Standard Methods for Analysis and Reporting of VAS or NRS Derived Pain Relief Response Scores

James R. Johnson, PhD
Sr. Director – Biostatistics and Data Sciences
Flexion Therapeutics, Inc.

Presentation at PhUSE 2016
Barcelona, Spain
Paper: SP01

Pain and Pain Intensity Instruments

Pain is a Subjective Experience composed of two complementary features:

- Localized sensation afflicting a particular part of the body
- An unpleasant quality of varying degrees of severity associated with behavior and treatments directed at relieving the pain experience (Pain Relief)

"Divine is the task to relieve pain."
-Hippocrates

Two main Pain Classifications:

- Acute pain (<3 months duration)
- Chronic pain (>3 months)

Clinical Trials Designed for Assessment of PR (Analgesics)

Trial Design Considerations:

- **Design must consider that Pain is a subjective response and fluctuates over time.**
  - Acute pain, post-operative, trials: PI decreases rapidly over days
  - Chronic Pain, Osteoarthritis, trials: PI may decrease slowly over time
- **High Placebo Response Rates are evident in Analgesic trials**
  - High drop-out rates should be expected.
  - Drop-out rates likely to be associated with lack of efficacy (Chronic trials) or adverse events (Chronic and Acute trials).
    - These are Non-random dropout patterns.
    - Must make every effort to minimize drop-outs
- **Rescue Medication used to minimize dropouts from lack of efficacy**
- **Many trial designs used:**
  - Parallel, cross-over, Add-on designs (adjunctive analgesic therapies)
  - Titration to effect designs and enrichment designs
  - Examination of Single-Dose and/or Multiple-Dose Characteristics

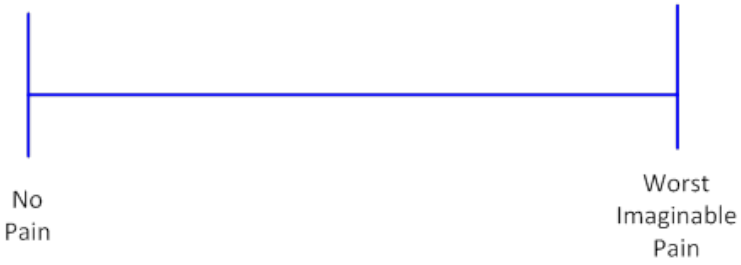
Pain Intensity Instruments: NRS and VAS

Numerical Rating Scale (NRS)

- 11-point scale, 0 to 10
- 0 = No Pain, 10 = Worst Imaginable Pain
- PI recorded in increments of 1 between 0 and 10

Visual Analog Scale (VAS)

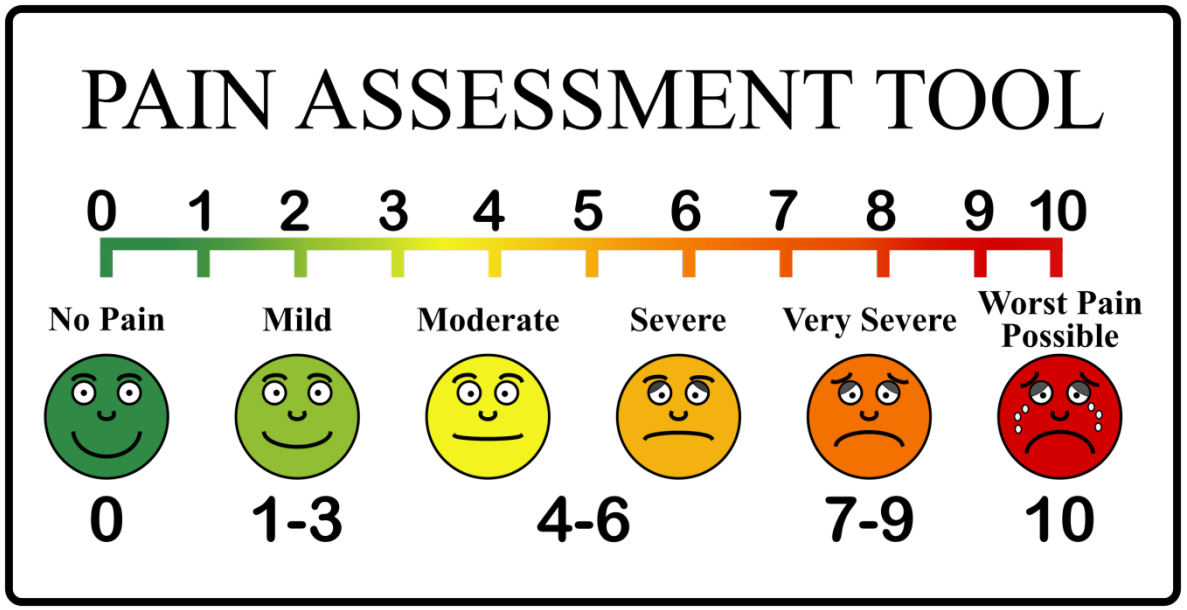
- 100 millimeter (mm) scale
- 0 mm = No Pain, 100 mm = Worst Imaginable pain
- PI measured in mm, rounded to nearest 1mm unit, Continuous between 0 and 100 mm



