wcg **ANALGESIC** SOLUTIONS

Improving Clinical Trial Outcomes

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

Introduction

- 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

  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**."

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

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

Why did my trial fail?

ME

Your Trial

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?

Reliability	How close repeated measurements are to each other

Accuracy	How close a measurement is to the true value

Accurate, reliable

Accurate, unreliable

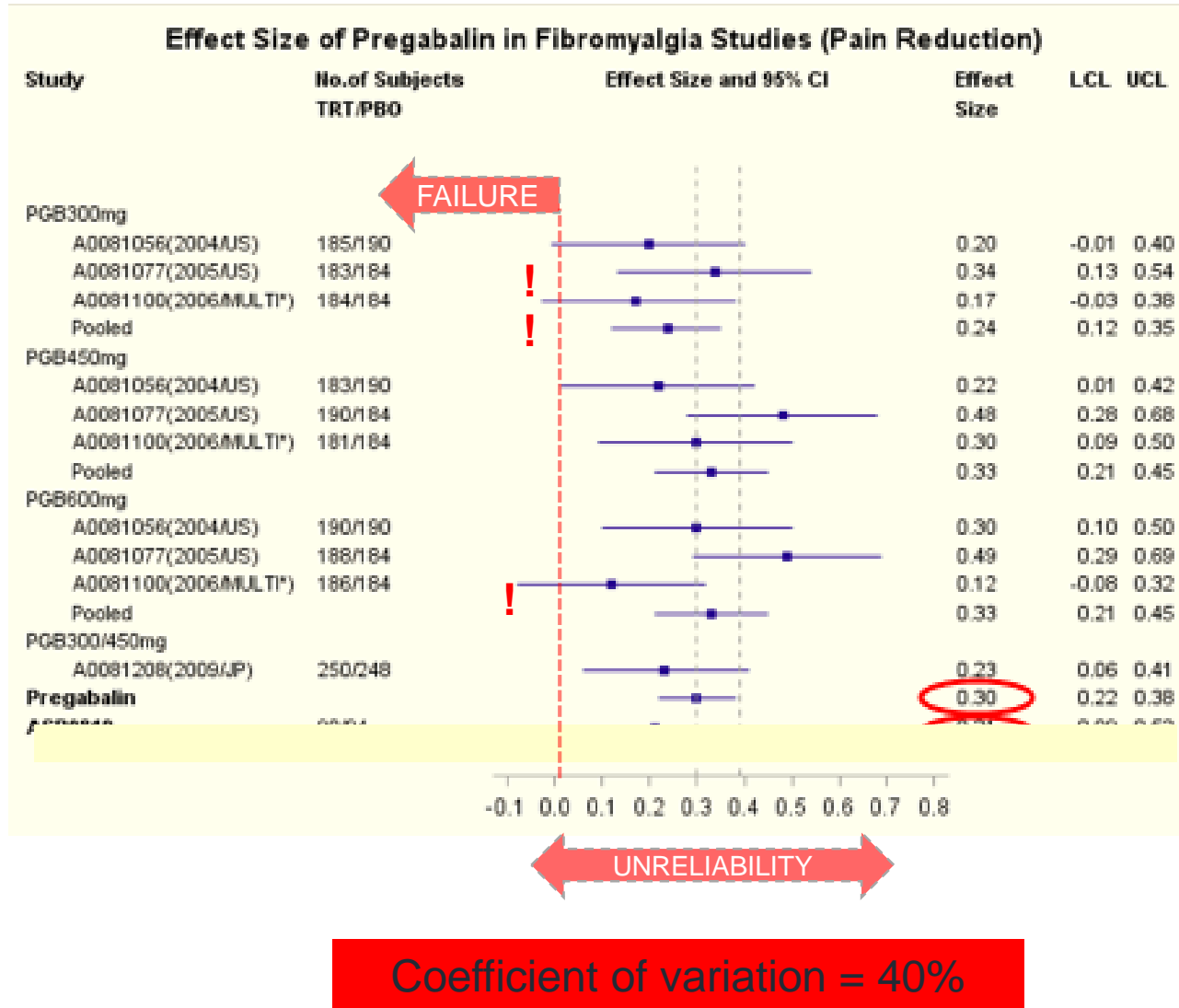Inaccurate, reliable

Inaccurate, unreliable

ISO 5725, 1994; JCGM, 2012

Accuracy and reliability can be applied to:

- A <u>method</u> for conducting clinical trials
  - Dental pain studies
  - Dental pain studies at a specific site
- The <u>results</u> of a <u>specific</u> trial
  - The results of this study were accurate (close to accepted standard)
  - Reliability is <u>undefined</u> with respect to a single result
- The <u>results</u> of a <u>group</u> 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 <u>critical procedures</u> 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

Effect Size of Pregabalin in Fibromyalgia Studies (Pain Reduction)

Study	No.of Subjects TRT/PBO	Effect Size	LCL	UCL
PGB300mg				
A0081056(2004/US)	185/190	0.20	-0.01	0.40
A0081077(2005/US)	183/184	0.34	0.13	0.54
A0081100(2006/MULTI*)	184/184	0.17	-0.03	0.38
Pooled		0.24	0.12	0.35
PGB450mg				
A0081056(2004/US)	183/190	0.22	0.01	0.42
A0081077(2005/US)	190/184	0.48	0.28	0.68
A0081100(2006/MULTI*)	181/184	0.30	0.09	0.50
Pooled		0.33	0.21	0.45
PGB600mg				
A0081056(2004/US)	190/190	0.30	0.10	0.50
A0081077(2005/US)	188/184	0.49	0.29	0.69
A0081100(2006/MULTI*)	186/184	0.12	-0.08	0.32
Pooled		0.33	0.21	0.45
PGB300/450mg				
A0081208(2009/JP)	250/248	0.23	0.06	0.41
Pregabalin		0.30	0.22	0.38

FAILURE

UNRELIABILITY

Coefficient of variation = 40%

- 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

Courtesy of Paul Blahunka, Astellas Inc.

A speedometer with a CoV of 40%:

