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)

Two main Pain Classifications:
- Acute pain (<3 months duration)
- Chronic pain (>3 months)
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

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0 (BL)</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>90</th>
<th>120</th>
<th>150</th>
<th>180</th>
<th>240</th>
<th>300</th>
<th>360</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI (0-10)</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td>LOCF</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>WOCF</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>BOCF</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

- 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 subject's PR profile.

![Graph showing numerical rating scale (NRS) over time, with comparisons between LOCF, WOCF, and BOCF imputations.]
Rescue Medication Adjustment

Imputation of “Rescue Medication Adjusted” PI scores

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0 (BL)</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>90</th>
<th>120</th>
<th>150</th>
<th>180</th>
<th>240</th>
<th>300</th>
<th>360</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI (0-10)</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Rescue Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Rescue PI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rescue Adjusted PI</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

- 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.
Consider the impact of “Rescue Medication Adjusted” imputation methods on an individual subject's PR profile.
Efficacy Endpoints: Three Common Endpoints

Pain Intensity Difference:
• Two methods for calculation, depending upon direction:

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

Time-Weighted Sum Pain Intensity Differences (SPID):
• Serial assessments of PI over time, weighted by time differences

\[SPID_{t_i-t_{i+n}} = \sum_{t_i}^{t_{i+n}} (PID_i) \times (t_{i+1} - t_i)\]

Sum Pain Intensity Differences Area Calculation (AUE):
• Linear trapezoid calculation of an area under PID

\[AUE_{t_i-t_{i+n}} = \sum_{t_i}^{t_{i+n}} ((PID_i + PID_{i+1})/2) \times (t_{i+1} - t_i)\]
SPID and AUE Calculations: Rescue Medication Adjustment

Single Subject Example:

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

- **Endpoints:**
  
<table>
<thead>
<tr>
<th>Efficacy Endpoint</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Weighted SPID$_{0-180}$</td>
<td>-1205</td>
</tr>
<tr>
<td>Time Weighted SPID$_{0-360}$</td>
<td>-2045</td>
</tr>
<tr>
<td>Time Weighted SPID$_{0-480}$</td>
<td>-2525</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficacy Endpoint</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUE$_{0-180}$</td>
<td>-1155</td>
</tr>
<tr>
<td>AUE$_{0-360}$</td>
<td>-2115</td>
</tr>
<tr>
<td>AUE$_{0-480}$</td>
<td>-2595</td>
</tr>
</tbody>
</table>
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 trt0lp avisitn;
  model chg = base trt0lp avisitn trt0lp*avisitn / DDFM=KR;
  repeated avisitn / type = un subject=usubjid;
  lsmeans trt0lp / diff=control ('Placebo') cl;
  lsmeans trt0lp*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=rem1;
  where avist) ^= '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)

```sas
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

<table>
<thead>
<tr>
<th>Time (Minutes)</th>
<th>BL</th>
<th>30</th>
<th>60</th>
<th>90</th>
<th>120</th>
<th>150</th>
<th>180</th>
<th>240</th>
<th>300</th>
<th>360</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain Intensity Difference</td>
<td>-4.0</td>
<td>-3.5</td>
<td>-3.0</td>
<td>-2.5</td>
<td>-2.0</td>
<td>-1.5</td>
<td>-1.0</td>
<td>-0.5</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment A</td>
<td>200</td>
</tr>
<tr>
<td>Treatment B</td>
<td>197</td>
</tr>
<tr>
<td>Placebo</td>
<td>99</td>
</tr>
</tbody>
</table>
### Recommend Standard Displays: PID Tabular Summaries

<table>
<thead>
<tr>
<th>Time Point / Statistic</th>
<th>Treatment A (N=xxx)</th>
<th>Treatment B (N=xxx)</th>
<th>Placebo (N=xxx)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed Change from Baseline</td>
<td>Observed Change from Baseline</td>
<td>Observed Change from Baseline</td>
</tr>
<tr>
<td>Baseline</td>
<td>n xxx</td>
<td>n xxx</td>
<td>n xxx</td>
</tr>
<tr>
<td></td>
<td>Mean (SD) xxx.x (xx.xx)</td>
<td>Mean (SD) xxx.x (xx.xx)</td>
<td>Mean (SD) xxx.x (xx.xx)</td>
</tr>
<tr>
<td></td>
<td>Median xxx.x</td>
<td>Median xxx.x</td>
<td>Median xxx.x</td>
</tr>
<tr>
<td></td>
<td>Min, Max xxx, xxx</td>
<td>Min, Max xxx, xxx</td>
<td>Min, Max xxx, xxx</td>
</tr>
<tr>
<td>&lt;Time Point&gt;</td>
<td>n xxx xxx</td>
<td>n xxx xxx</td>
<td>n xxx xxx</td>
</tr>
<tr>
<td></td>
<td>Mean (SD) xxx.x (xx.xx)</td>
<td>Mean (SD) xxx.x (xx.xx)</td>
<td>Mean (SD) xxx.x (xx.xx)</td>
</tr>
<tr>
<td></td>
<td>Min, Max xxx, xxx</td>
<td>Min, Max xxx, xxx</td>
<td>Min, Max xxx, xxx</td>
</tr>
<tr>
<td></td>
<td>LSM (SE) xxx.x (xx.xx)</td>
<td>LSM (SE) (xx.xx)</td>
<td>LSM (SE) (xx.xx)</td>
</tr>
<tr>
<td></td>
<td>95% CI xxx, xxx</td>
<td>95% CI xxx, xxx</td>
<td>95% CI xxx, xxx</td>
</tr>
<tr>
<td></td>
<td>p-value, 2-sided 0.xxx</td>
<td>p-value, 2-sided 0.xxx</td>
<td>p-value, 2-sided 0.xxx</td>
</tr>
</tbody>
</table>

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

**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

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

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>
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, SPIE, 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.