Variations of the NRS

Numerical Rating Scale (NRS)

- 11-point scale, 0 to 10
- 0 = No Pain, 10 = Worst Possible Pain
- Increments of Mild, Moderate, Severe, and Very Severe specified.
- PI recorded in increments of 1 between 0 and 10
- Modified for pediatric usage



NRS vs VAS Precision: Which Instrument to Use?

NRS: 11-point scale, 0-10

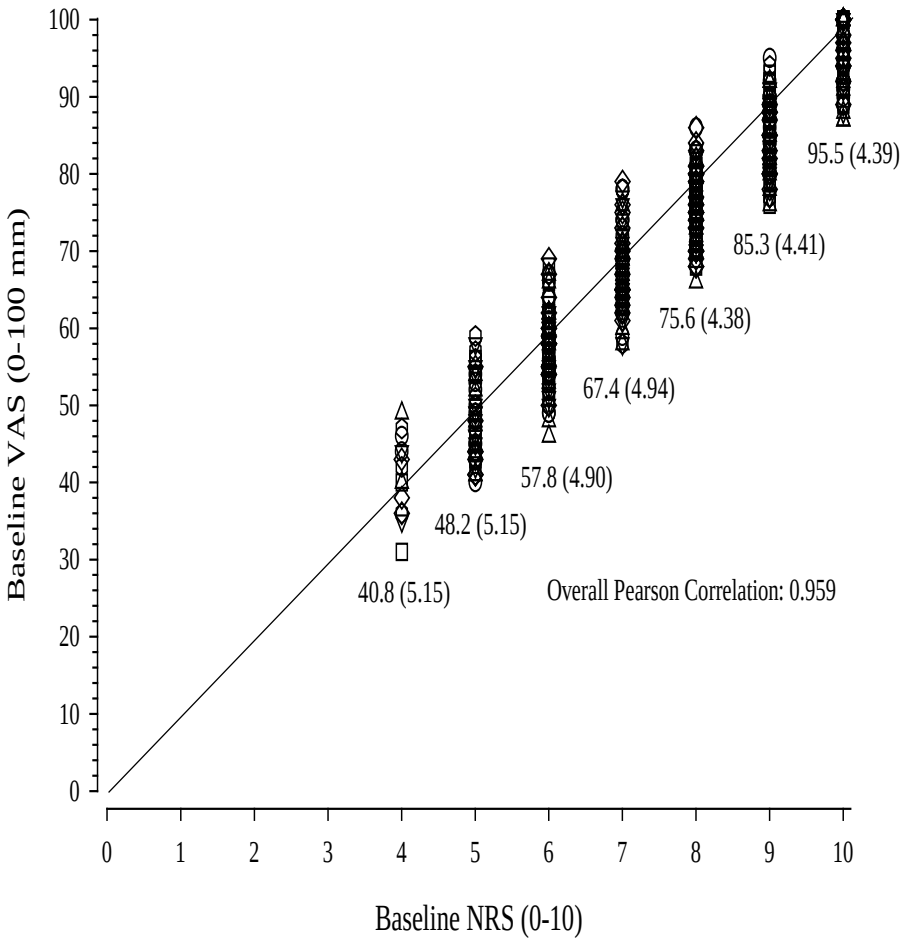
- Ordinal scale with 0=No Pain and 10-Worst Imaginable pain