True speed
55 mph

40 mph

30 mph

80 mph

50 mph

70 mph

Replicate trials often do not replicate
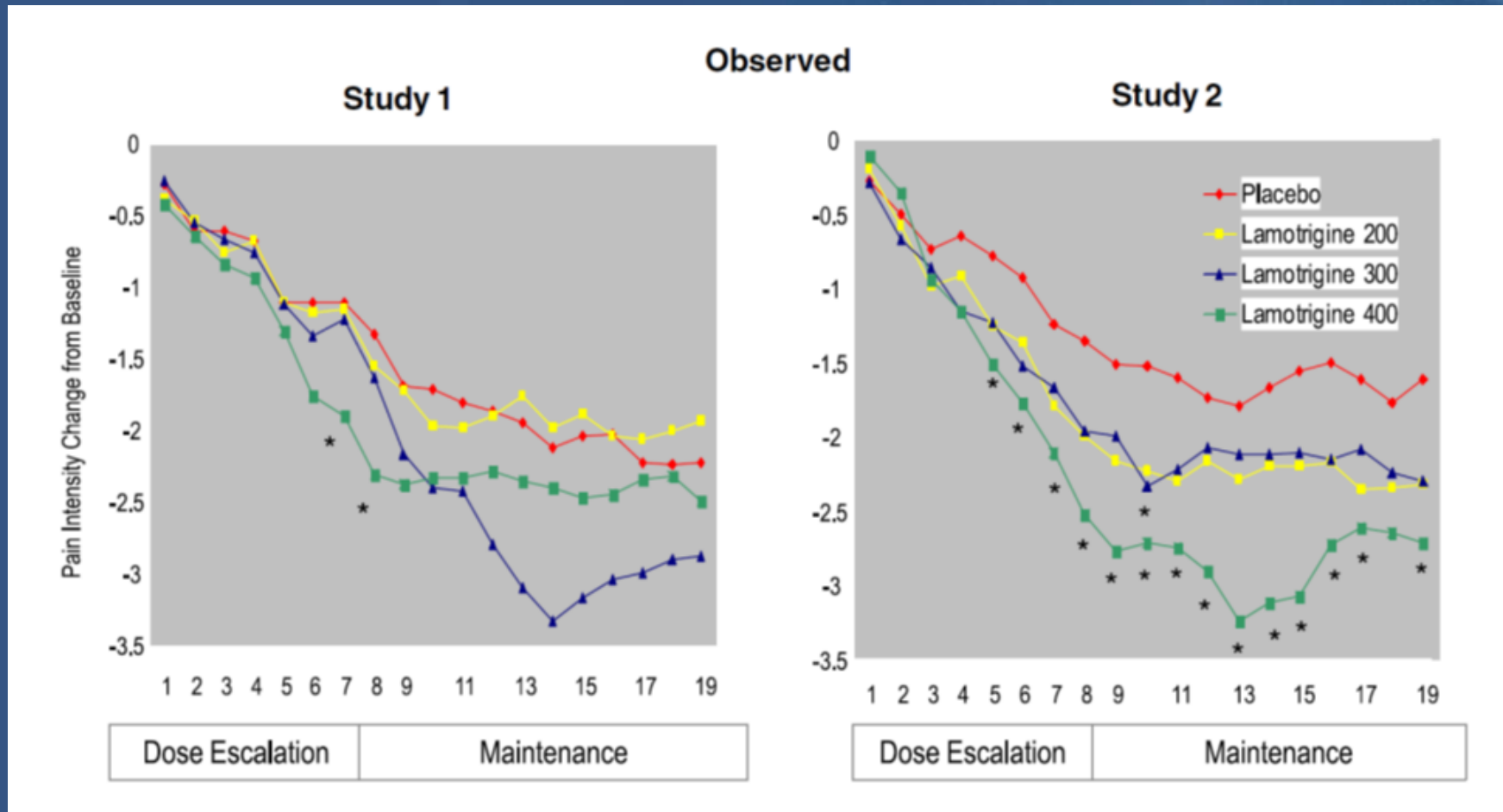

Vinik AI et al, Pain, 2007

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
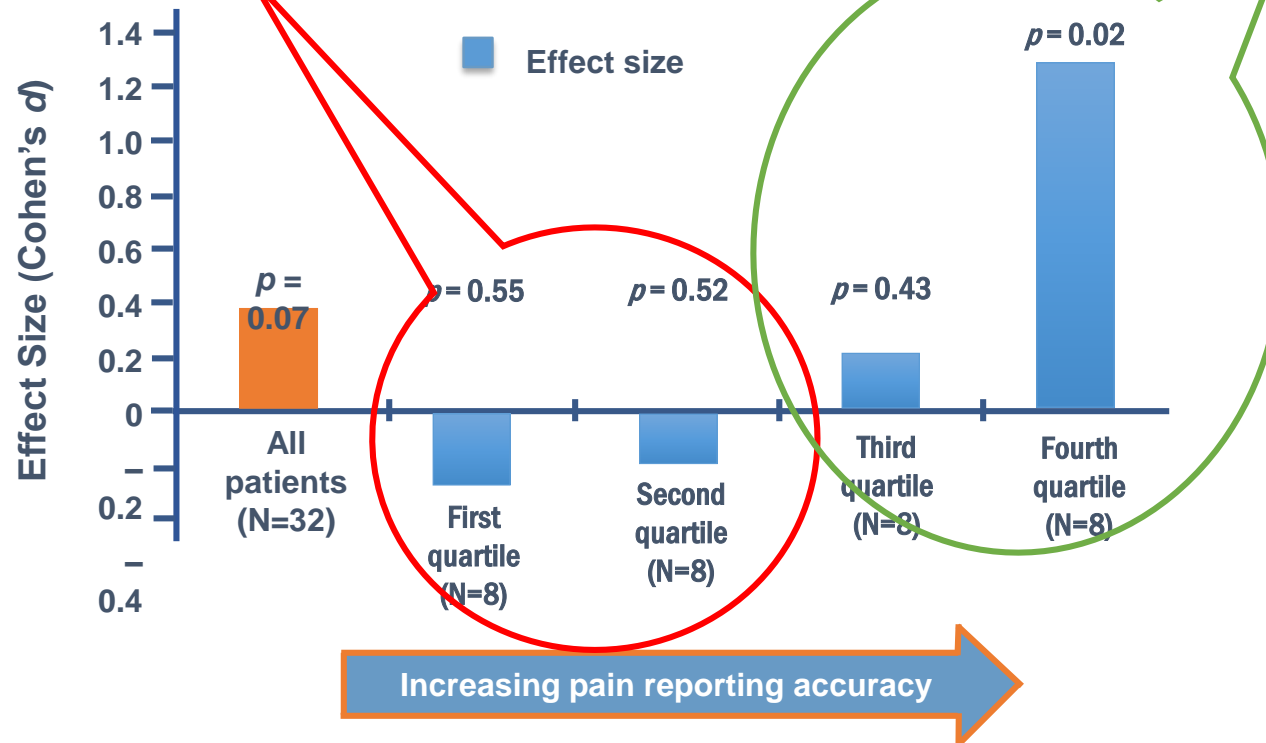- Performing diagnostic assessments
- Consistent study conduct across sites

, 2005; Harris, 2005; Wise, 2009; Dworkin, 2012, 2014; Treister, 2019

Consistency of Pain Reporting Varies from Person to Person: An Experimental Paradigm
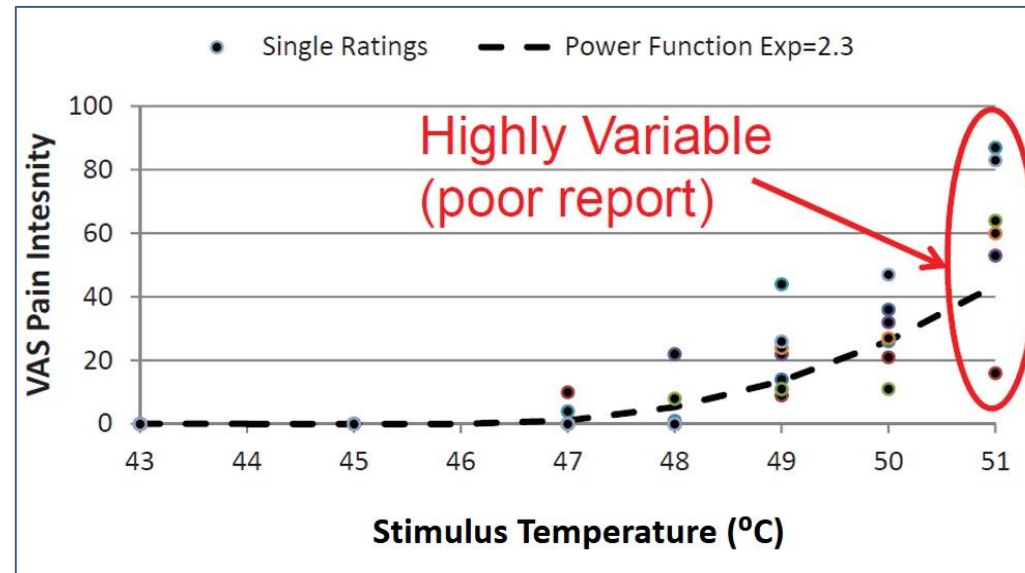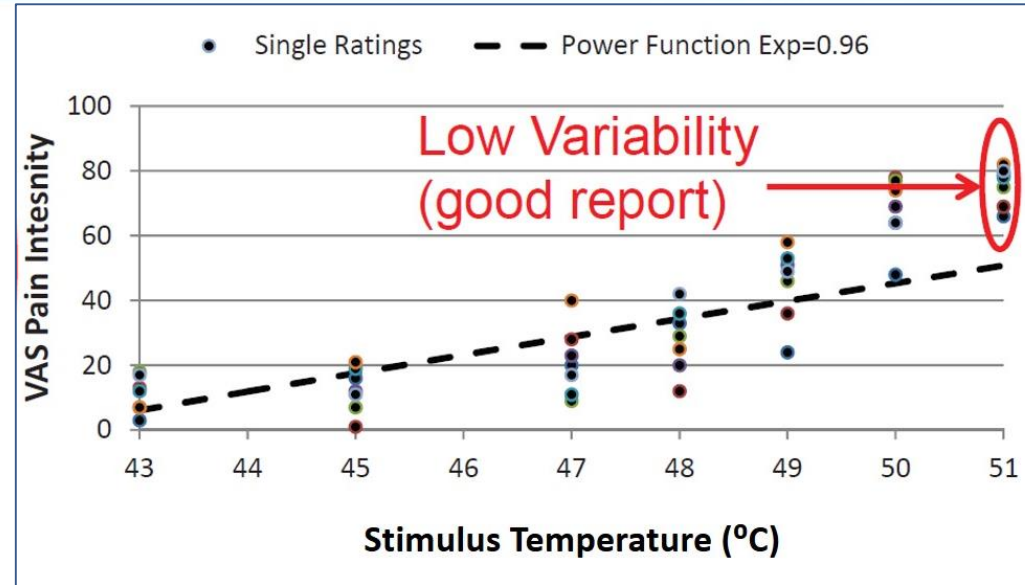
Poor Pain Reporters Cannot Differentiate Naproxen from Placebo

Poor pain reporters:
Drug no better than placebo

Good pain reporters:
Drug beats placebo

Effect Size (Cohen's *d*)

Effect size

p = 0.07 — All patients (N=32)

p = 0.55 — First quartile (N=8)

p = 0.52 — Second quartile (N=8)

p = 0.43 — Third quartile (N=8)

p = 0.02 — Fourth quartile (N=8)

