Assessing the Minimally Important Difference of Health-Related Quality of Life Measures

Ron D. Hays, Ph.D.

NCI

April 4, 2007

11-11:59 am

Responsiveness to Change

- **HRQOL measures should be responsive to interventions that change HRQOL**
- **Need external indicators of change (Anchors)**
  - **mean change in HRQOL scores among people who have changed**

Self-Report Indicator of Change

- Overall has there been any change in your asthma since the beginning of the study?

Much improved; Moderately improved; Minimally improved

No change

Much worse; Moderately worse; Minimally worse

Clinical Indicator of Change

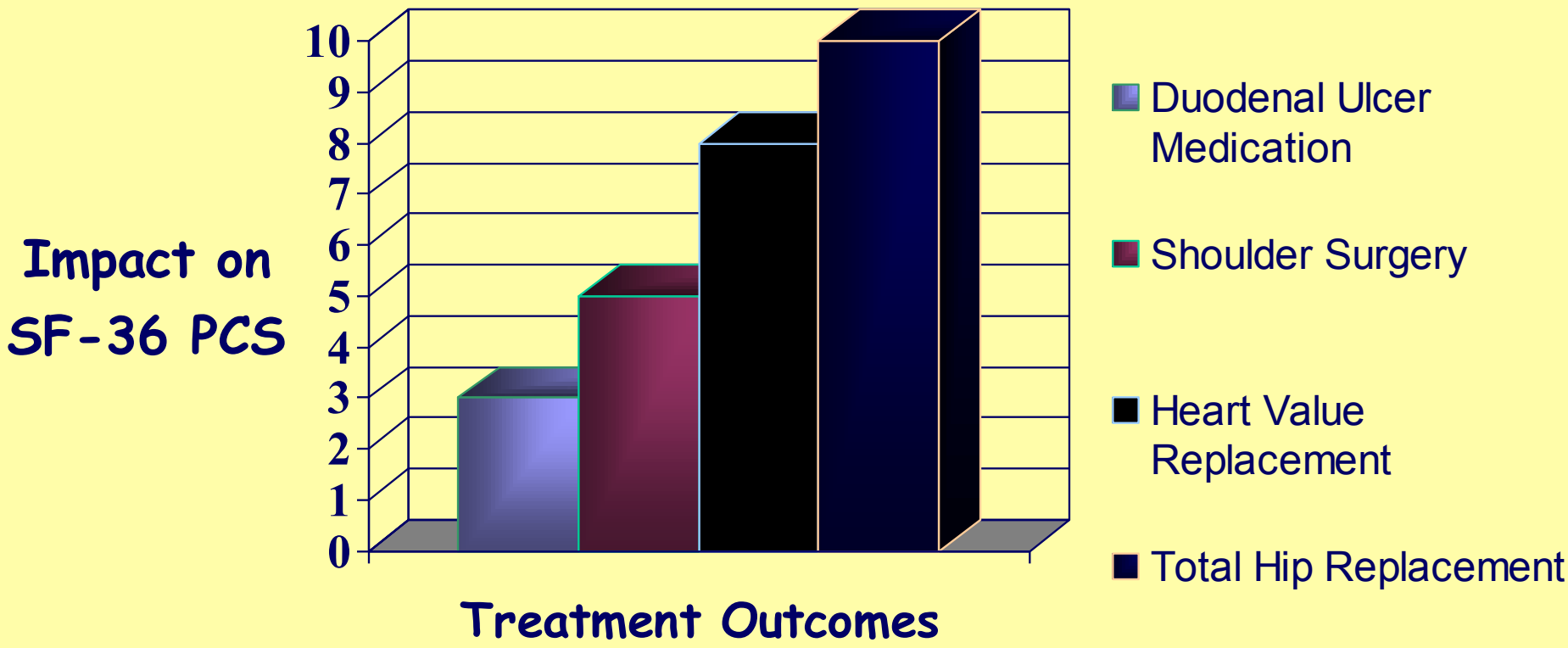
- "changed" group = seizure free (100% reduction in seizure frequency)

