# Can assay sensitivity be increased in analgesic trials? Maximizing the reliability and validity of outcome measures

Laurie B Burke
Associate Director for Study Endpoints and Labeling
Office of New Drugs, CDER, FDA
and
Donald L Patrick
Professor, University of Washington

# Summary

- Reduction of variability is the next frontier in better measurement and clinical study efficiency
- Small effect sizes demand attention to variability of the outcome measure
- Attention to good measurement principles (validity and reliability) can minimize variability and increase assay sensitivity

# Assay sensitivity

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

International Conference on Harmonization, E10: *Choice of control groups and related issues in clinical trials.* www.fda.gov/downloads/regulatoryinformation/guidances/ucm125912.pdf

# Assay sensitivity

- Requires adequate statistical power
- Power is a function of
  - Sample size
  - True magnitude of the effect
  - Variability of the outcome assessment
  - Significance level (alpha)

# Options for decreasing sample size while keeping power fixed

- Lower the variability (and SD) of the outcome assessment
- Increase the magnitude of the treatment effect

# 4 Types of Clinical Trial Outcome Assessments

- Clinical Outcome Assessments
  - Patient reported (PRO)
  - Clinician reported (ClinRO)
  - Observer reported (ObsRO)
- Biomarkers

# Outcome Assessment = "Concept" for "Context of Use"

- **Concept** = the "thing" that is measured
  - Score = Concept = Claim
  - Latent (pain intensity) or Observed
  - Direct or indirect measure of treatment benefit
  - Treatment benefit = how patients feel and function

- **Context of Use** = the components of the study objectives and design that influence the claim (eg, population, disease, endpoint)

# What generates the variability of any outcome assessment?

- Patient variability
- Measurement error (random only)
- Measurement mistakes (systematic, non-random)
- Experiment error

# ICH E5 *Ethnic Factors in the Acceptability of Foreign Clinical Data*

- Ethnic factors relate to race or larger populations grouped according to common traits and customs
  - May affect a product's safety, efficacy, dosage, and dose regimen
  - The impact can vary depending upon the product's pharmacologic class, indication, age and gender of the patients, probably many other things
- Ethnic factors are classified as
  - Intrinsic or Extrinsic
- Study results are reviewed for important heterogeneity in response related to these factors

# ICH E5: Classification of intrinsic and extrinsic factors

**Appendix A: Classification of intrinsic and extrinsic ethnic factors**

| INTRINSIC | | EXTRINSIC |
|---|---|---|
| Genetic | Physiological and pathological conditions | Environmental |
| Gender | Age (children-elderly) | Climate |
| Height | | Sunlight |
| Bodyweight | | Pollution |
| | Liver | **Culture** |
| | Kidney | Socioeconomic factors |
| | Cardiovascular functions | Educational status |
| ADME | | Language |
| Receptor sensitivity | | |
| Race | | Medical practice |
| | | Disease definition/Diagnostic |
| Genetic polymorphism of the drug metabolism | | Therapeutic approach |
| | | Drug compliance |
| | Smoking | |
| | Alcohol | |
| | Food habits | |
| Genetic diseases | Diseases | Stress |
| | | Regulatory practice/GCP |
| | | Methodology/Endpoints |

# Intrinsic Heterogeneity Includes:

- Genetics
  - Sex
  - Race
  - Genetic diseases
- Pathophysiological conditions
  - Age
  - Organ function
  - Disease subtype and severity
  - Comorbidities
- Phenotype

# Extrinsic (Environmental) Heterogeneity Includes:

- Culture (SES, occupation, education)
- Language
- Personality (eg, willingness to disclose, attention to detail)
- Medical practice norms
- Disease definition
- Therapeutic approach
- Concurrent meds
- Clinical trials/GCP/regulatory environment
- Data collection format
- Instrument format and content

# Heterogeneity: Intrinsic versus extrinsic factors

- Important variation of both intrinsic and extrinsic factors
- Most focus on intrinsic factors
- Extrinsic factors often overlooked and may not be addressed by randomization
- Primary and key secondary endpoint assessments need to be well-defined and reliable in all subgroups to avoid measurement mistakes

# Impact of Heterogeneity of Treatment Effect Findings

- Non-approvals because of regional heterogeneity, OR
- Need to request more data or another study because of regional heterogeneity, OR
- Need to include information in labeling about regional heterogeneity