Increasing pain reporting accuracy

Mayorga AJ, Scand J Pain, 2017; Trudeau J, World Congress on Pain, 2012, # PF

Training improves pain reporting accuracy (and decreases placebo response)
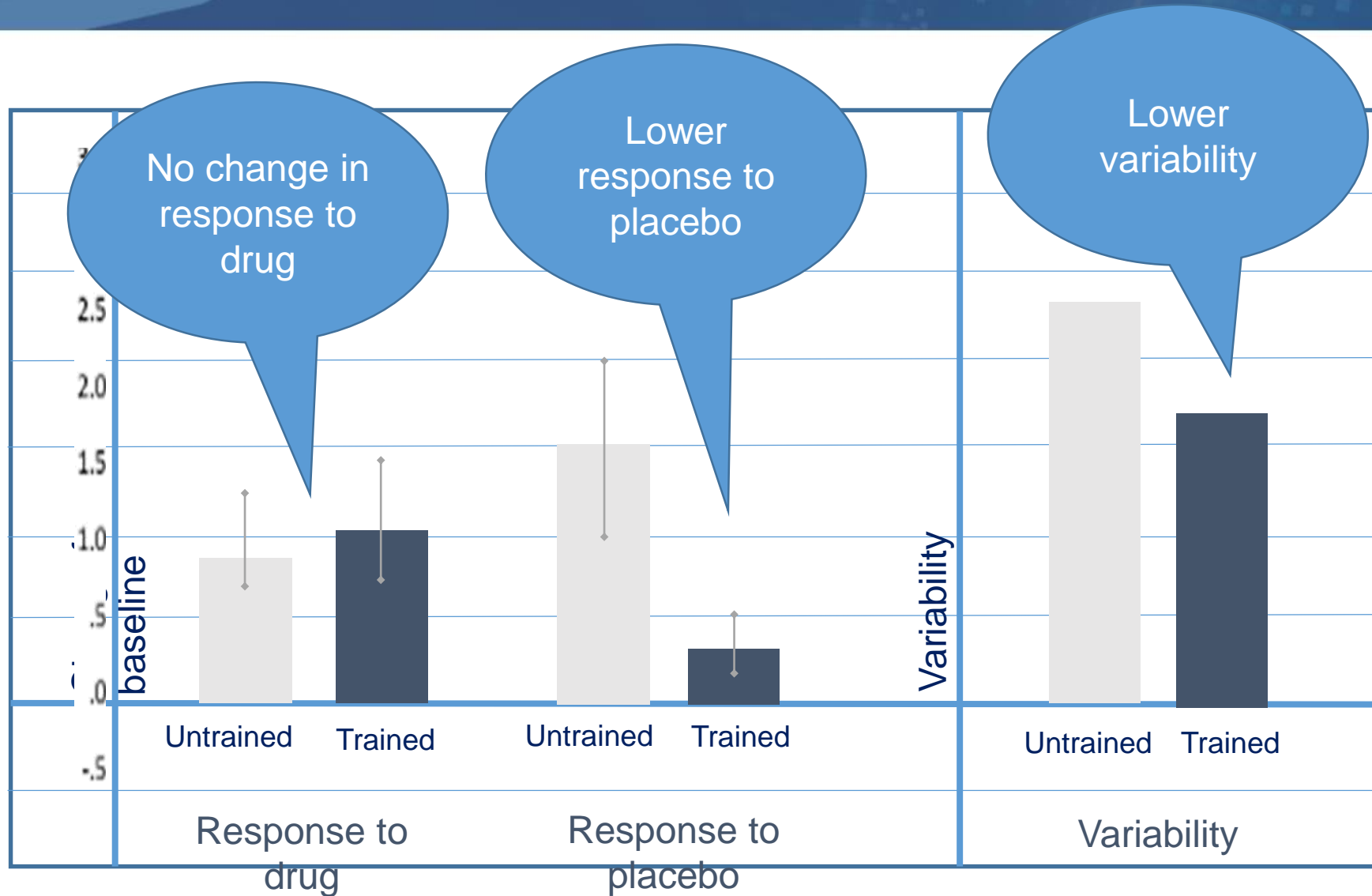
Accurate Pain Reporting Training: Basic Principles

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

ANALGESIC
S O L U T I O N S

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.

Training Focuses Patients on How to Use the Pain Scales

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."

NO PAIN

WORST POSSIBLE PAIN

0 1 2 3 4 5 6 7 8 9 10

11-point, 0-10 NRS scale

Less pain

More pain
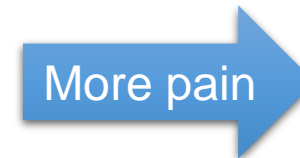
Training Provides Opportunities to Practice Pain Reporting

NO PAIN — WORST POSSIBLE PAIN

0 1 2 3 4 5 6 7 8 9 10

Stubbed toe

NO PAIN — WORST POSSIBLE PAIN

0 1 2 3 4 5 6 7 8 9 10

Gall bladder surgery

NO PAIN — WORST POSSIBLE PAIN

0 1 2 3 4 5 6 7 8 9 10

Kidney stone
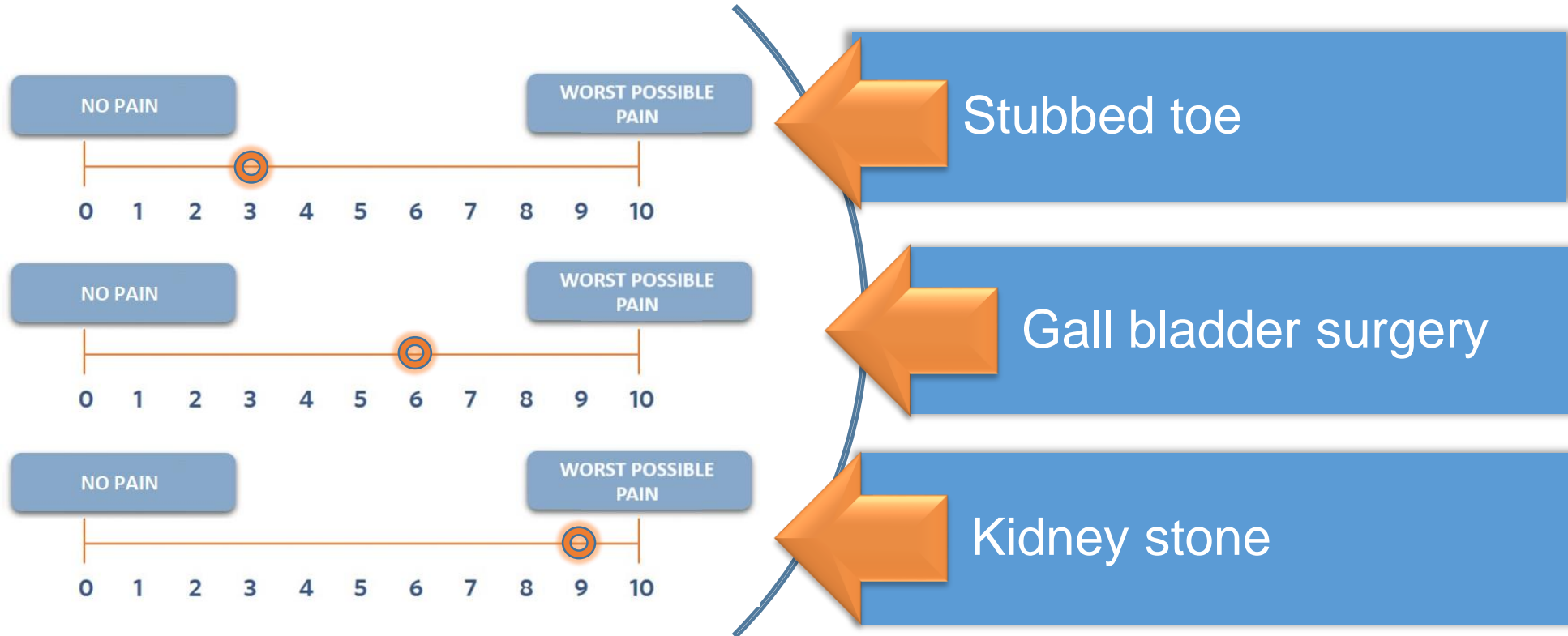
Patients should refer to their 3, 6, and 9 scores whenever they rate their pain in order to achieve consistency and accuracy.

The Placebo Response: Why is it a Problem?

