Measurement in Clinical Research

David Luckenbaugh
Medical Statistician
Experimental Therapeutics and Pathophysiology Branch
National Institute of Mental Health, IRP

OBJECTIVE
Enhance understanding of key principles of measurement relevant for clinical research

Performance
- Time
- Distance
- Speed
- Calories
CONSTRUCT
• A theoretical concept

MEASUREMENT
• A system of defining the level of a construct
  • Operational Definition
    • The method used for examining some domain

Examples
1. Depression
   A. Hamilton Depression Rating Scale
   B. Beck Depression Inventory

2. Tremor
   A. Judge rated spirals
   B. Computer evaluated spirals

3. Heart Disease
   A. Cholesterol
   B. C-Reactive Protein

OUTLINE
1. Validity
2. Reliability
3. Sensitivity to Change
4. Scale
5. Feasibility
VALIDITY
How well does the measure reflect the construct?

VALIDITY

VALIDITY: Types
1. Construct
   A. Face
   B. Content

2. Criterion-related
   A. Convergent
   B. Divergent
Consistency of measurement

Valid Not Valid

Reliable Not Reliable

RELIABILITY and VALIDITY

Valid Not Valid

Reliable Not Reliable
1. Internal Consistency
2. Inter-Rater
3. Test-Retest

Lack of reliability introduces error into your measurement
1. Less sensitive statistics
2. Larger sample size
3. Uninterpretable results

Reliability and Sample Size

![Bar chart showing minimum total sample size for different reliability coefficients]
RELIABILITY: Improving
1. Provide standardized procedures
2. Train raters
3. Monitor raters
4. Use multiple raters for each rating
5. Take repeated observations


SENSITIVITY to CHANGE
Ability to detect improvement or worsening

SENSITIVITY to CHANGE
Can assess with effect size
Cohen’s $d = (\text{Mean}_2 - \text{Mean}_1)/\text{SD}$

• Standard Interpretation
  .8 Large
  .5 Moderate
  .2 Small
SENSEITIVITY to CHANGE


SENSEITIVITY to CHANGE


SCALE
1. Nominal
   No order
2. Ordinal
   Ordered (ranked)
3. Interval
   Ordered + Equal spacing
4. Ratio
   Ordered + Equal spacing + Absolute zero
SCALE
1. Nominal
   Diagnostic status (yes/no)
2. Ordinal
   Stage of illness
3. Interval
   Severity of illness
4. Ratio
   Number of doctor visits

SCALE
1. Continuous
2. Categorical

3. When should you use these?
   A. Continuous – study outcome
   B. Categorical – clinical relevance

Clinical Relevance
1. Sensitivity
   If have illness, how often is test positive?
2. Specificity
   If no illness, how often is test negative?
3. Positive Predictive Value
   If test positive, how often have illness?
4. Negative Predictive Value
   If test negative, how often no have illness?
Sensitivity and Specificity

<table>
<thead>
<tr>
<th>Illness</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td>118</td>
<td>39</td>
</tr>
<tr>
<td>Test Negative</td>
<td>30</td>
<td>47</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.797</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>0.547</td>
<td></td>
</tr>
</tbody>
</table>

Olie, et al., J Aff Dis, 2011

Positive and Negative Predictive Value

<table>
<thead>
<tr>
<th>Illness</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td>118</td>
<td>39</td>
</tr>
<tr>
<td>Test Negative</td>
<td>30</td>
<td>47</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>0.732</td>
<td></td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>0.630</td>
<td></td>
</tr>
</tbody>
</table>

Olie, et al., J Aff Dis, 2011

FEASIBILITY

1. Cost
2. Time
3. Environment
RESOURCES

Kraemer (1991). To increase power in randomized clinical trials without increasing sample size. 
Psychopharmacology Bulletin.
