

---

## Sources of Measurement Error in Clinical Trials

**Nathaniel Katz, MD, MS**

Analgesic Solutions, Inc.

Tufts University School of Medicine

FDA Clinical Outcome Assessments Workshop

October 19, 2011

---

## Objectives

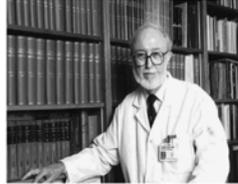
---

- To present a conceptual framework for approaching the problem of measurement error in clinical trials
  - To present early data on potential solutions to these problems
  - Focus on PROs with considerations for Clin-ROs and Obs-ROs
-

Abraham Sunshine  
January 3, 1928–January 2, 2007



Ray Houde  
(1916-2006)



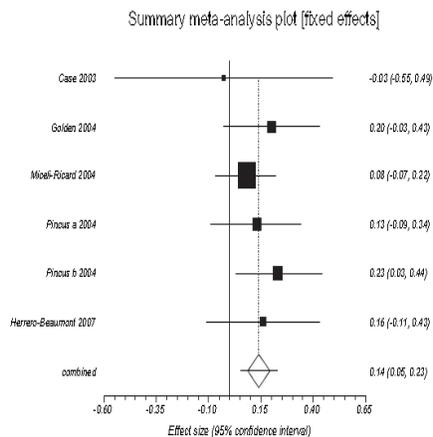
Louis Lasagna  
(1923-2003)



Mitchell B. Max  
(1949-2008)



## Why do effect sizes of identical treatments differ across studies?



- Actual biological effect of drug differs when studied by different authors
- Random chance: God rolls dice in our studies
- The way we measure influences observed effect size

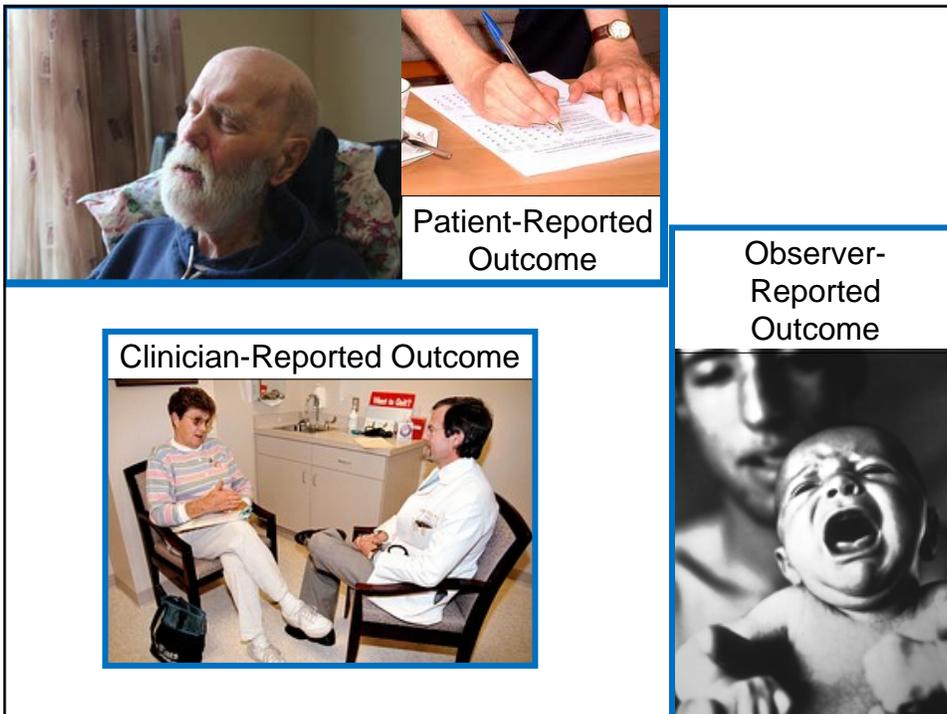
Zhang et al, Osteoarthritis & Cartilage, 2010

## Assay Sensitivity

“a property of a clinical trial defined as the ability to distinguish an effective treatment from a less effective or ineffective treatment”