The Powerful Placebo
Henry K. Beecher,
1955

Increasing
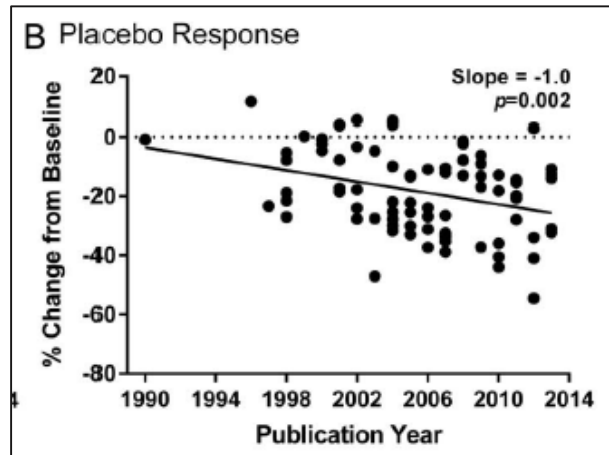
Predicts failure
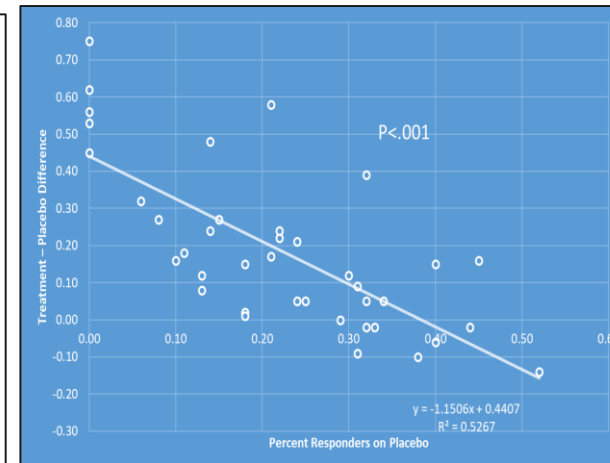
Variable
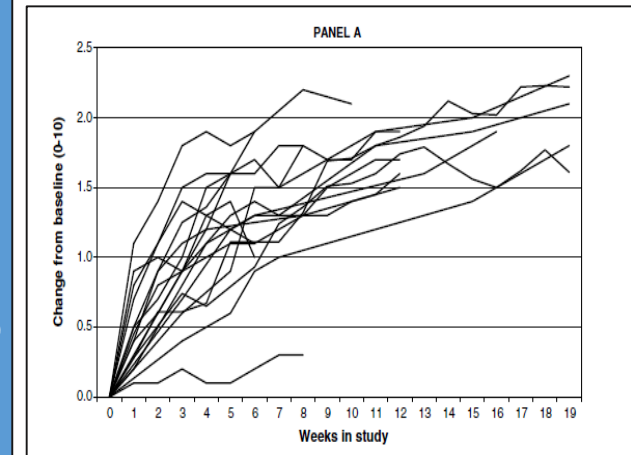
Beecher, 1955; Evans K, 2020; Tuttle AH, 2015; Quessy S, 2008

Neutralize Expectations to Reduce the Placebo Response

▶ The main driver of the placebo response is expectation of benefit by the patients *and* study staff

Expectation of benefit → Pain relief

▶ "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

Investigator Expectation Is Transmitted to Patients

Placebo patients: Investigators thought patients COULD NOT get active drug

Placebo patients: Investigators thought patients COULD get active drug
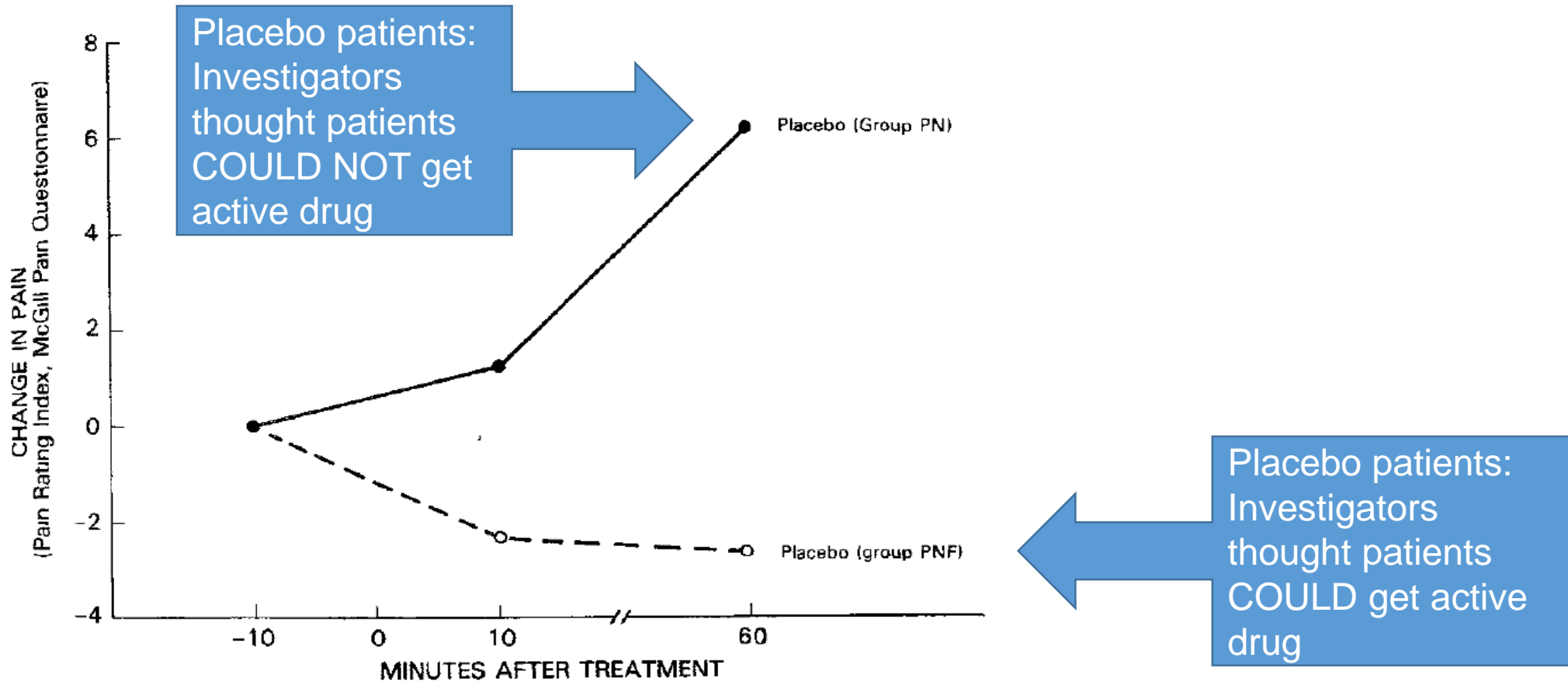
Change in pain rating index between baseline (10 min before injection) and 10 and 60 min after administration of placebo.

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

Change in pain score (%)

Labeling: P, U, M, P, U, M

Treatment: NT, Placebo pill, Maxalt pill
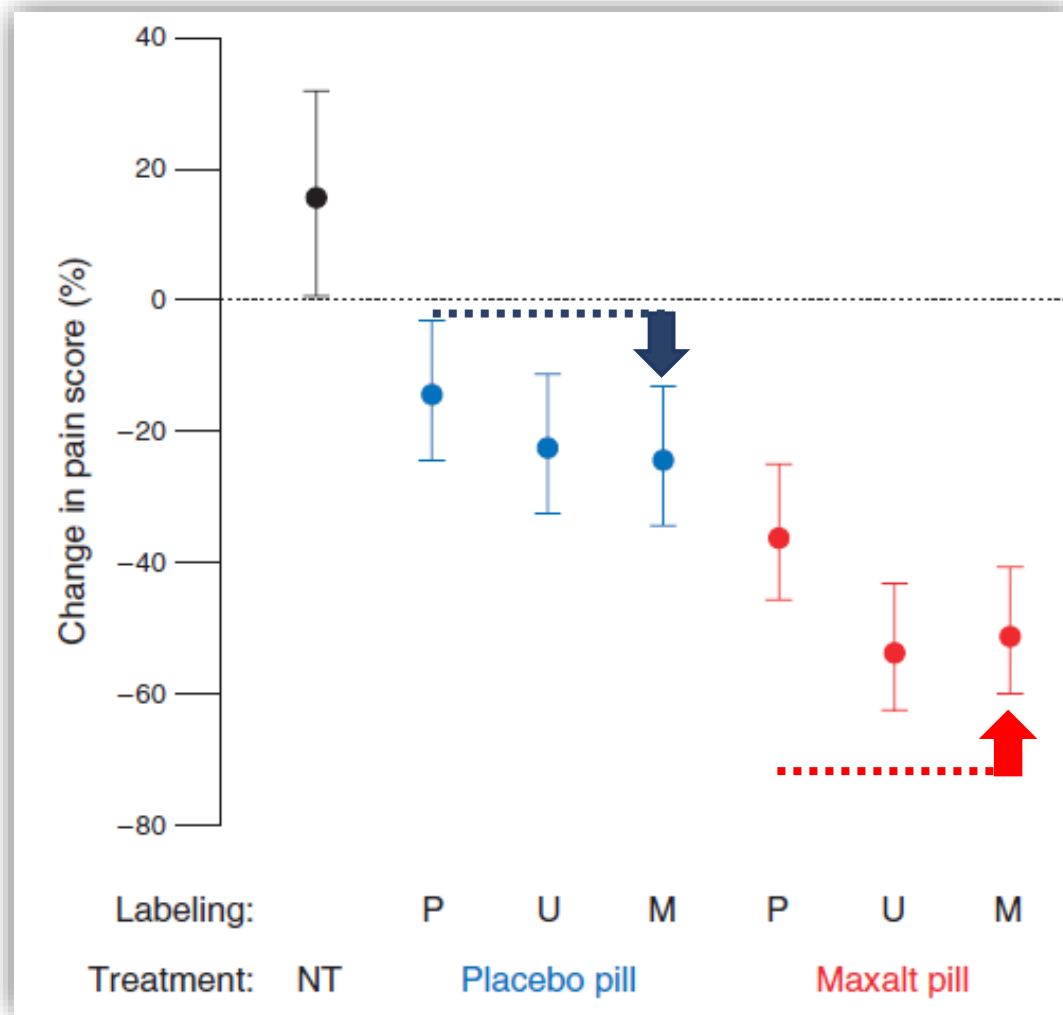
NT = No Treatment; **P** = Placebo; **U** = Placebo or Maxalt; **M** = Maxalt

