

An Overview of Minimally Important Difference Estimation in Health-Related Quality of Life Studies

Ron D. Hays, Ph.D. (hays@rand.org)

October 14, 2004, 4-5pm

UCLA Departments of Medicine and Public Health

RAND Health Program

U.S.A.

How do we know how the patient is doing?

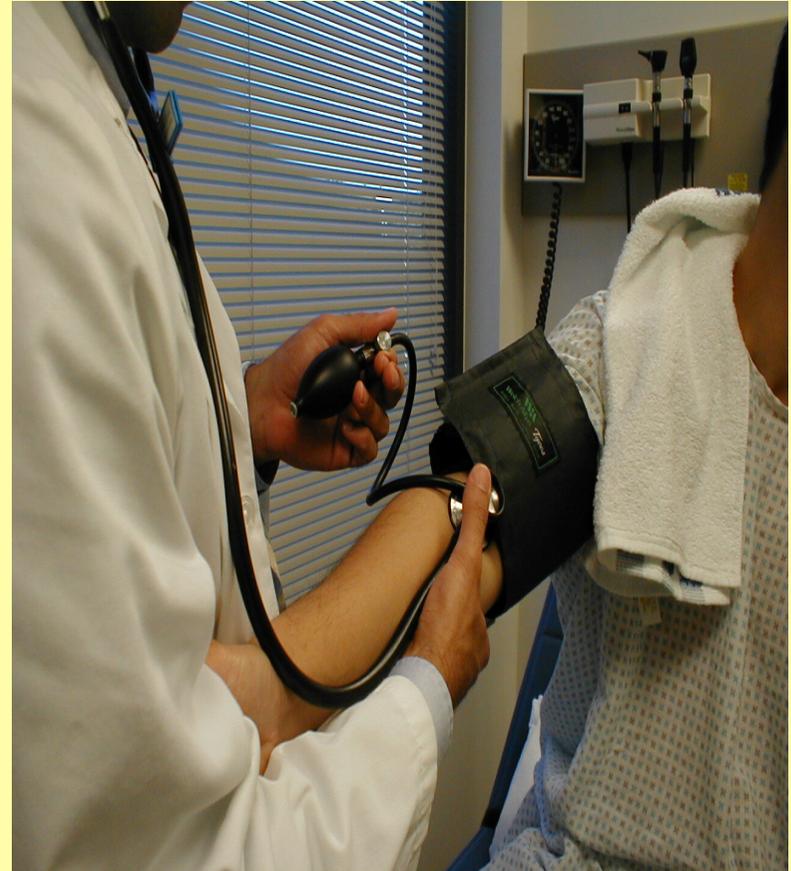
Temperature

Respiration

Pulse

Weight

Blood pressure



And by asking her or him about ...



◆ Symptoms

First RCT of Treatment for Newly Diagnosed Prostate Cancer (NEJM, 2002)

- ◆ Radical prostatectomy vs. watchful waiting
- ◆ - Trend to reduction in all-cause mortality
- ◆ (18% versus 15%; RR 0.83, 0.57 to 1.2, $p = 0.31$)

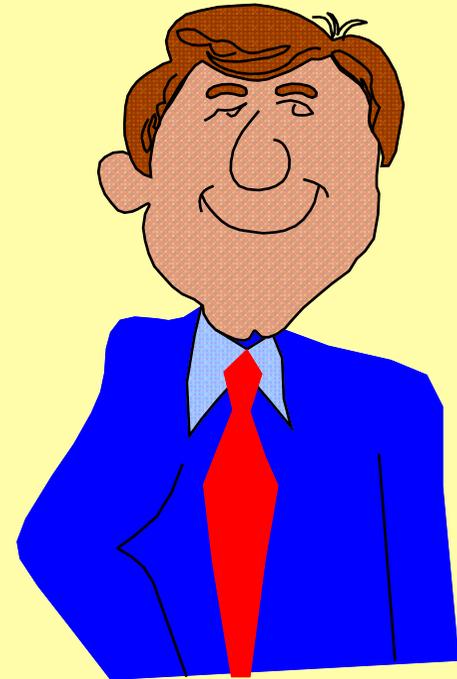
Impact on Symptoms

- ◆ Urinary obstruction (weak stream)
 - 44% waiting, 28% prostatectomy
- ◆ Sexual dysfunction and urinary leakage
 - 80% prostatectomy vs. 45% waiting
 - 49% prostatectomy vs. 21% waiting

Also, by talking to her or him about ...