# Example of Need for More Data

## **S**atraplatin and **P**rednisone **A**gainst **R**efractory **C**ancer (SPARC)

- Proposed Indication
  - Treatment of patients with hormone refractory prostate cancer (HRPC) who have failed prior chemotherapy
- Study design
  - Multinational (16 countries), randomized, double-blind, placebo-controlled
  - Patients randomized to either: satraplatin + prednisone OR placebo + prednisone every 5 weeks

# Pain Progression

- Definition of worsening based on 2 consecutive 7-day averages of pain intensity OR analgesic use compared to baseline
- Present Pain Intensity (PPI)
  - Report *average* pain intensity over the past 24 hours:

    *0-None*

    *1-Mild*

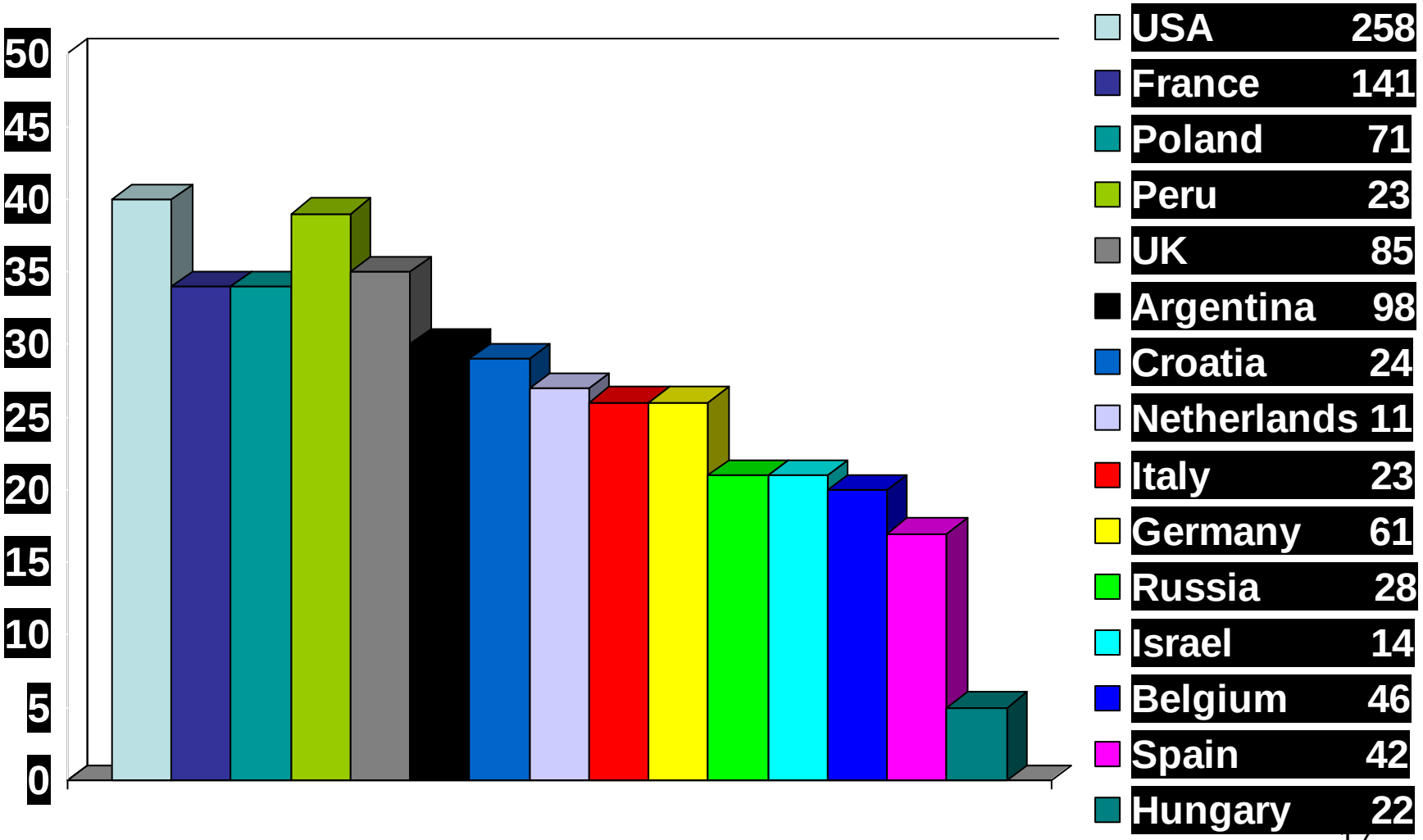
    *2-Discomforting*

    *3-Distressing*

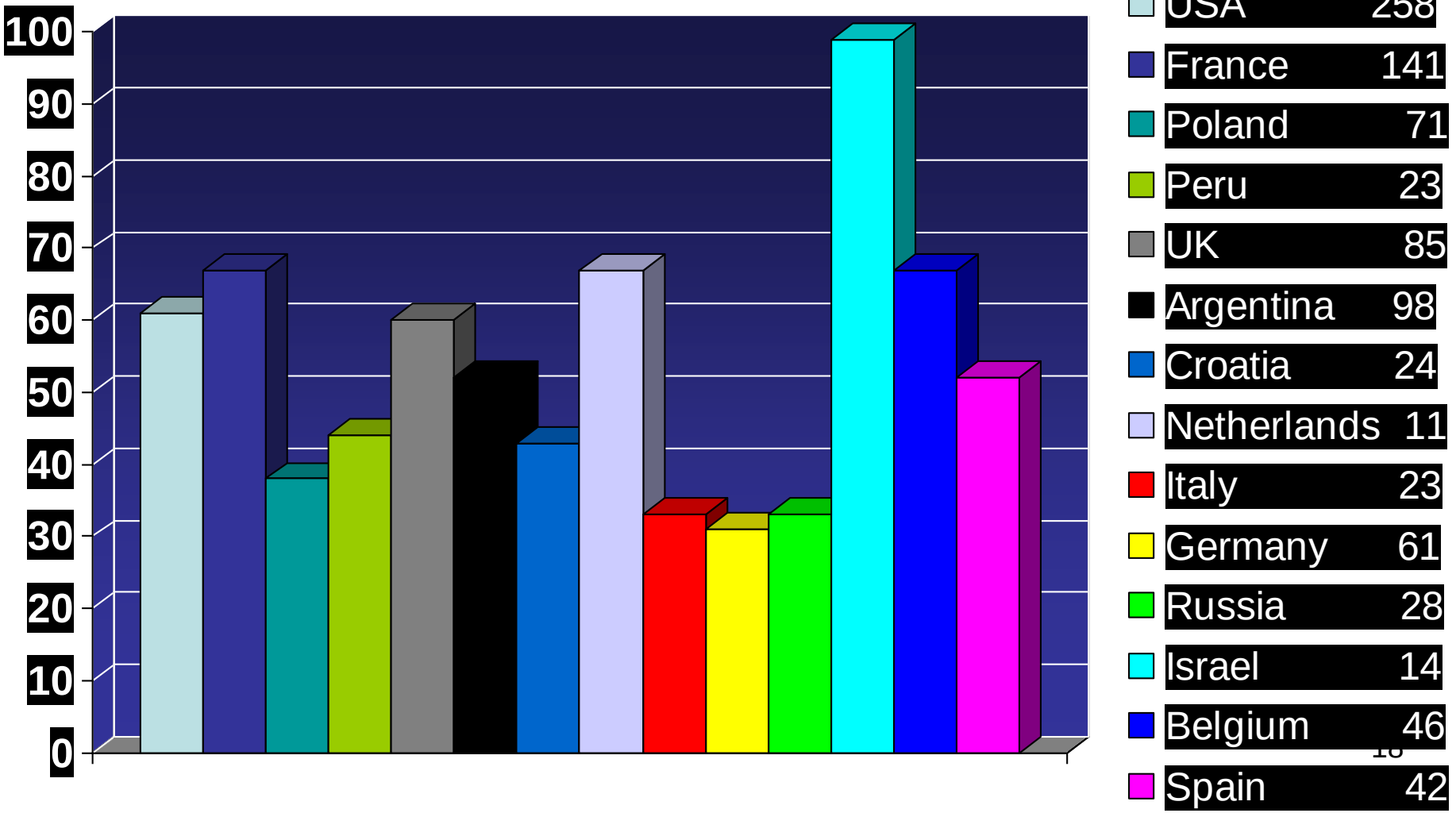
    *4-Horrible*

    *5-Excruciating*

# Percentage of Patients With Pain Progression



# Percent of Pain Progression Attributed to Increased Analgesic Score



# Validation

- There is no such thing as a validated measure
  - Validation results apply to the concept in the context of use tested, not to the instrument
- Traditional approach addresses reliability before validity
- PRO guidance approach recommends that validity is confirmed before psychometric testing (reliability, construct validity)