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

Warmth And Empathy Enhance Placebo Response

Kelley, Psychosomatic Medicine, 2009
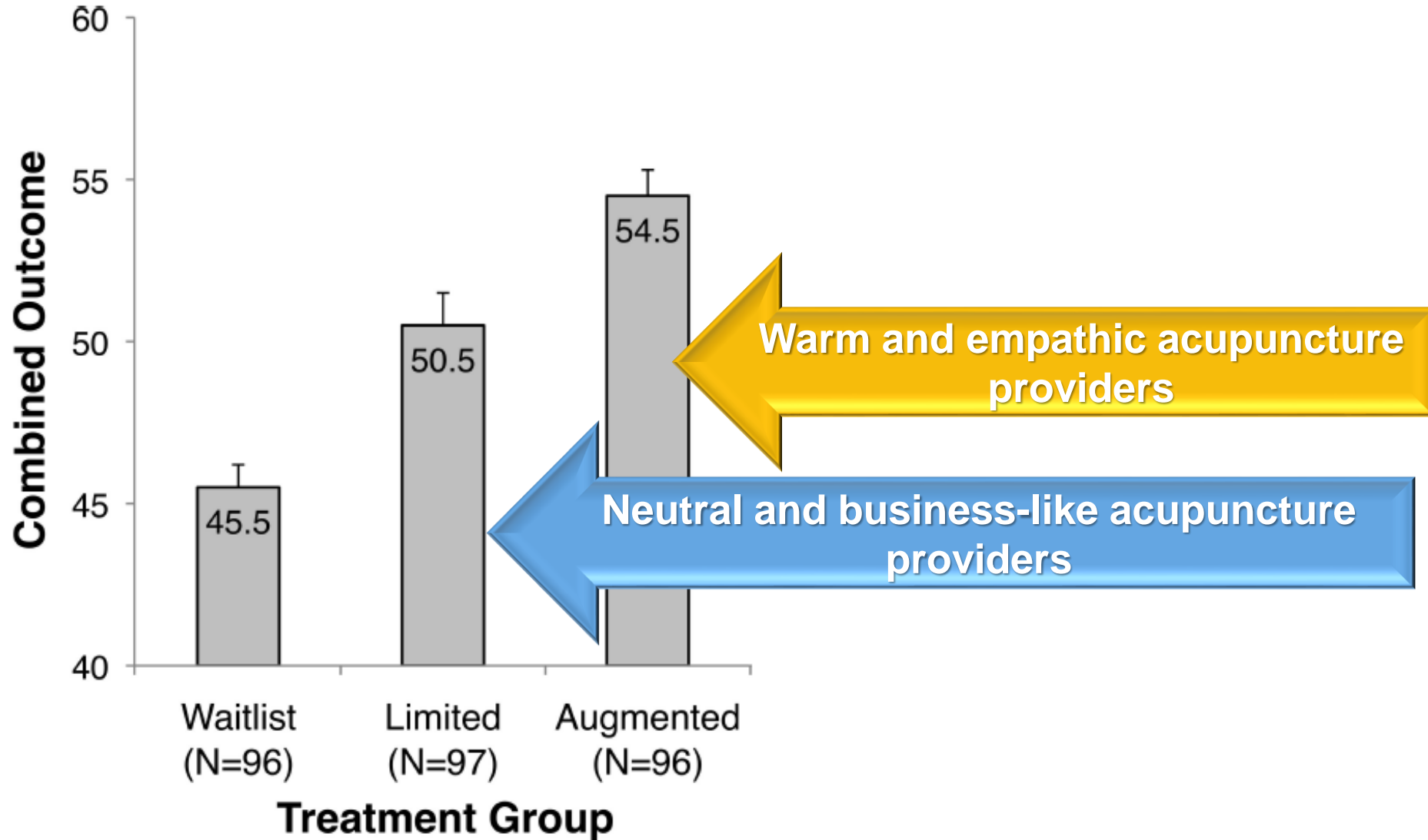
Interact with Patients in a Neutral Manner