VAS: 0-100 millimeter scale

- Continuous scale with 0 mm = No Pain, 100 mm = Worst Imaginable Pain. Measured and Rounded to increments of 1 mm.

Precision in Measurement

- Better correlation between NRS and VAS at anchors 0 and 10
- Small differences in VAS can have profound effects on PI scores and PID endpoints.



Timing of PI Assessments

Acute Pain Management Trials:

- Shorter Term duration generally <3 months in Pain duration.
- Pain associated with Injury, post-operative procedures, or short-term idiopathic conditions.
- Medications often administered multiple times per day, with pharmacokinetic profiles that warrant repeat dosing.
- Time points associated with single dose administration generally range up to 72 hours, with greater sampling intensity at early onset times.
    - Early onset time associated with the PK of the medication.

Timing of PI Assessments

Chronic Pain Management Trials:

- Longer Term duration generally >3 months in Pain duration.
- Pain associated with syndromes (cancer, fibromyalgia, Osteoarthritis, chronic migraine), or long-term idiopathic conditions.
- Medications often administered multiple times per day for long periods of time (up to a year or longer), with pharmacokinetic profiles that warrant repeat dosing.
- Time points associated with multiple dose Rx administration generally range up to 1 week, with greater sampling intensity at early onset times for the first dose.
  - Early onset time associated with the PK of the medication for the first dose.
- Time points associated long term duration measured in Days and Weeks (e.g. daily PI assessment for 12 weeks).
  - Long term sampling schema associated with repeat dosing and assessing impact of medication at “steady state”.
  - Single dose chronic pain studies examining extent and duration of PR response following single Rx administration.
    - Usually associated with Extended Release medications or medications with very long half-life or residence times as part of their PK profile.

Missing PI Assessments

Imputation for Missing Data is needed for Computing some endpoints:

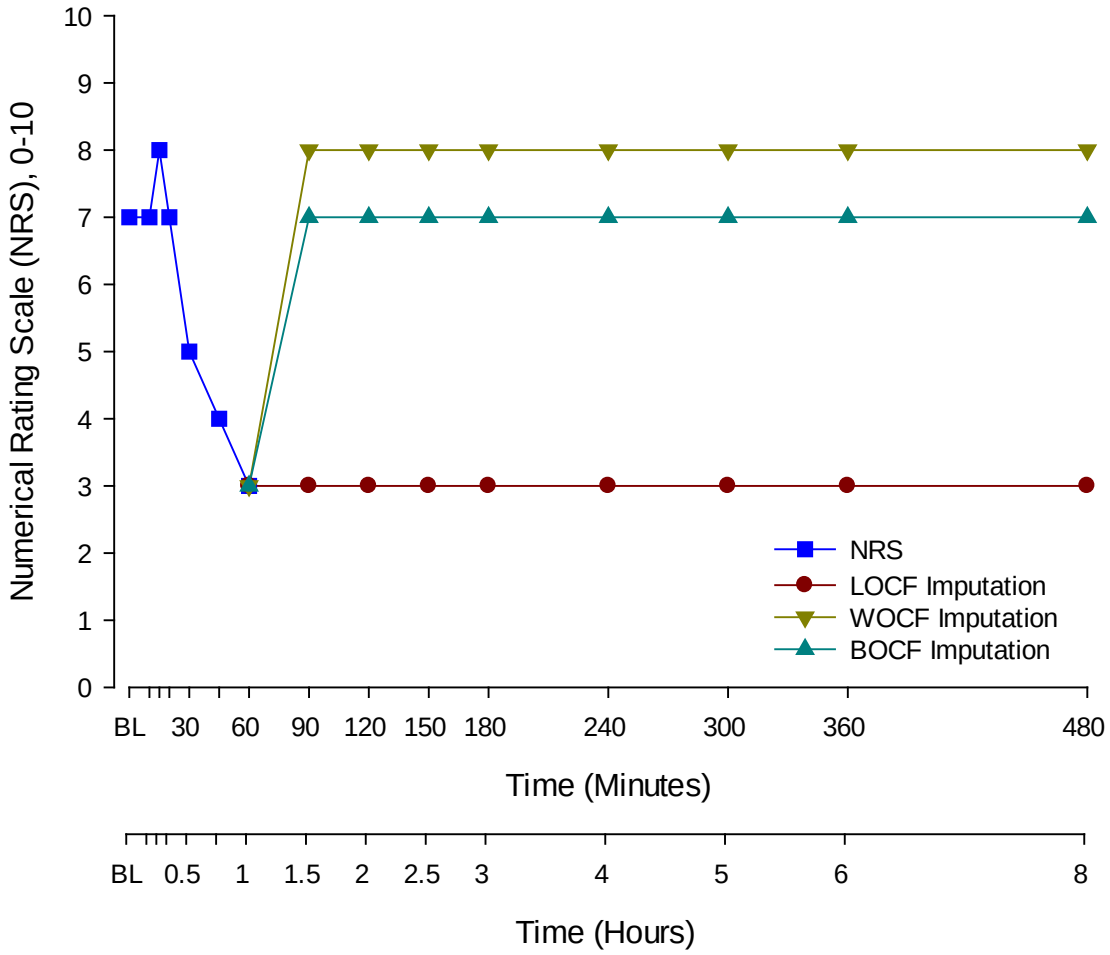
- Common include, LOCF, BOCF, WOCF

Time (min)	0 (BL)	10	15	20	30	45	60	90	120	150	180	240	300	360	480
PI (0-10)	7	7	8	7	5	4	3	Missing	Missing	Missing	Missing	Missing	Missing	Missing	Missing
LOCF	7	7	8	7	5	4	3	3	3	3	3	3	3	3	3
WOCF	7	7	8	7	5	4	3	8	8	8	8	8	8	8	8
BOCF	7	7	8	7	5	4	3	7	7	7	7	7	7	7	7

- Calculation of static endpoints: SPID, AUE will utilize imputation methods for sensitivity purposes.
- The use of LOCF has many statistical difficulties and should be avoided as a primary method for imputation of missing data.
- No one method for imputation should be used. Also consider Multiple Imputation methods.
- Wherever possible “Observed Cases” is the preferred method with no imputation.

Missing PI Assessments

- Consider the impact of imputation methods on an individual subjects PR profile.