- "unchanged" group = <50% change in seizure frequency

Effect Size = D/SD

- Small: 0.20->0.49
- Moderate: 0.50->0.79
- Large: 0.80 or above

Treatment Impact on PCS



Responsiveness Indices

(1) Effect size (ES) = D/SD

(2) Standardized Response Mean (SRM) = D/SD†

(3) Guyatt responsiveness statistic (RS) = D/SD‡

D = raw score change in "changed" group;

SD = baseline SD;

SD† = SD of D;

SD‡ = SD of D among "unchanged"

Minimally Important Difference (MID)

- One can observe a difference between two groups or within one group over time that is statistically significance but small.
- With a large enough sample size, even a tiny difference could be statistically significant.
- The MID is the smallest difference that we care about.

FDA Guidance for Industry (2006)

"Patient-Reported outcome measures: Use in medical product development to support labeling claims"

- Describes how the FDA will evaluate the appropriateness and adequate of PRO measures used as effectiveness endpoints in clinical trials

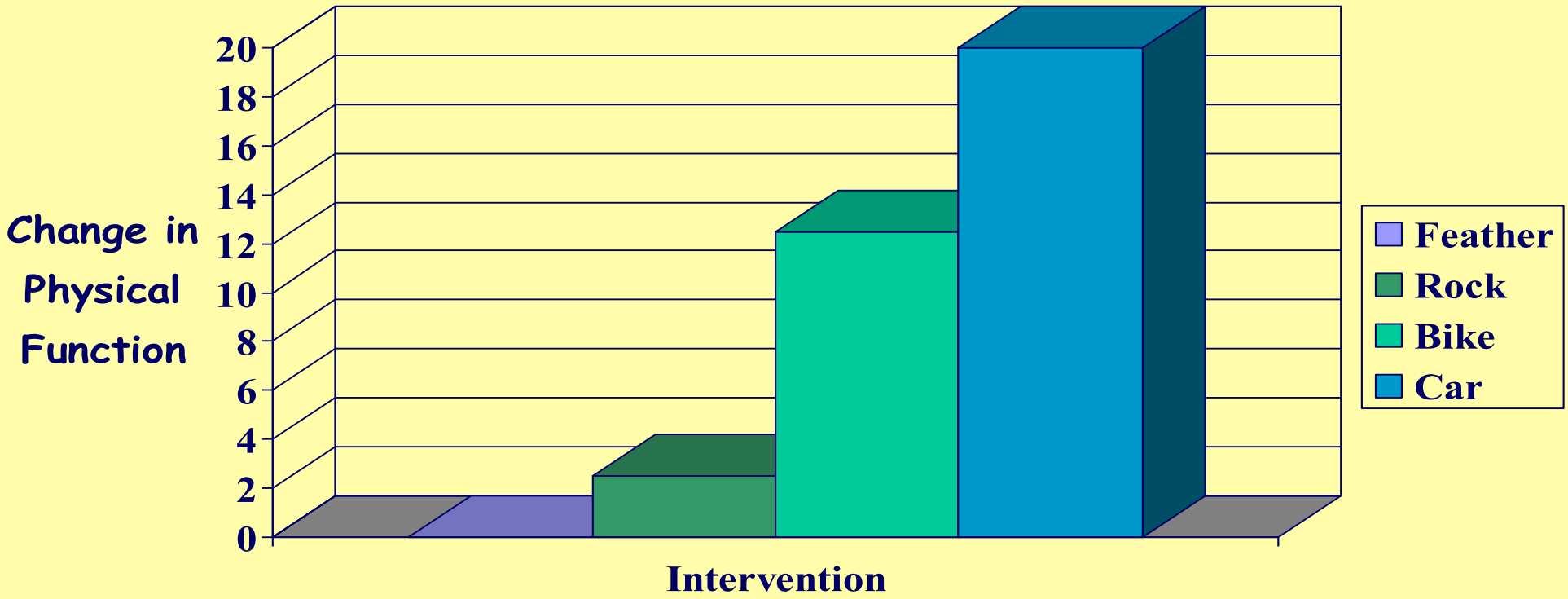
FDA Draft Guidance Document

For many wide used measures (pain, treadmill distance, HamD), the ability to show any difference between treatment groups has been considered evidence of a relevant treatment effect. If PRO instruments are to be considered more sensitive than past measures, it can be useful to specify a minimum important difference (MID) as a benchmark for interpreting mean differences" (p. 19)