What she or he is able to do

And how they feel about their life

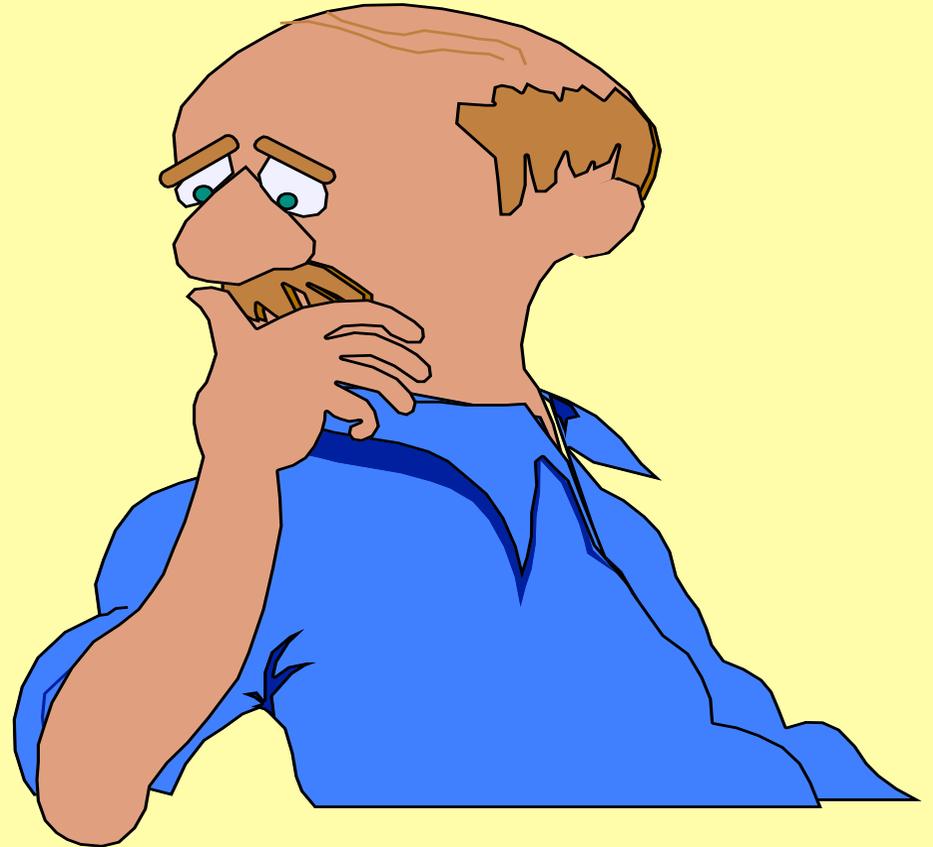


Does your health now limit you in walking more than a mile?

- ◆(If so, how much?)
- ◆*Yes, limited a lot*
- ◆*Yes, limited a little*
- ◆*No, not limited at all*

How much of the time during the past 4 weeks have you been happy?

- ◆ *None of the time*
- ◆ *A little of the time*
- ◆ *Some of the time*
- ◆ *Most of the time*
- ◆ *All of the time*



In general, how would you rate your health?

