How can we generate reliable data in clinical trials?

Nathaniel Katz, MD, MS
Chief Science Officer
WCG Analgesic Solutions
Adjunct Associate Professor of Anesthesia
Tufts University School of Medicine

Autumn, 2020
Developing new therapeutics requires an accurate and reliable method to measure their efficacy – i.e. reliable clinical trials.

Regulatory guidance (FDA, 2013; ICH 2016; ICH 2019) requires sponsors to take action to ensure the reliability of clinical trial results.

No guidance documents define reliability or provide clear direction on how to achieve it.

ICH, Guideline for good clinical practice E6(R2), Dec 2016

This guideline has been amended to encourage implementation of improved and more efficient approaches to clinical trial design, conduct, oversight, recording and reporting while continuing to ensure human subject protection and reliability of trial results.”
Why did my trial fail?

CONSEQUENCES OF FAILED TRIALS
• Need to repeat studies
• Massive waste of resources
• Premature death of programs
• Exposure of human subjects to risk without gaining scientific knowledge
Reliability:

the quality of being trustworthy or of performing consistently well
What is reliability in science?

<table>
<thead>
<tr>
<th>Reliability</th>
<th>How close repeated measurements are to each other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>How close a measurement is to the true value</td>
</tr>
</tbody>
</table>

ISO 5725, 1994; JCGM, 2012
Accuracy and reliability can be applied to:

- A method for conducting clinical trials
  - Dental pain studies
  - Dental pain studies at a specific site

- The results of a specific trial
  - The results of this study were accurate (close to accepted standard)
  - Reliability is undefined with respect to a single result

- The results of a group of trials
  - The results of dental pain studies performed by Al Sunshine were reliable
  - The results of lamotrigine studies in PDN studies were not reliable

- The performance of critical procedures in a trial
  - This assessment is being performed reliably
  - The results of these assessments are accurate
  - This activity is being performed consistently (e.g. medication adherence)

- Assay sensitivity
  - Differentiation in a trial between active and control
  - Indirect measure of accuracy and reliability of study methods

ISO 5725, 1994
Clinical trials are unreliable and have a high failure risk

- Pregabalin is FDA-approved for fibromyalgia and represented in all treatment guidelines
- Observed effect sizes in similar studies ranged from 0.12 to 0.48
- 3/10 studies failed (slipped below p<.05)
- The coefficient of variation of this set of studies is 40% - unacceptable in any other area
- Failure is common: 53% of Phase 3 trials fail to confirm efficacy observed in Phase 2

Coefficient of variation = 40%
A speedometer with a CoV of 40%:
Replicate trials often do not replicate

Vinik AI et al, Pain, 2007
Why are clinical trials unreliable and prone to failure?

The reliability of clinical trials is determined by the reliability of “critical procedures” in the trials.

- Reporting pain accurately
- Setting level of expectation
- Adherence to study & rescue medication
- Compliance with diaries
- Performing diagnostic assessments
- Consistent study conduct across sites

Katz, 2005; Harris, 2005; Wise, 2009; Dworkin, 2012, 2014; Treister, 2019
Consistency of Pain Reporting Varies from Person to Person: An Experimental Paradigm

Low Variability (good report)

Highly Variable (poor report)
Poor Pain Reporters Cannot Differentiate Naproxen from Placebo

Poor pain reporters: Drug no better than placebo

Good pain reporters: Drug beats placebo

Effect Size (Cohen's $d$)

- All patients (N=32) $p=0.07$
- First quartile (N=8) $p=0.55$
- Second quartile (N=8) $p=0.52$
- Third quartile (N=8) $p=0.43$
- Fourth quartile (N=8) $p=0.02$

Increasing pain reporting accuracy

Training improves pain reporting accuracy (and decreases placebo response)

No change in response to drug

Lower response to placebo

Lower variability

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Untrained</th>
<th>Trained</th>
<th>Untrained</th>
<th>Trained</th>
<th>Variability</th>
<th>Untrained</th>
<th>Trained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response to drug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response to placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Treister R et al, Pain, 2018
Accurate Pain Reporting Training: Basic Principles

**Accuracy**
Mindfully report the most accurate, precise report of your pain every time

**Consistency**
- Use the pain scale the same way every time

**Open-Mindedness**
- Don’t be biased by expectations about your pain

**Specificity**
- Answer the question being asked: location of pain, recall period, pain aspect, etc.

---

Reporting Your Pain
What you need to know for this clinical research study

© Analgesic Solutions 2017
232 Pond Street, Natick, MA 01760 USA
All rights reserved.

Version 4.2 02 MAR 2017
In this trial, patients will be asked to describe their average and worst pain due to diabetic neuropathy by picking a number from 0 to 10.

For example:

“Please rate your pain due to diabetic neuropathy by selecting the number that best describes your pain on the average in the last 24 hours.”