▶ 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

Evans K, 2019
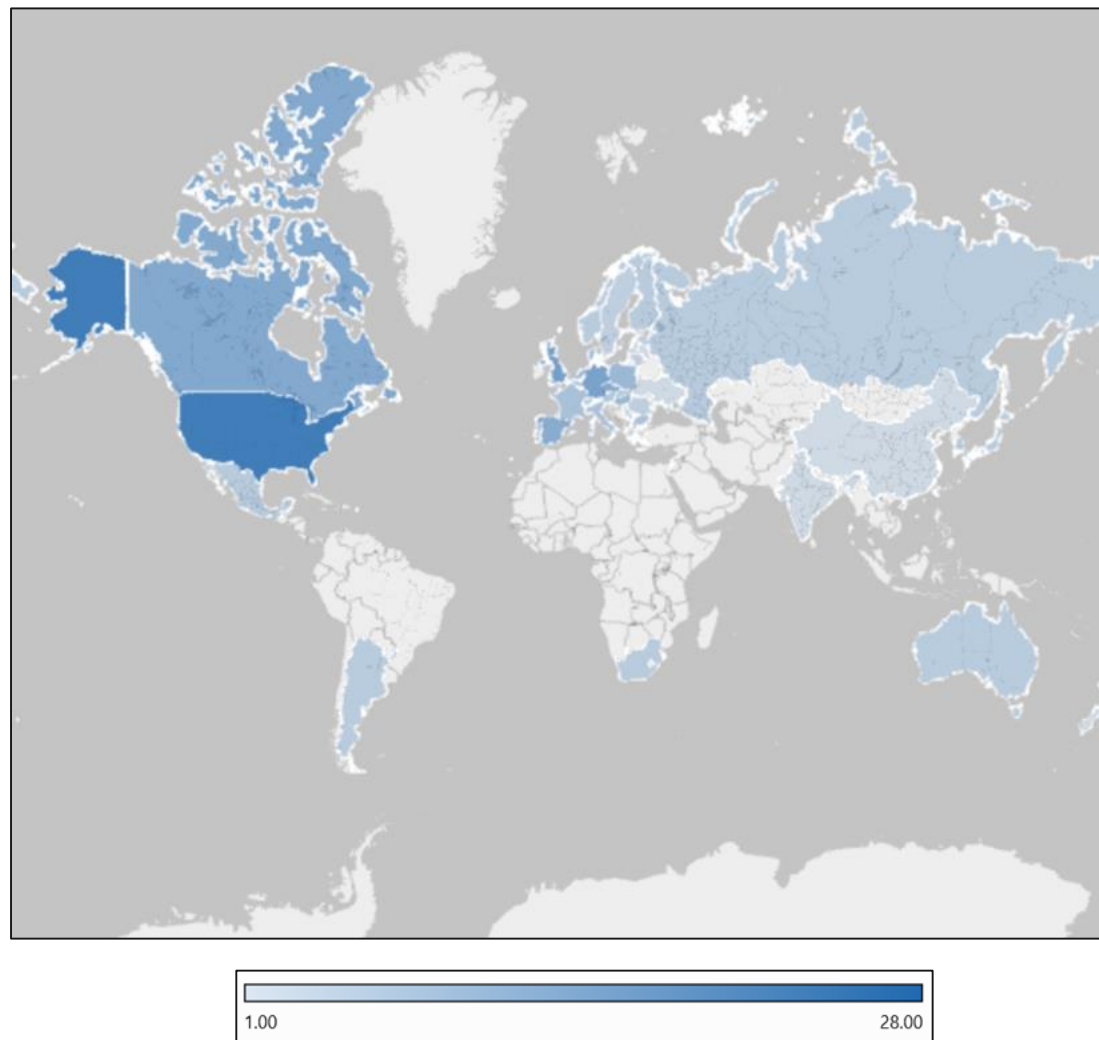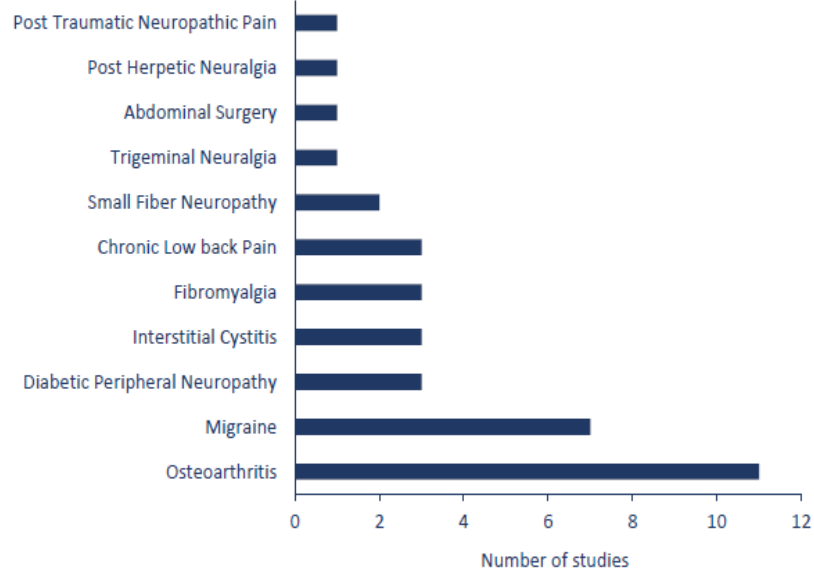
Combining Accurate Pain Reporting and Placebo Response Reduction Training: a clinical trial in lumbosacral radiculopathy demonstrated a low placebo response vs. published studies
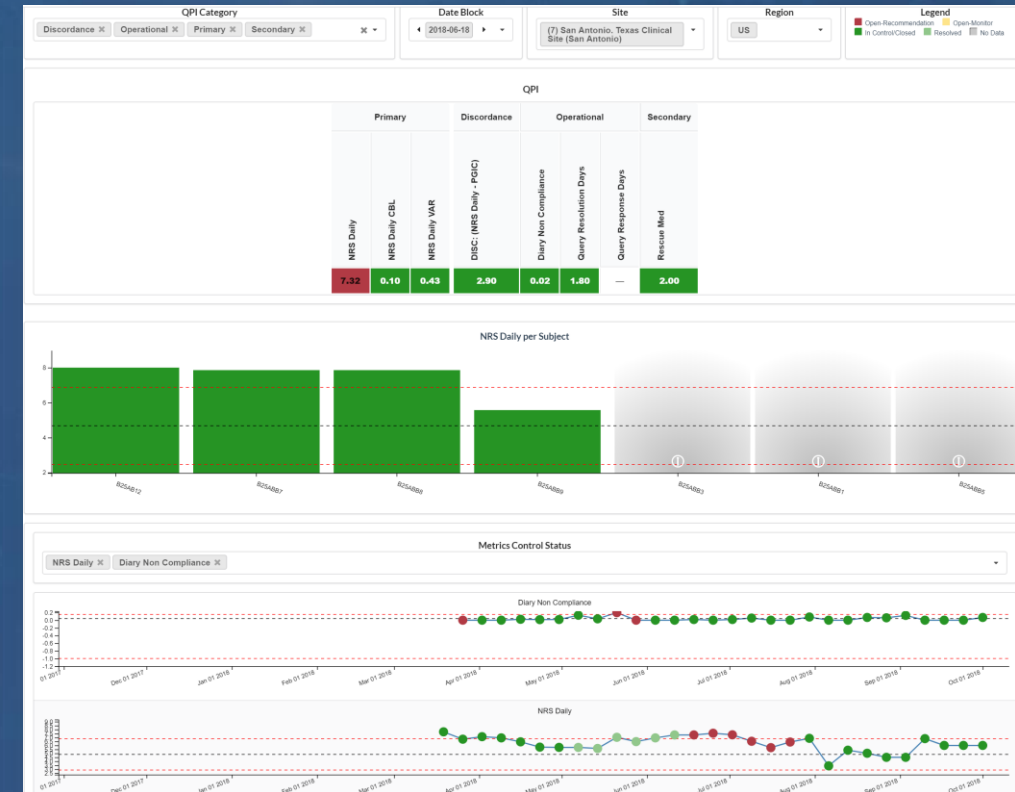

% of Patients in PL Group with ≥30% Reduction in PI
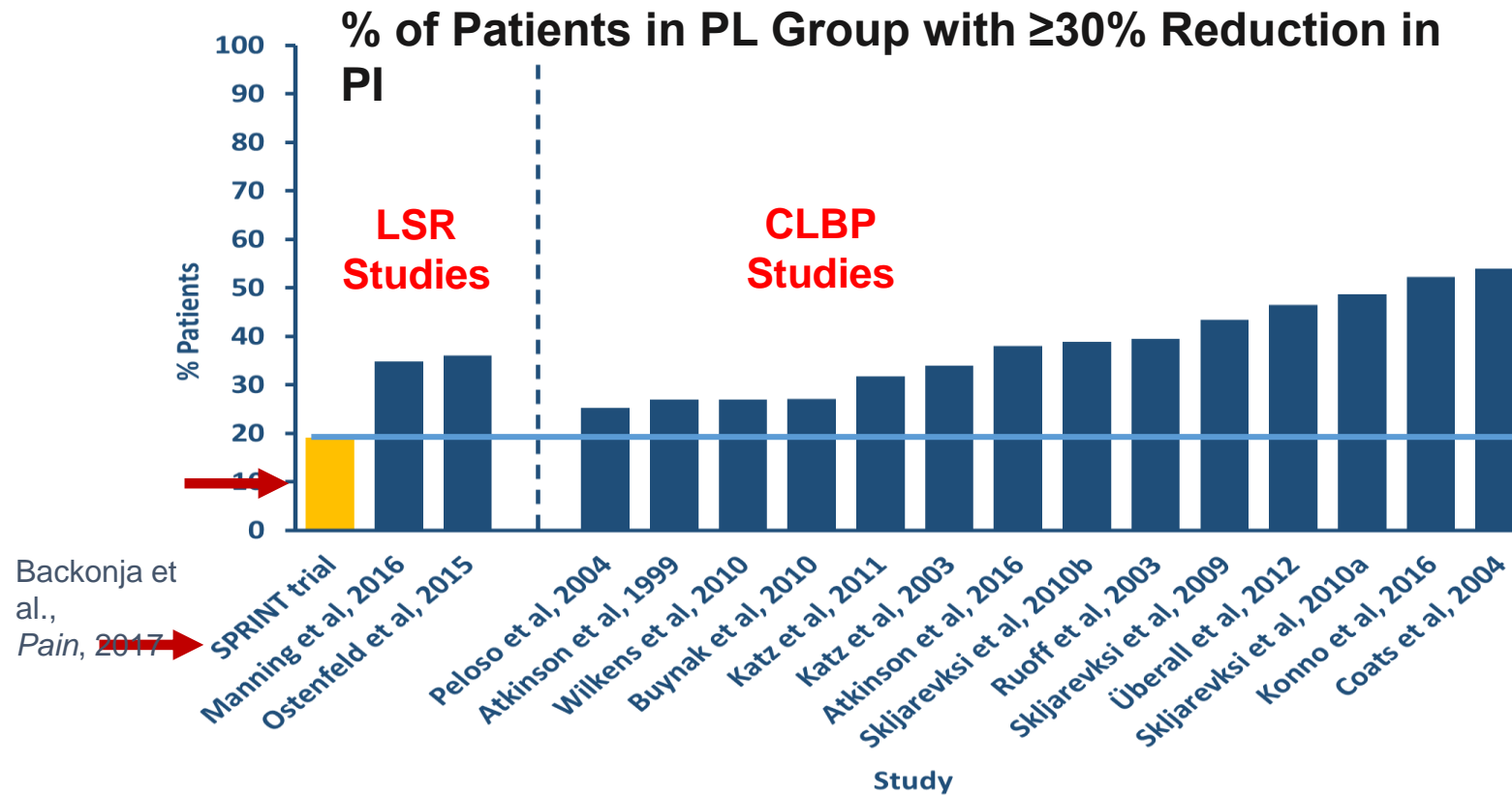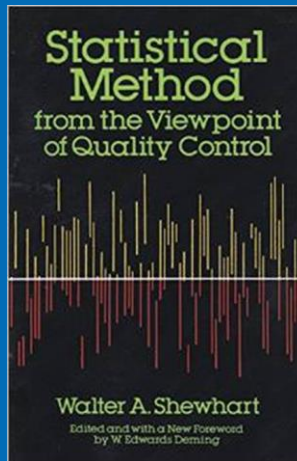
Central statistical monitoring of live clinical trials