Rescue Medication Adjustment

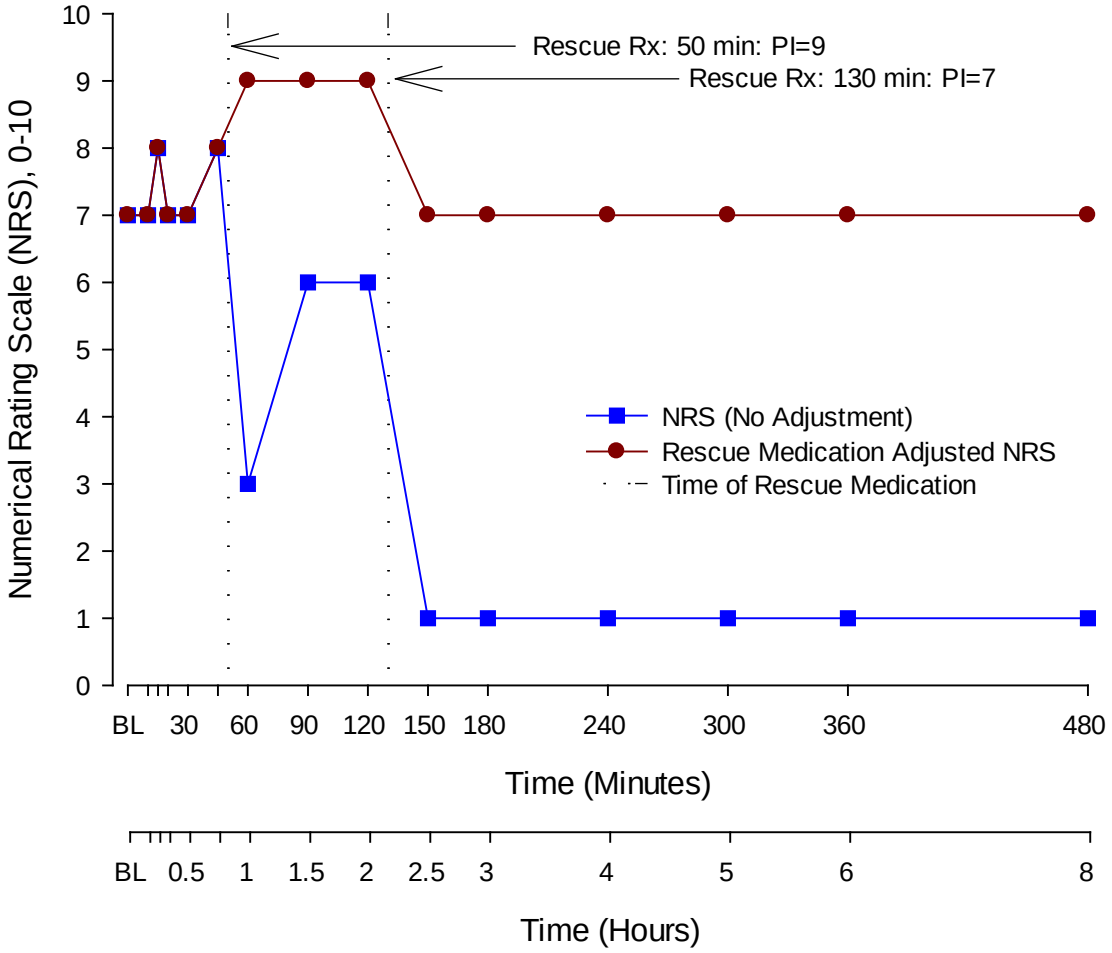
Imputation of “Rescue Medication Adjusted” PI scores

Time (min)	0 (BL)	10	15	20	30	45	60	90	120	150	180	240	300	360	480
PI (0-10)	7	7	8	7	7	8	3	6	6	1	1	1	1	1	1
Rescue Time						50			130						
Pre-Rescue PI						9			7						
Rescue Adjusted PI	7	7	8	7	7	8	9	9	9	7	7	7	7	7	7

- In Acute Pain trials adjustment of PI scores for “Pre-Rescue” PI assessment provides method for least bias in calculation of efficacy endpoints.
- Preferred method for implementation in Acute Pain management trials.

Rescue Medication Adjustment

- Consider the impact of “Rescue Medication Adjusted’ imputation methods on an individual subjects PR profile.