11-point, 0-10 NRS scale
Training Provides Opportunities to Practice Pain Reporting

Patients should refer to their 3, 6, and 9 scores whenever they rate their pain in order to achieve consistency and accuracy.
The Placebo Response: Why is it a Problem?

Increasing Predicts failure Variable

B Placebo Response

The main driver of the placebo response is expectation of benefit by the patients and study staff. “Neutralizing” staff and patient expectation decreases the placebo response, and improves discrimination between active drug and placebo.

Colloca L, Barsky A. New England Journal of Medicine, 2020
In the experiment, investigators expected patients could not get active drug, but placebo patients thought they could. The graph shows the change in pain rating index between baseline (10 min before injection) and 10 and 60 min after administration of placebo.

PN = group that could have either received placebo or naloxone. PNF = group that could have received placebo, naloxone, or fentanyl (PNF).
Positive Information Increases Placebo Effect

Telling patients that placebo was Maxalt doubled the response.

Telling patients that Maxalt was placebo cut the response in half.

Your words may be as powerful as the drug.

NT = No Treatment; P = Placebo; U = Placebo or Maxalt; M = Maxalt

Warmth And Empathy Enhance Placebo Response

Kelley, Psychosomatic Medicine, 2009

Warm and empathic acupuncture providers lead to higher combined outcomes compared to neutral and business-like providers.
How You Act Matters!

Interact with Patients in a **Neutral** Manner

▶ Warmth and empathy enhance the placebo effect
▶ Your expectation of outcome is conveyed by:

- Facial expression
- Voice tone
- Body language
- Words
- Physical contact
- Time spent outside of protocol with patients
Accurate pain reporting and placebo response reduction training experience

- >80 studies
- >70,000 subjects trained
- >15 indications
- Acceptance by regulatory authorities and IRBs

Figure 3. APR and PRR indications

<table>
<thead>
<tr>
<th>Indication</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Traumatic Neuropathic Pain</td>
<td></td>
</tr>
<tr>
<td>Post Herpetic Neuralgia</td>
<td></td>
</tr>
<tr>
<td>Abdominal Surgery</td>
<td></td>
</tr>
<tr>
<td>Trigeminal Neuralgia</td>
<td></td>
</tr>
<tr>
<td>Small Fiber Neuropathy</td>
<td></td>
</tr>
<tr>
<td>Chronic Low Back Pain</td>
<td></td>
</tr>
<tr>
<td>Fibromyalgia</td>
<td></td>
</tr>
<tr>
<td>Interstitial Cystitis</td>
<td></td>
</tr>
<tr>
<td>Diabetic Peripheral Neuropathy</td>
<td></td>
</tr>
<tr>
<td>Migraine</td>
<td></td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td></td>
</tr>
</tbody>
</table>
Combining Accurate Pain Reporting and Placebo Response Reduction Training: a clinical trial in lumbosacral radiculopathy demonstrated a low placebo response vs. published studies.

Backonja et al., *Pain*, 2017
Central statistical monitoring of live clinical trials

Study View

Site View
When you don’t have validated fixed thresholds, use SPC
Case 1: Fixed threshold
E-Diary Compliance <85% per week
Case 2: Statistical threshold
Rescue Medication Consumption
Case 3: Statistical Threshold
Staircase-evoked pain: key efficacy endpoint

StEPP Results by Site

StEPP Results Before and After Retraining
A subject with multiple critical process failures

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of Serious Adverse Events (SAEs)</th>
<th>Pain Discordance</th>
<th>Low Back Pain Intensity (LBP)</th>
<th>Patient Global Assessment</th>
<th>Roland Morris Disability</th>
<th>Joint Pain Questionnaire</th>
<th>Missing Post-Randomization Data (%)</th>
<th>Missing Pre-Randomization Data (%)</th>
<th>Number of Adverse Events</th>
<th>Rescue Medication Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-Jan-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>21-Jan-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>28-Jan-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>4-Feb-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>11-Feb-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>18-Feb-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>25-Feb-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>4-Mar-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>11-Mar-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>18-Mar-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
<tr>
<td>25-Mar-15</td>
<td>0.00</td>
<td>7.30</td>
<td>10.00</td>
<td>2.00</td>
<td>6.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>162.50</td>
</tr>
</tbody>
</table>

Early intervention can potentially correct these issues and lead to better assay sensitivity.
The accuracy and reliability of clinical trials as a method for measuring treatment effects merits quantification, but appears poor.

Reliability of trial results is determined by reliability of critical processes within that trial.

Protocols should include a section on reliability:
- What are the critical processes that impact reliability of study results?
- What procedures will be utilized during the study to monitor reliability of critical processes?
- What corrective actions will be taken to remediate performance issues?

Validated training of subjects and staff, and targeted central statistical monitoring of critical processes, are our major tools to achieve study reliability.
Thank You

nkatz@wcgclinical.com