Estimating the MID

- External anchors
  - Self-report
  - Provider report
  - Clinical measure
  - Intervention
- Anchor correlated with change on target measure at 0.371 or higher
- Anchor indicates “minimal” change

Hypothetical Change in Physical Function (T-score units) by magnitude of intervention



The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

1. Vigorous activities, such as running, lifting heaving objects, participating in strenuous sports
2. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
3. Lifting or carrying groceries
4. Climbing several flights of stairs
5. Climbing one flight of stairs
6. Bending, kneeling, or stooping
7. Walking more than a mile
8. Walking several blocks
9. Walking one block
10. Bathing or dressing yourself

Yes, limited a lot (0)/Yes, limited a little (50)/No, not limited at all (0)
Mean = 87; 75th percentile = 100 for U.S. males

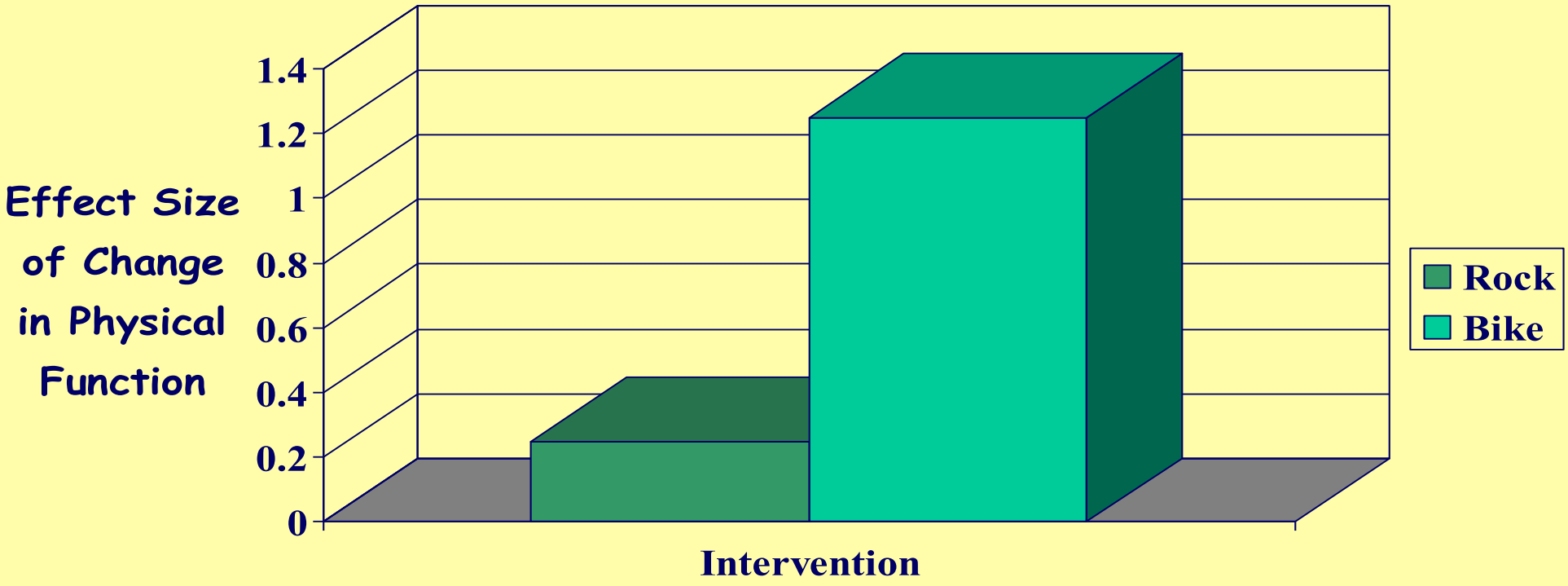
Change in Physical Function from Baseline

Baseline = 100 (U.S. males mean = 87, SD = 20)

- Hit by **Rock** causes me to be *limited a little* in vigorous activities and physical functioning drops to 95 (- 0.25 SD)

- Hit by **Bike** causes me to be *limited a lot* in vigorous activities, *limited a little* in moderate activities, and *limited a lot* in climbing several flights of stairs. Physical functioning drops to 75 (- 1.25 SD)

Getting Hit By Bike is > Minimal
Getting Hit by Rock is Closer to MID



Self-Report Anchor

- People who report a "minimal" change
- How is your physical health now compared to 4 weeks ago?