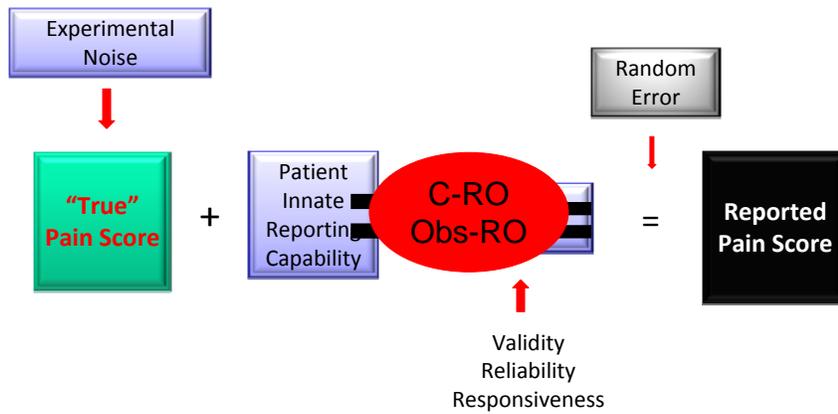
$$\frac{\text{Pain}_{\text{ACTIVE}} - \text{Pain}_{\text{PBO}}}{\text{Std Dev}_P}$$

International Conference on Harmonization. E10: Choice of control groups and related issues in clinical trials.



# Sources of Error in the Measurement System

What generates a pain score?



## THE EXPERIMENT



## Approaches to Optimizing Experimental Design

---

### *Fundamental question:*

- What is the relationship between specific aspects of study design/conduct and observed effect size of treatment?

### *Two approaches:*

- Meta-analysis of the relationship between study design features and outcome
  - Experimental studies on experimental methods
- 



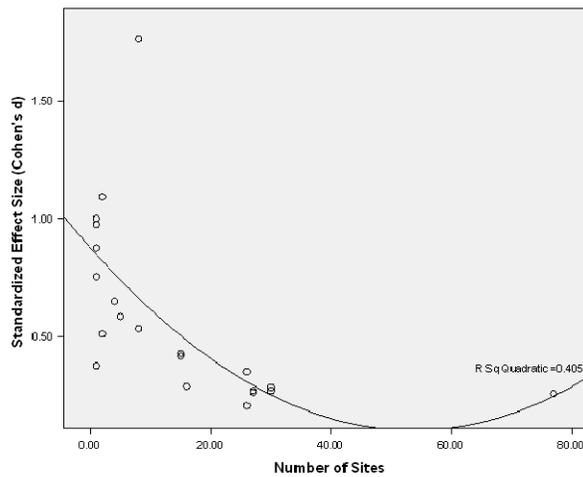
## Reasons for Failure: Opioid Trials

---

- Trial structure
    - Crossover and withdrawal better than parallel treatment
  - Dosing
    - Titration better than non-titration
    - Flexible better than fixed
  - Concomitant analgesics
    - Prohibited better than allowed
  - Rescue
    - Prohibited better than allowed
  - Primary endpoint
    - AUC better than landmark
  - Number of sites
    - The fewer the better
- 

Katz N, et al, Neurology, 2005

## Standardized effect size vs. number of sites, opioid trials



Adapted from Katz N, et al, Neurology, 2005

## Sample size requirements to detect differences in SES by methodologic feature

Difference in SES	Total N of studies	Total N of Patients
0.1	74	12561
0.2	20	3142
0.3	12	1398
0.4	8	787
0.5	6	505

Increasing SES from 0.3 to 0.4 can decrease sample size requirements **per arm** from 175 to 100

### How do clinical trials help?

When some people take a new drug at the same time that other people take a placebo (fake drug), we can find out if:

- the new drug works at all.
- the new drug works better than placebo.
- the new drug works same or worse than placebo.

All "over-the-counter" and prescription medicines, like cough syrup or pain killers, were once tested in clinical trials. They now help thousands of people every day.

### What is a placebo?

A placebo will be used in this clinical trial.

Placebos have *no treatment benefit* because they do not contain any medicines. The placebo will look, smell, taste, and feel like the new drug being tested.

Sometimes participants taking a placebo experience side effects, or their symptoms might get better. Many people assume this means they are taking the real drug.

Although this is possible, you might experience changes in symptoms because you:

- use other treatments or medicines.
- are hopeful about the study.
- feel worried about side effects.
- want to help the research staff.

Please complete the quiz below!