Study View
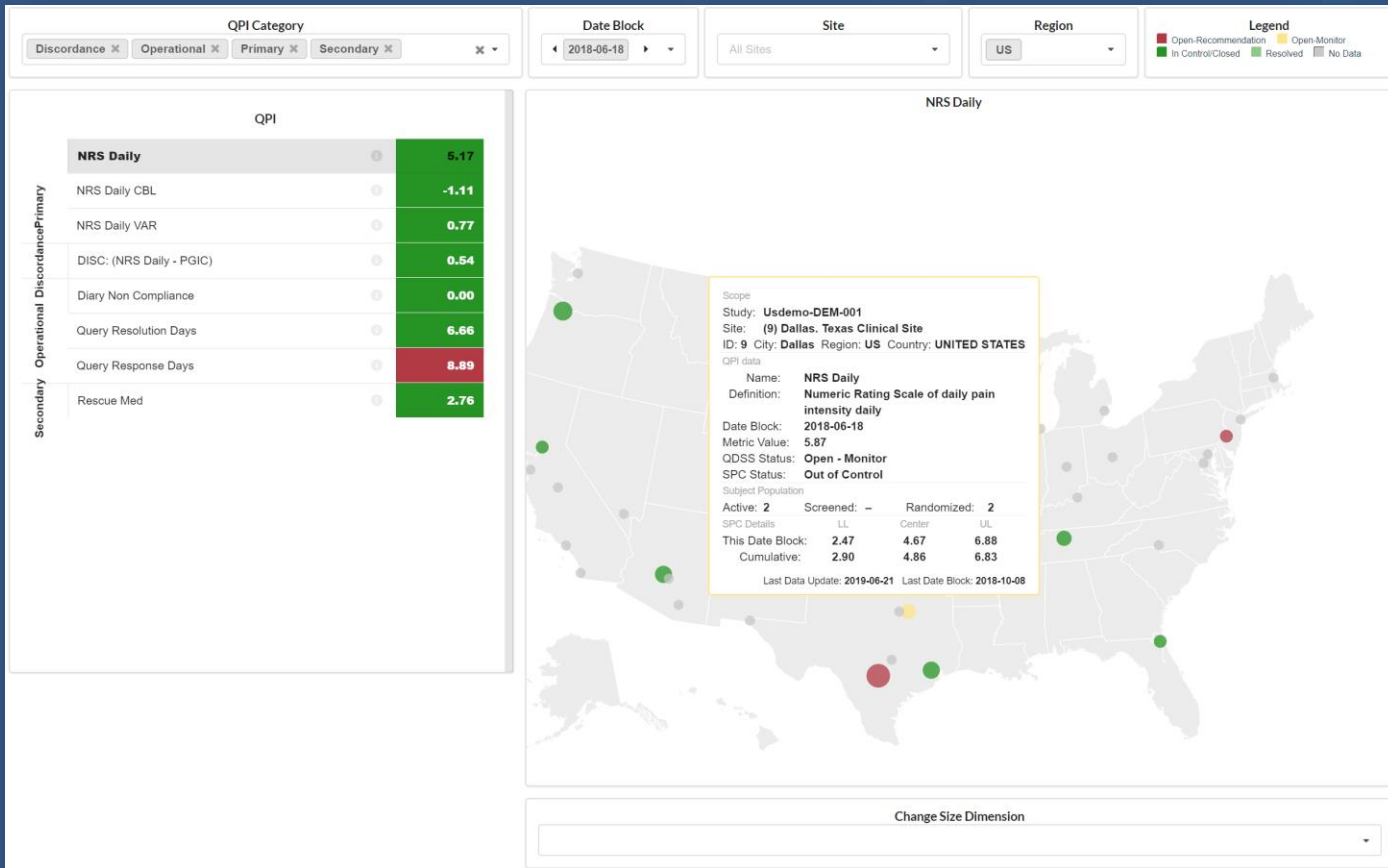
Site View

When you don't have validated fixed thresholds, use SPC
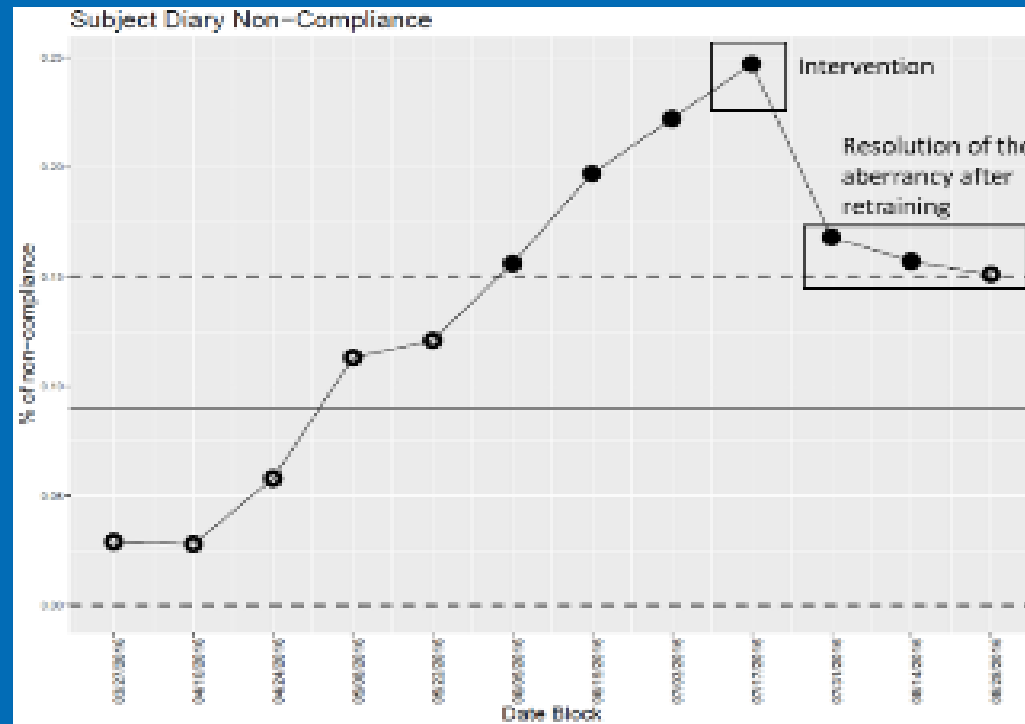

Walter A. Shewhart
1891-1967

Statistical Method from the Viewpoint of Quality Control — Walter A. Shewhart, Edited and with a New Foreword by W Edwards Deming

W. Edwards Deming
1900-1993

Metric Appeared In Control due to lack of data to define accurate control limits at the time, As trial progressed, detection accuracy improved
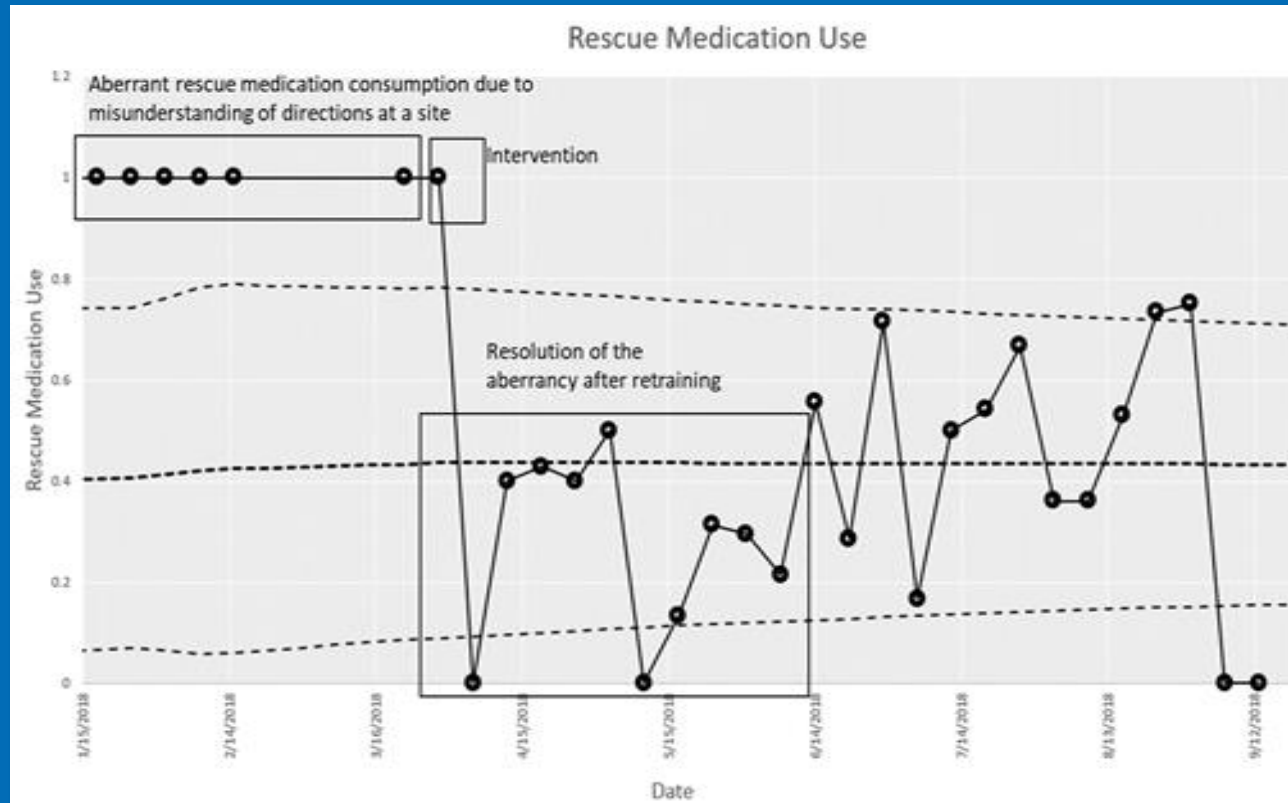
UCL