- *Much improved; Moderately Improved;*
- ***Minimally Improved****;*
- *No Change;*
- ***Minimally Worse****;*
- *Moderately Worse; Much Worse*

Example with Multiple Anchors

- 693 RA clinical trial participants evaluated at baseline and 6-weeks post-treatment.
- Five anchors:
  - 1) patient global self-report;
  - 2) physician global report;
  - 3) pain self-report;
  - 4) joint swelling;
  - 5) joint tenderness

Kosinski, M. et al. (2000). Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis and Rheumatism, 43, 1478-1487.

Patient and Physician Global Reports

- How the patient is doing, considering all the ways that RA affects him/here?

Very good (asymptomatic and no limitation of normal activities)

Good (mild symptoms and no limitation of normal activities)

Fair (moderate symptoms and limitation of normal activities)

Poor (severe symptoms and inability to carry out most normal activities)

Very poor (very severe symptoms that are intolerable and inability to carry out normal activities)

--> Improvement of 1 level over time

Global Pain, Joint Swelling and Tenderness

- 0 = no pain, 10 = severe pain; 10 centimeter visual analog scale
- Number of swollen and tender joints

-> 1-20% improvement over time

Effect Sizes (mean = 0.34) for SF-36 Changes Linked to Minimal Change in Anchors

Scale	Self-R	Clin.-R	Pain	Swell	Tender	Mean
PF	.35	.33	.34	.26	.32	.32
Role-P	.56	.52	.29	.35	.36	.42
Pain	.83	.70	.47	.69	.42	.62
GH	.20	.12	.09	.12	.04	.12
EWB	.39	.26	.25	.18	.05	.23
Role-E	.41	.28	.18	.38	.26	.30
SF	.43	.34	.28	.29	.38	.34
EF	.50	.47	.22	.22	.35	.35
PCS	.49	.48	.34	.29	.36	.39
MCS	.42	.27	.19	.27	.20	.27