Circle T for true and F for false. Check back in this brochure if you aren't sure. When you finish, check your answers on the bottom of the page.

1. A placebo contains only inactive substances.	T/F
2. Participants in a clinical trial have a good chance of receiving placebo.	T/F
3. I should give the most accurate opinion of how I feel.	T/F
4. I should drop out of the study if I do not have any changes in my symptoms.	T/F
5. None of the study staff know if I am actually getting the real drug or a placebo.	T/F
6. If I feel a side effect, it must mean that I am in the real drug group.	T/F



## CLINICAL TRIALS

What you need to know as a research participant




---

# THE PATIENT

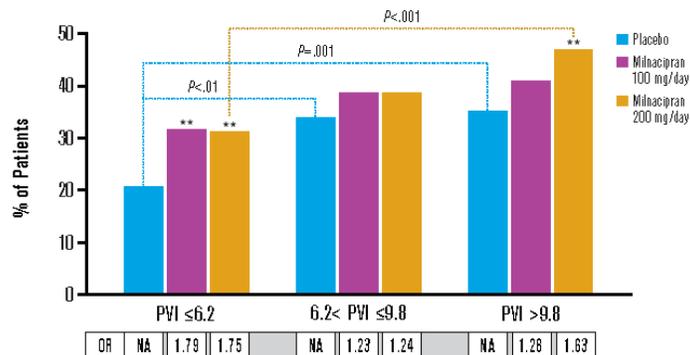
---

## How patients contribute to measurement imprecision

- True within-patient variability of pain
- True between-patient heterogeneity in the capability of responding to treatment
- Patient heterogeneity in appropriateness of the outcome instrument
- Patient pain reporting capability

## Baseline pain variability vs. effect size of milnacipran in fibromyalgia

### B. Pain Responders



Palmer R, et al, presented at 13<sup>th</sup> World Congress on Pain, Montreal, Canada, 2010

## Defining OA Patient Subtypes: the Bedside Sensory Testing Kit - OA



## Sensory Categories in OA: Pilot Study

	No hyperalgesia	1° hyperalgesia	2° hyperalgesia	1° and 2° hyperalgesia
Intact DNIC	N=3	N=1	N=2	N=2
Dysfunctional DNIC	N=0	N=1	N=2	N=9

Inter-rater reliability = .59 - .72

Katz N, presented at American Pain Society Annual Meeting, 2010



## WOMAC Knee and Hip Osteoarthritis Index: Pain Subscale

---

QUESTION: How much pain have you had...	None	Mild	Moderate	Severe	Extreme
1. when walking on a flat surface?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
2. when going up or down stairs?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
3. at night while in bed? (that is - pain that disturbs your sleep)	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
4. while sitting or lying down?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
5. while standing?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4

---

Bellamy N, Clin Exp Rheum, 2005



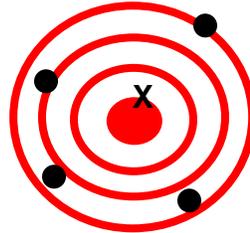
---

## A BRIEF DETOUR THROUGH TERMINOLOGY BEFORE WE GET TO PT. PERFORMANCE

---

## Terminology 1

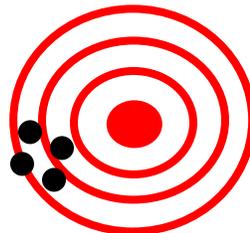
- Validity:
  - Evidence that the instrument measures the concept of interest including evidence from qualitative studies.<sup>1</sup>
- Accuracy
  - The degree of closeness of measurements of a quantity to the quantity's true value<sup>2</sup>



1. FDA PRO guidance. 2. JCGM, 1997

## Terminology 2

- Reliability:
  - Stability of scores over time when no change is expected in the concept of interest<sup>1</sup>
- Precision
  - The degree to which repeated measurements under unchanged conditions show the same results<sup>2</sup>



For an individual measurement to be close to the “true value,” the measurement system must be valid (accurate) AND reliable (precise)

