72	4081	4022	-	4141
70 hours	73	4201	4142	-	4261
72 hours	74	4321	4262	-	4441
76 hours	75	4561	4442	-	4681
Interim	x.x			---	

Note: Unscheduled evaluation time points and time points for premature termination are considered as regularly scheduled time points. The time point closest to the target time will be used to determine the evaluation time-specific measurement. If there are two measures that obtained at the same time distance away from the target time for the same evaluation time point number, the one prior to the target time will be used in the summary. Measures that do not fall within the evaluation windows will be considered as interim data and assigned a decimal evaluation time point number between two evaluation time point numbers in integers (e.g., 2.1 or 3.9).