Efficacy Endpoints: Three Common Endpoints

Pain Intensity Difference:

- Two methods for calculation, depending upon direction:

$$(1)\ PID_t = PI_{baseline} - PI_t \quad \text{or} \quad (2)\ PID_t = PI_t - PI_{baseline}$$

Time-Weighted Sum Pain Intensity Differences (SPID):

- Serial assessments of PI over time, weighted by time differences

$$SPID_{t_i - t_{i+n}} = \sum_{t_i}^{t_{i+n}} (PID_i) * (t_{i+1} - t_i)$$

Sum Pain Intensity Differences Area Calculation (AUE):

- Linear trapezoid calculation of an area under PID

$$AUE_{t_i - t_{i+n}} = \sum_{t_i}^{t_{i+n}} ((PID_i + PID_{i+1})/2) * (t_{i+1} - t_i)$$

SPID and AUE Calculations: Rescue Medication Adjustment

Single Subject Example:

	BL	10 Mins.	15 Mins.	20 Mins.	30 Mins.	45 Mins.	1 Hr.	1.5 Hrs.	2 Hrs.	2.5 Hrs.	3 Hrs.	4 Hrs.	5 Hrs.	6 Hrs.	8 Hrs.
Time (minutes)	0	10	15	20	30	45	60	90	120	150	180	240	300	360	480
TD min (t_{i+1} - t_i)		10	5	5	10	15	15	30	30	30	30	60	60	60	120
Rescue Adjusted Pain Intensity (0-10)[1]	9	9	7	6	5	2	2	2	1	1	1	3	5	5	5
PID (t_i-BL)	0	0	-2	-3	-4	-7	-7	-7	-8	-8	-8	-6	-4	-4	-4
PIDi*(t_{i+1}-t_i)		0	-10	-15	-40	-105	-105	-210	-240	-240	-240	-360	-240	-240	-480
(PID_{i+1}+PID_i)/2		0	-1	-2.5	-3.5	-5.5	-7	-7	-7.5	-8	-8	-7	-5	-4	-4
[(PID_i+PID_{i+1})/2]*TD min		0.0	-5.0	-12.5	-35.0	-82.5	-105.0	-210.0	-225.0	-240.0	-240.0	-420.0	-300.0	-240.0	-480.0

- Endpoints:

Efficacy Endpoint	Value
Time Weighted $SPID_{0-180}$	-1205
Time Weighted $SPID_{0-360}$	-2045
Time Weighted $SPID_{0-480}$	-2525

Efficacy Endpoint	Value
AUE_{0-180}	-1155
AUE_{0-360}	-2115
AUE_{0-480}	-2595

Analysis of PID Endpoint

Completed with a longitudinal Mixed Model for Repeated Measures (MMRM)

```
proc mixed data=<ADNRS> method=reml;
  where avisit) ^= 'baseline' and paramcd="PIDRA";
  class usubjid trt01p avisitn;
  model chg = base trt01p avisitn trt01p*avisitn / DDFM=KR;
  repeated avisitn / type = un subject=usubjid;
  lsmeans trt01p / diff=control ('Placebo') cl;
  lsmeans trt01p*avisitn / pdiff cl;
  ods output lsmeans = lsmeans
              diffs = diffs
             Tests3 = test;
run;
```

Best Implemented with standardized ADaM datasets: (topic for next paper)

Analysis of PID Endpoint: Residuals Examination

Completed with a longitudinal Mixed Model for Repeated Measures (MMRM)

```
proc mixed data=<ADNRS> method=reml;
  where avisit) ^= 'baseline' and paramcd="PIDRA";
  class usubjid trt01p avisitn;
  model chg = base trt01p avisitn trt01p*avisitn / DDFM=KR
              out=PRED1 residual solution;
  repeated avisitn / type = un subject=usubjid;
  lsmeans trt01p / diff=control ('Placebo') cl;
  lsmeans trt01p*avisitn / pdiff cl;
  ods output lsmeans = lsmeans
             diffs = diffs
             Tests3 = test;
run;
```

Evaluating the standardized residuals is a very useful technique for detection of outliers: Generally expect 95% of the standardized residuals will be ± 2 SD. Ideal technique for understating impact of covariates