LCL

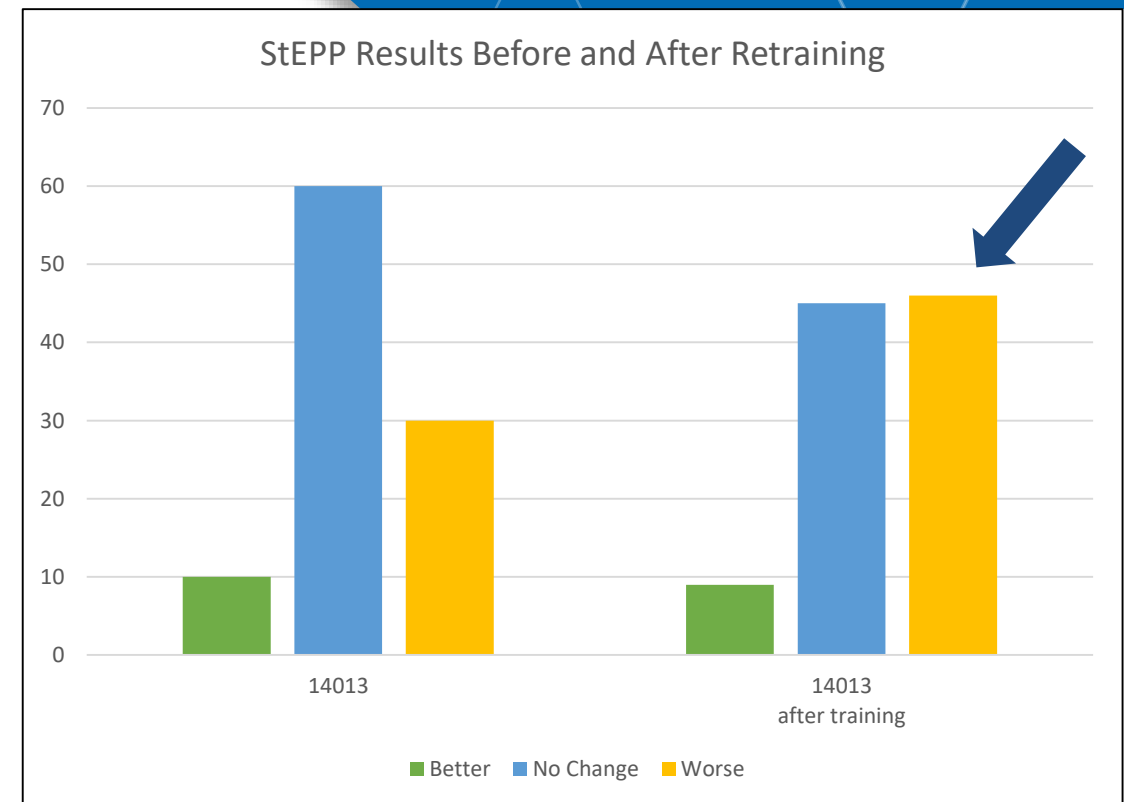Metric value

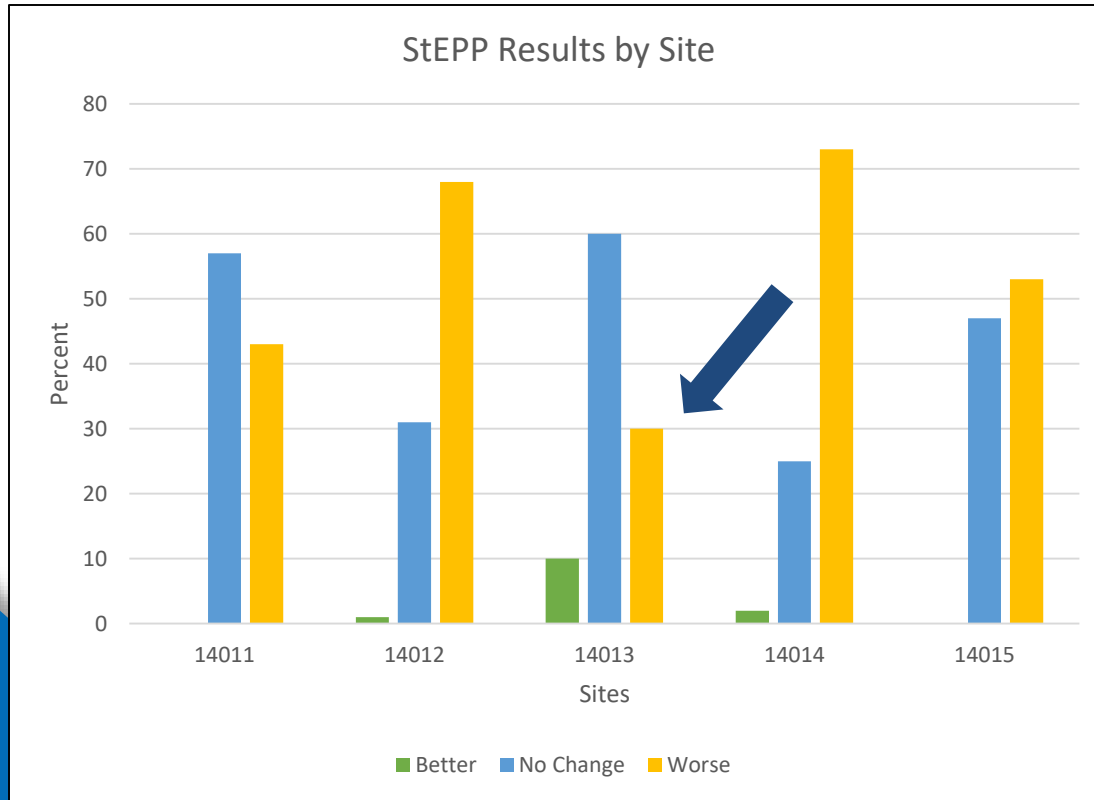Date Block
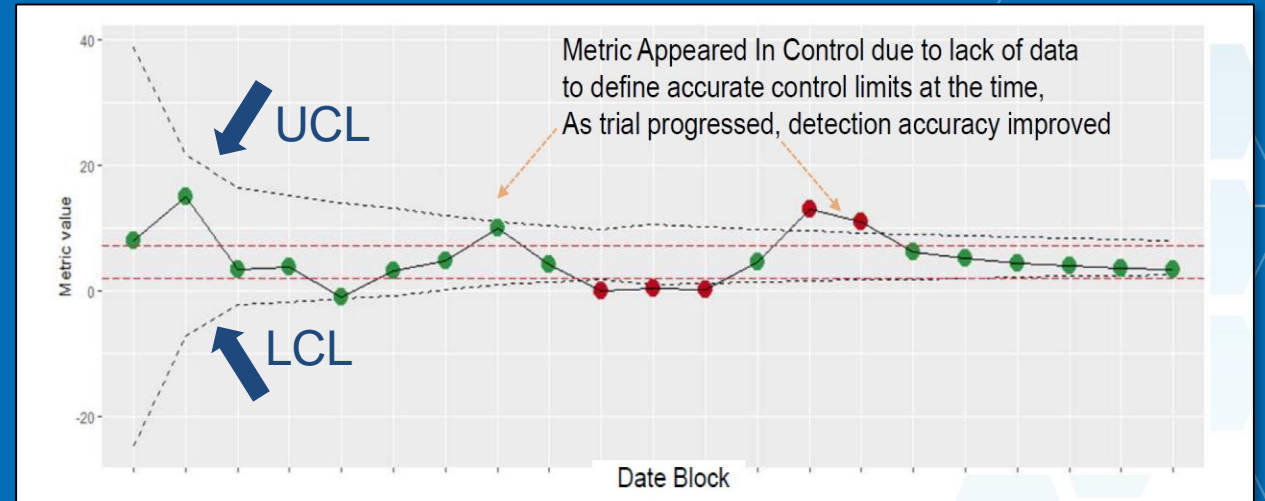
Case 1: Fixed threshold
E-Diary Compliance <85% per week

Case 2: Statistical threshold
Rescue Medication Consumption

Case 3: Statistical Threshold
Staircase-evoked pain: key efficacy endpoint

A subject with multiple critical process failures

Recommendations and considerations

- 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

wcg