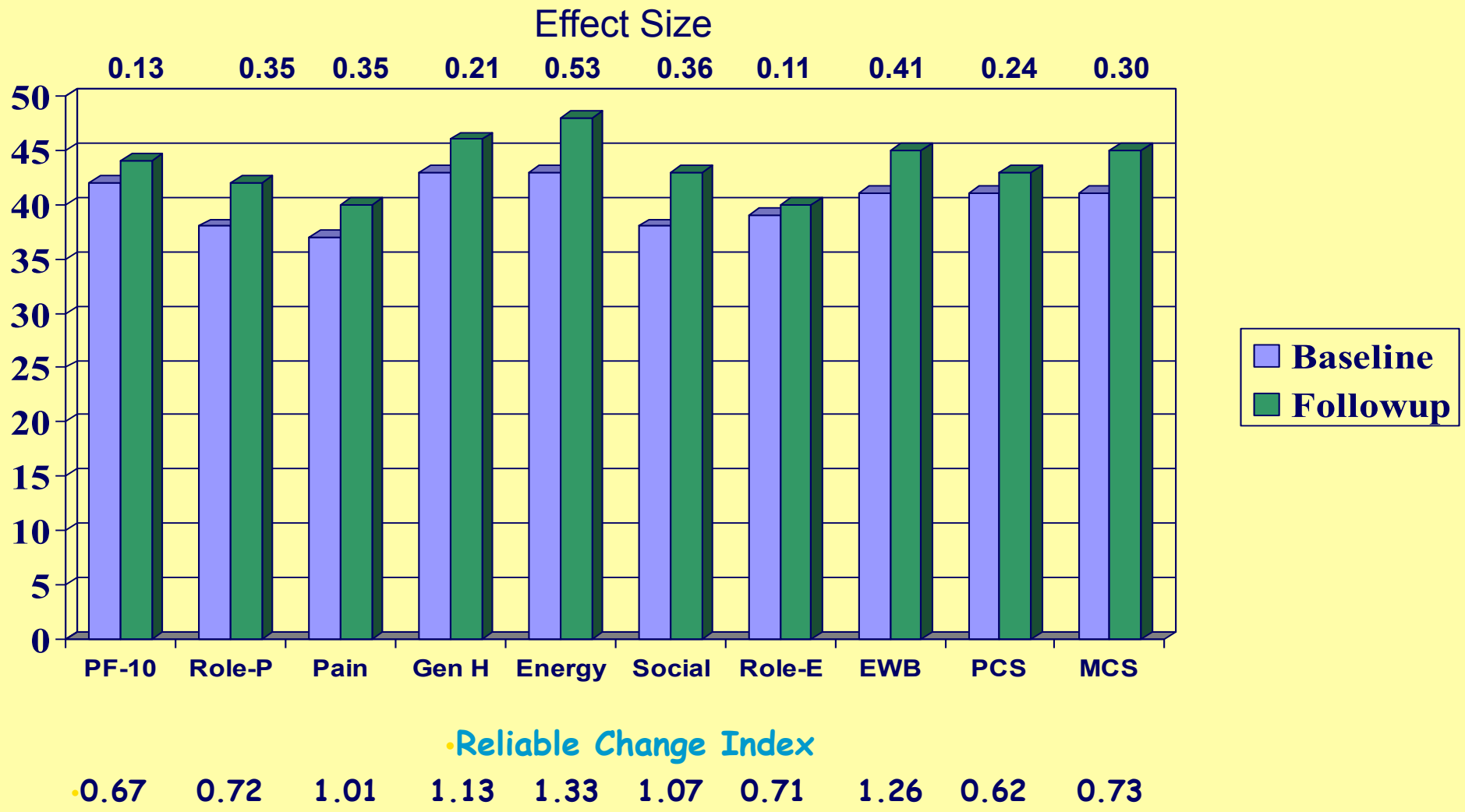
Use of "No Change" Group in Estimating MID

	Change #1 MID = ?	Change #2 MID = ?	Change #3 **MID = 4**
No Change on Anchor	Doesn't matter	+ 2	0, +1, or + 2
Minimal Change on Anchor	**0**	**+ 2**	**+ 4**

FDA Draft Guidance Document

- “When defining a meaningful change on an individual patient basis (i.e., a responder), that definition is generally larger than the minimum important difference for application to group mean comparisons” (p. 30).

Change in SF-36 Scores Over Time (n = 54)



Distribution-Based "Estimation" of MID

- Is not an estimate of the MID
- Is raw score difference derived from prior information about the MID
  - e.g., $D_{measure} = ES * SD_{measure}$
- Distribution-based formulas
  - Effect size (ES) = D/SD
  - Standardized Response Mean (SRM) = $D/SD^{\dagger}$
  - Responsiveness statistic (RS) = $D/SD^{\ddagger}$

SD = baseline SD; $SD^{\dagger}$ = SD of D; $SD^{\ddagger}$ = SD of D among "unchanged"

Standard Error of Measurement

- SEM = SD * SQRT (1-reliability)
- 95% CI = Estimated true score +/- 1.96 * SEM
- 1 SEM = 0.50 SD when reliability is 0.75

Summary

- No single best way to estimate MID
  - Use multiple anchors
  - Use anchors that represent minimum change
- Wide variation in estimates of MID
  - Report range, inter-quartile range, and confidence intervals around mean estimates.
- It is easier to conclude that a difference is clearly or obviously important than it is to say it is always unimportant.

Formulas for Significance of Individual Change

SEM 95% CI	$1.96 * SD_b * (1- \text{reliability})^{1/2}$
SEp 90% CI	$1.64* SD_b * (1- \text{reliability}^2)^{1/2}$
SEp 95% CI	$1.96* SD_b * (1- \text{reliability}^2)^{1/2}$
Estimated true score	Mean + reliability (score – mean)
Reliable change index	$X_2-X_1 / \sqrt{2}SEM.$

SD_b = standard deviation at baseline