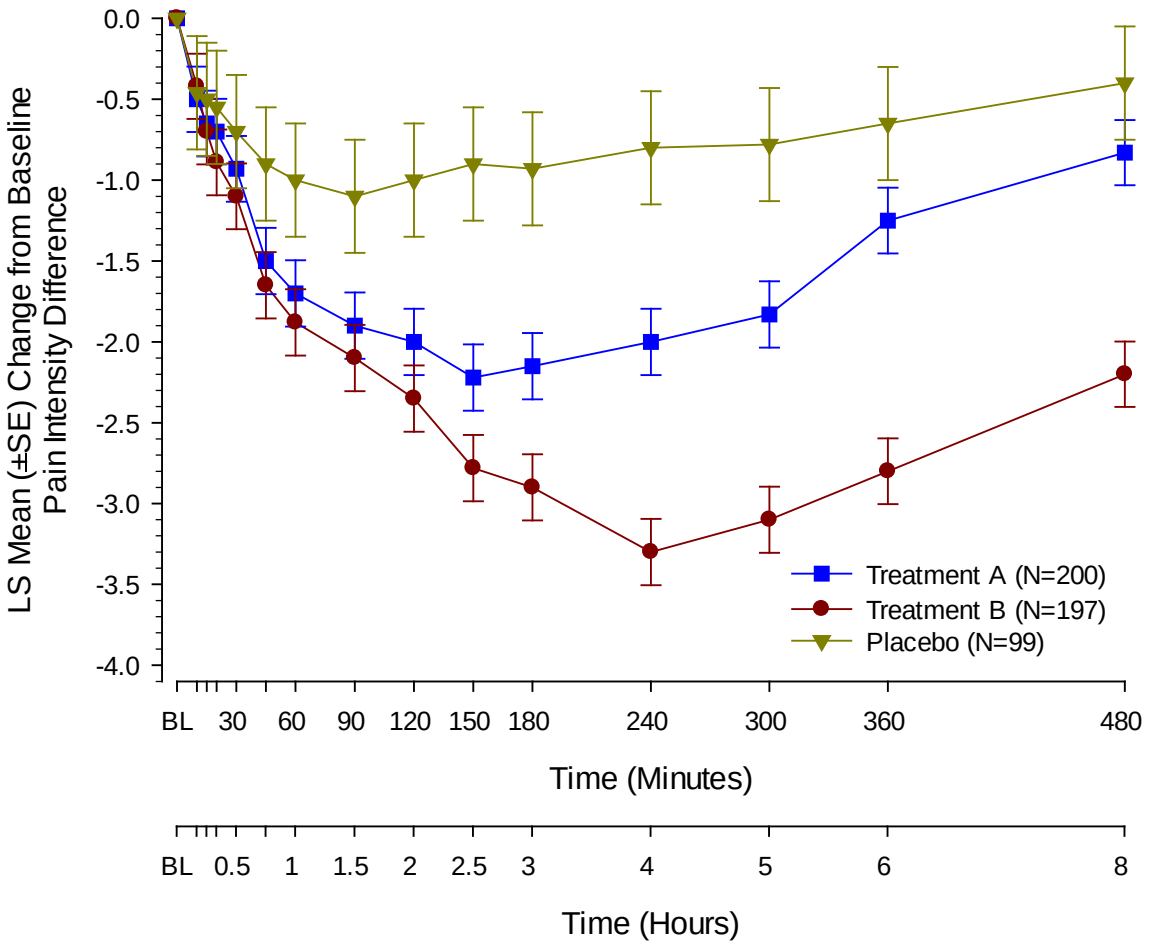
Analysis of SPID and AUE Endpoints

Completed with a Linear Models (ANOVA)

```
proc mixed data=<ADEFF>;
  where paramcd = "<endpoint code>;
  class trt01p;
  model aval = trt01p base / ddfm=kr;
  lsmeans trt01p / pdiff cl;
  estimate 'Treatment A v Placebo' trt01p -1  0 1 / cl alpha=0.05;
  estimate 'Treatment B v Placebo' trt01p  0 -1 1 / cl alpha=0.05;
  ods output lsmeans = lsmeans
             diffs = diffs
            Tests3 = test
         Estimates = est;
run;
```