# Validity

- Evidence to support the conclusion that
  - The score represents the intended concept in the context of use studied
  - The items in the assessment adequately cover the “thing” being evaluated
- Decreased validity leads to increased variability

# How can the variability of a clinical outcome assessment be minimized?

- Qualitative research in the targeted respondents (patient, clinicians, observers) to support content validity
  - Subject variability identified
  - Contributors to measurement error identified
  - Measurement mistakes avoided
  - Contributors to experiment error avoided

# Patient Heterogeneity Affects Variability

- How patients experience the symptoms of interest (e.g., high pain thresholds)
- Capability for careful self-observation
- Willingness to be truthful in reporting
- Reading level and literacy
- Educational status
- Interpretation of the response scale, e.g., willingness to use the extremes
- Expectation of outcomes

# Instrument Attributes Affect Variability

- Clarity or relevance of items
- Literacy level
- Response range
- Response options
- Recall period
- Length of questionnaire
- Formatting, font size
- Length of questionnaire
- Other

# Administration Environment Affects Variability

- Diary versus interview
- Privacy of the setting
- Time to complete questionnaire
- Invasive questions
- Interviewer behavior and interaction
- Need for physical help in responding
- Other

# Increasing Reliability

- Decrease random error
  - Training, eliminate extremes, improve scale design (eg, ePRO)
- Address subject variability
  - Eliminate ceiling, floor and wasted items
  - Add more items in relevant portions of the scale
  - Alter response options to fit the population
- Increase the number of items on the test

# Wasted items may decrease reliability

| | Response | |
|---|---|---|
| MSWS-12 Item | Off Fampridine | On Fampridine |
| 1. Ability to walk | Quite a bit | Moderately |
| 2. Ability to run | Extremely | Extremely |
| 3. Ability to climb stairs | Quite a bit | Moderately |
| 4. Made standing difficult | Moderately | Moderately |
| 5. Limited balance standing or walking | Quite a bit | Moderately |
| 6. Limited walking distance | Quite a bit | Quite a bit |
| 7. Increased effort needed to walk | Quite a bit | Moderately |
| 8. Support walking INDOORS | Quite a bit | Moderately |
| 9. Support walking OUTDOORS | Quite a bit | Quite a bit |
| 10. Slowed your walking | Quite a bit | Moderately |
| 11. Affected how smoothly you walk | Quite a bit | Quite a bit |
| 12. Concentrate on walking | Quite a bit | Quite a bit |
| | 6 items change (order: 1,5,7,8,3,10) | |

Source: Meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, October 14, 2009, www.fda.gov

# More items may reduce standard deviation

| Correlation coefficients between BPI pain scale and MDASI pain worst | | | | | |
|---|---|---|---|---|---|
| | Month 0 | Month 3 | Month 6 | Month 9 | Month 12 |
| r | 0.805 | 0.789 | 0.887 | 0.948 | 0.950 |
| *P* | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |

| Intraclass correlation (ICC) of month 3 and 4 | | | |
|---|---|---|---|
| | **ICC (95% CI)** | **Mean (SD)** | |
| | | **Month 3** | **Month 4** |
| BPI pain scale | 0.824 (0.689 – 0.903) | 1.61 (1.89) | 1.73 (1.85) |
| MDASI pain worst | 0.819 (0.681 – 0.901) | 2.15 (2.38) | 2.30 (2.20) |

Source: Cleeland C, PRO Consortium Workshop, Silver Spring, March 2011.[27]

# Summary

- Reduction of variability is the next frontier in better measurement and clinical study efficiency
- Small effect sizes demand attention to variability of the outcome measure
- Attention to good measurement principles (validity and reliability) can minimize variability and increase assay sensitivity