1. FDA PRO guidance. 2. JCGM, 1997

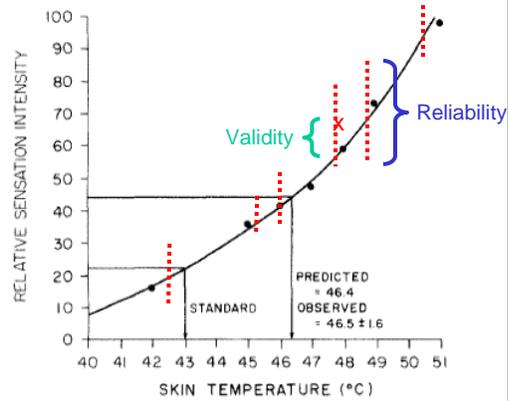
# Measuring Patient Pain Reporting Performance

## Reliability:

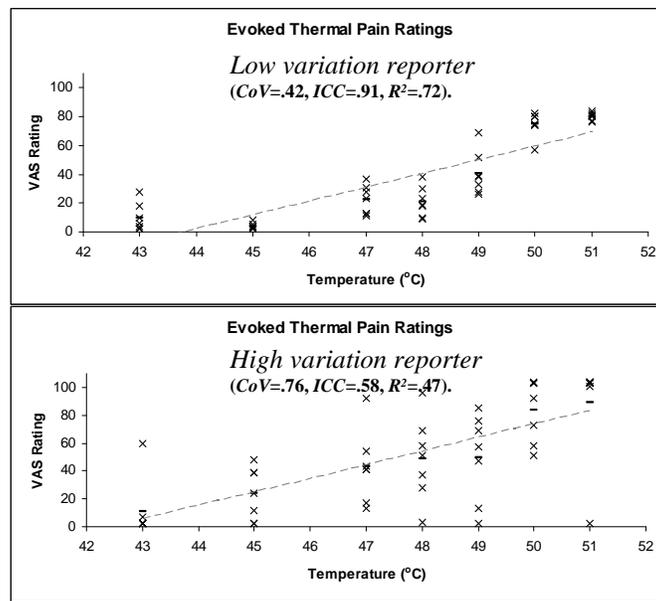
-How tightly clustered are the pain reports for the same stimulus

## Validity (accuracy):

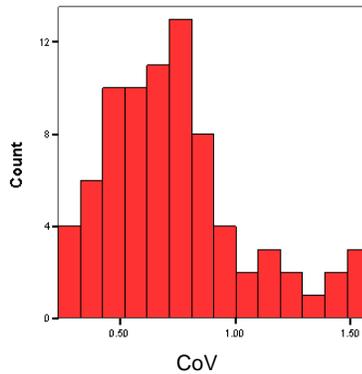
- How close the mean of the pain reports for a given stimulus are to the "true value" (the point on the psychophysical function)



# Psychophysical Profiles



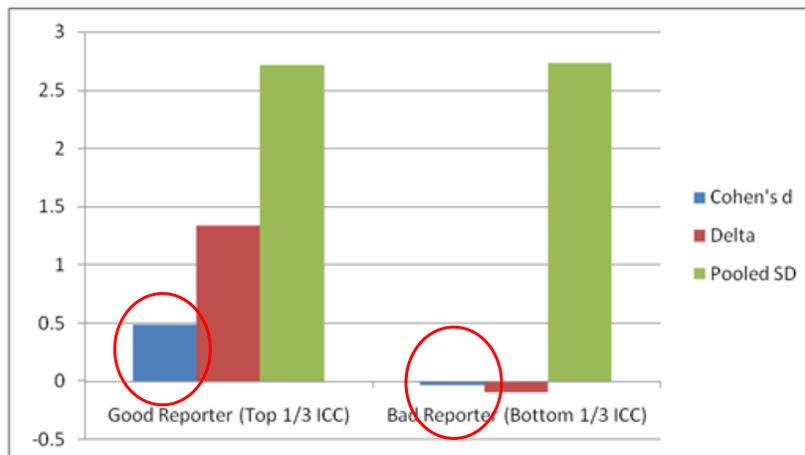
## Frequency Plots for Pain Reporting Skill



**N= 79**  
**Mean = .74**  
**SD= .31**

Subjects demonstrated a large range of performance in pain reporting skill as indexed by CoV, ICC, and R<sup>2</sup>.