Best Implemented with standardized ADaM datasets: (topic for next paper)

Recommended Standard Displays: LSM Change from Baseline



Recommend Standard Displays: PID Tabular Summaries

	Treatment A (N=xxx)		Treatment B (N=xxx)		Placebo (N=xxx)	
Time Point / Statistic	Observed	Change from Baseline	Observed	Change from Baseline	Observed	Change from Baseline
Baseline						
n	xxx		xxx		xxx	
Mean (SD)	xxx.x (xx.xx)		xxx.x (xx.xx)		xxx.x (xx.xx)	
Median	xxx.x		xxx.x		xxx.x	
Min, Max	xxx, xxx		xxx, xxx		xxx, xxx	
<Time Point>	xxx	xxx	xxx	xxx	xxx	xxx
n	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)
Mean (SD)	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x
Min, Max	xxx, xxx	xxx, xxx	xxx, xxx	xxx, xxx xxx.x	xxx, xxx	xxx, xxx xxx.x
LSM (SE)		xxx.x (xx.xx)		(xx.xx)		(xx.xx)
95% CI		xxx.x, xxx.x		xxx.x, xxx.x		xxx.x, xxx.x
LSM Difference from Placebo		xx.xx		xx.xx		
95% CI		xx.xx, xx.xx		xx.xx, xx.xx		
p-value, 2-sided		0.xxxx		0.xxxx		

Note: LSM (SE), mean difference from placebo, CI and p-values from mixed model, modeling Pain Intensity Difference from baseline with fixed effects of Treatment, time point, treatment by time interaction, and model covariates of baseline PI score and <covariates>

<Other Footnotes>

Programming Note:

- Display one time point per page for clarity of analysis
- Additional descriptive statistics may include %CV or Interquartile ranges if needed. Insert on separate lines

Recommend Standard Displays: SPID and AUE Summaries

Endpoint / Statistic	Treatment A (N=xxx)	Treatment B (N=xxx)	Placebo (N=xxx)
<Endpoint>	xxx	xxx	xxx
n	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)
Mean (SD)	xxx.x	xxx.x	xxx.x
Min, Max	xxx, xxx	xxx, xxx	xxx, xxx
LSM (SE)	xxx.x (xx.xx)	xxx.x (xx.xx)	xxx.x (xx.xx)
95% CI	xxx.x, xxx.x	xxx.x, xxx.x	xxx.x, xxx.x
LSM Difference from Placebo	xx.xx	xx.xx	
95% CI	xx.xx, xx.xx	xx.xx, xx.xx	
p-value, 2-sided	0.xxxx	0.xxxx	

Note: LSM (SE), mean difference from placebo, CI and p-values from linear model (ANOVA), modeling <Efficacy Endpoint> Difference from baseline with fixed effects of Treatment, and model covariates of baseline PI score and <covariates>

<Other Footnotes>

Programming Note:

- Display one efficacy endpoint per page for clarity of analysis
- Additional descriptive statistics may include %CV or Interquartile ranges if needed. Insert on separate lines

Conclusions and Recommendations

- NRS and VAS are powerful instruments for assessing PI, choose carefully based on the design and characteristics of the Rx under development.
- Methods for handling missing PI data has been the topic of many peer reviewed articles and appropriate methods are defined in the literature.
  - Must adjust for Rescue Medication usage. Rescue adjusted PI score follows best practice on how to handle these data.
- Efficacy endpoints (PID, SPID, AUE) well established and validated in the literature. Accepted by regulatory agencies.
  - Statistical models can be standardized for these analyses.
- No formal standards for presentation of these data.
  - Proposed methods for standardizing the presentation (graphical and tabular) of endpoints.
  - Propose development of PhUSE CSS sponsored White paper for Pain Endpoints.
    - Standard TLF approaches
    - Standard STDM and ADaM data sets for Pain data and endpoints.