## High-accuracy pain reporters discriminate pre-post exercise pain better than low-accuracy reporters



---

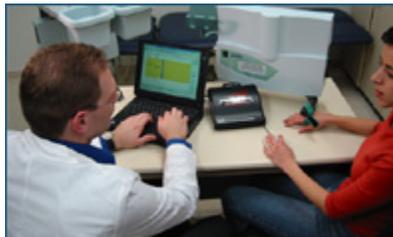
## THE MEASURE

---

## Pain Matching

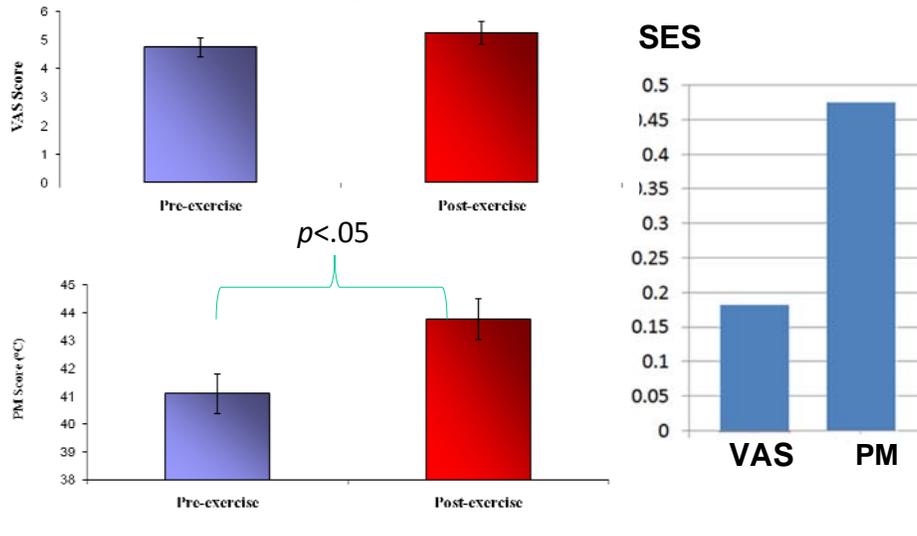
---

Subjects adjust thermode temp until  $\text{pain}_{\text{heat}} = \text{pain}_{\text{OA}}$  (forced choice staircase procedure)

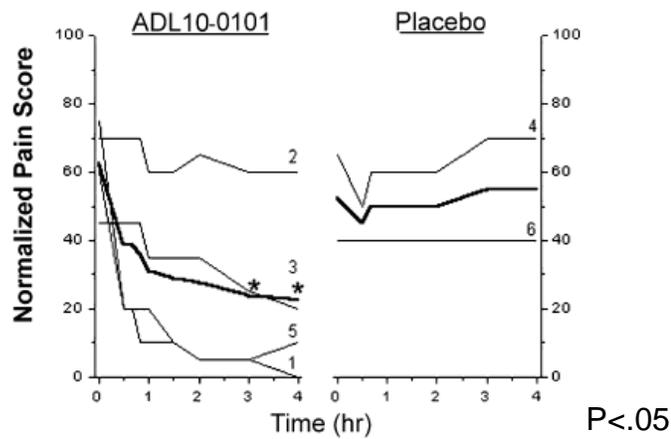


## Delta Exercise Pain Results:

Change in pain significantly different for PM not VAS



## ANALGESIC SOLUTIONS Other explorations of alternative pain measures



Eisenach J, et al, Pain, 2003



## Translating Concepts to Clinician and Observer Assessments

---

- Intermediaries (clinicians, non-clinician observers) introduce an additional element into the measurement system
  - Validity (the degree to which intermediaries' observations reflect concept of interest, and its variability) cannot be assumed
  - Reliability, and stability of reliability, cannot be assumed
- 



## Conclusions

---

- Validity (accuracy) and reliability are characteristics of *measurement systems*, that involve measure, patient, rater, and experimental factors
  - Improving validity/reliability of measurement systems is an important opportunity to accelerate development of therapeutics
  - Both can be measured and subject to rational study
-



---

Thanks to:  
Jeremiah Trudeau, PhD, Bob Dworkin, PhD,  
Dennis Turk, PhD, Mark Jensen, PhD

Nathaniel Katz, MD, MS  
Analgesic Solutions  
781-444-9605, x124  
[nkatz@analgesicsolutions.com](mailto:nkatz@analgesicsolutions.com)

---