- ◆ Excellent
- ◆ Very Good
- ◆ Good
- ◆ Fair
- ◆ Poor

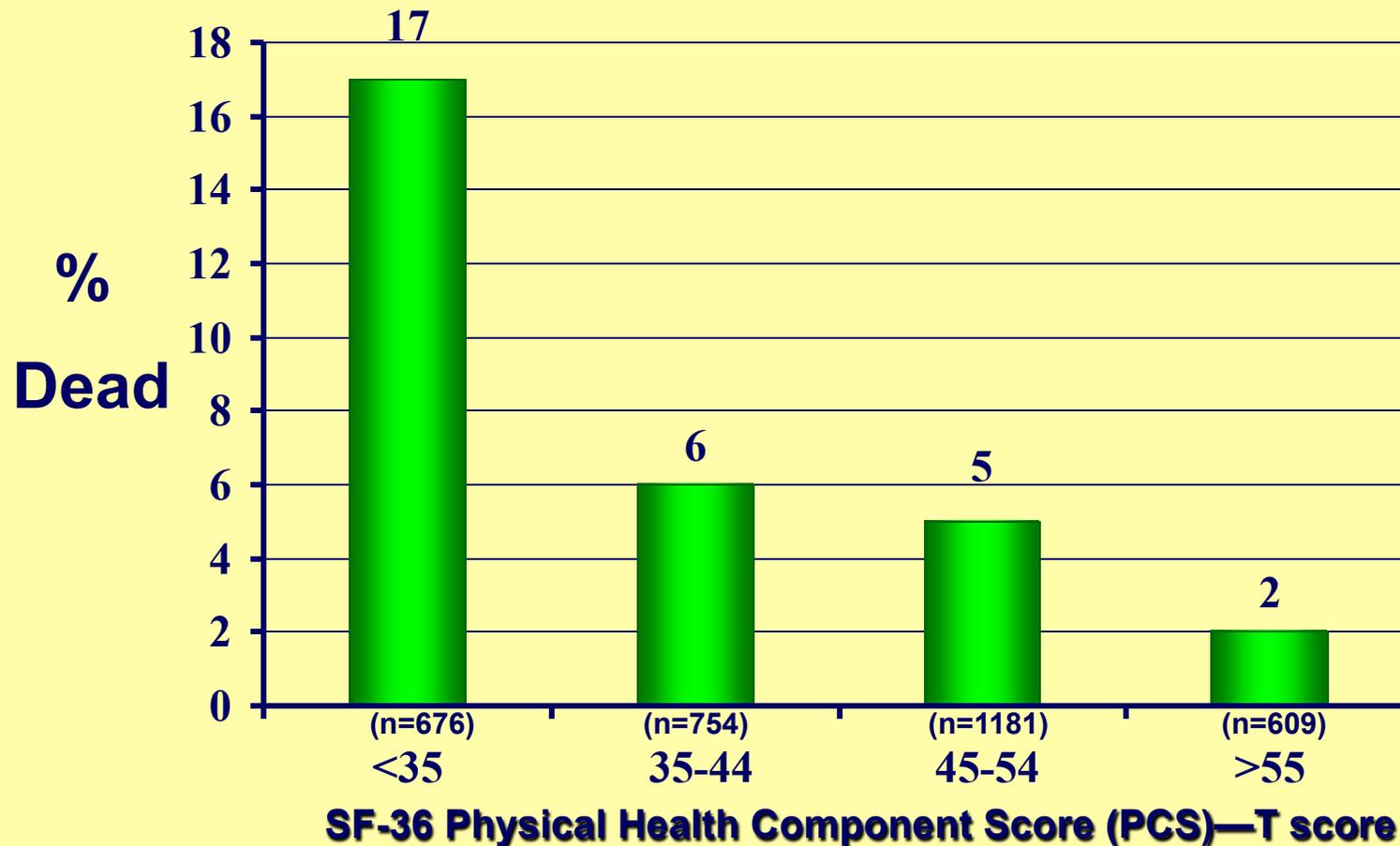
Health-Related Quality of Life is:

- ◆ How the person FEELS (well-being)
 - Emotional well-being
 - Pain
 - Energy
- ◆ What the person can DO (functioning)
 - Self-care
 - Role
 - Social

Self-Report Reliability Comparable to Traditional Clinical Measures

- ◆ 0.80-0.90 for blood pressure
- ◆ 0.70-0.90 for multi-item self-report scales

Self-Reports are Valid—For example, Physical Health Predicts 5-Year Mortality



Ware et al. (1994). SF-36 Physical and Mental Health Summary Scales: A User's Manual.

Minimally Important Difference (MID)

- ◆ One can observe a difference between two groups or within one group over time that is statistically significant, but the difference could be 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.

Terminology

Minimally Important Difference (MID) or Minimal difference (MD)

- > Minimally Detectable Difference (MDD)
- > Clinically Important Difference (CID)

Obviously Important Difference (OID)

Distribution-Based “Estimation” of MID

- ◆ Provides no direct information about the MID

- Effect size (ES) = D/SD
- Standardized Response Mean (SRM) = D/SD^\dagger
- Guyatt responsiveness statistic (RS) = D/SD^\ddagger

D = raw score change in “changed” group;

SD = baseline SD;

SD^\dagger = SD of D;

SD^\ddagger = SD of D among “unchanged”

Standard Error of Measurement

- ◆ $SEM = SD * \text{SQRT}(1 - \text{reliability})$
- ◆ 1 SEM = 0.50 SD when reliability is 0.75

Estimating the MID

- ◆ External anchor to determine there has been “minimal” change
 - Self-report
 - Provider report
 - Clinical measure
 - Intervention
- ◆ Estimate change in HRQOL among those with minimal change on anchor

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/her?

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

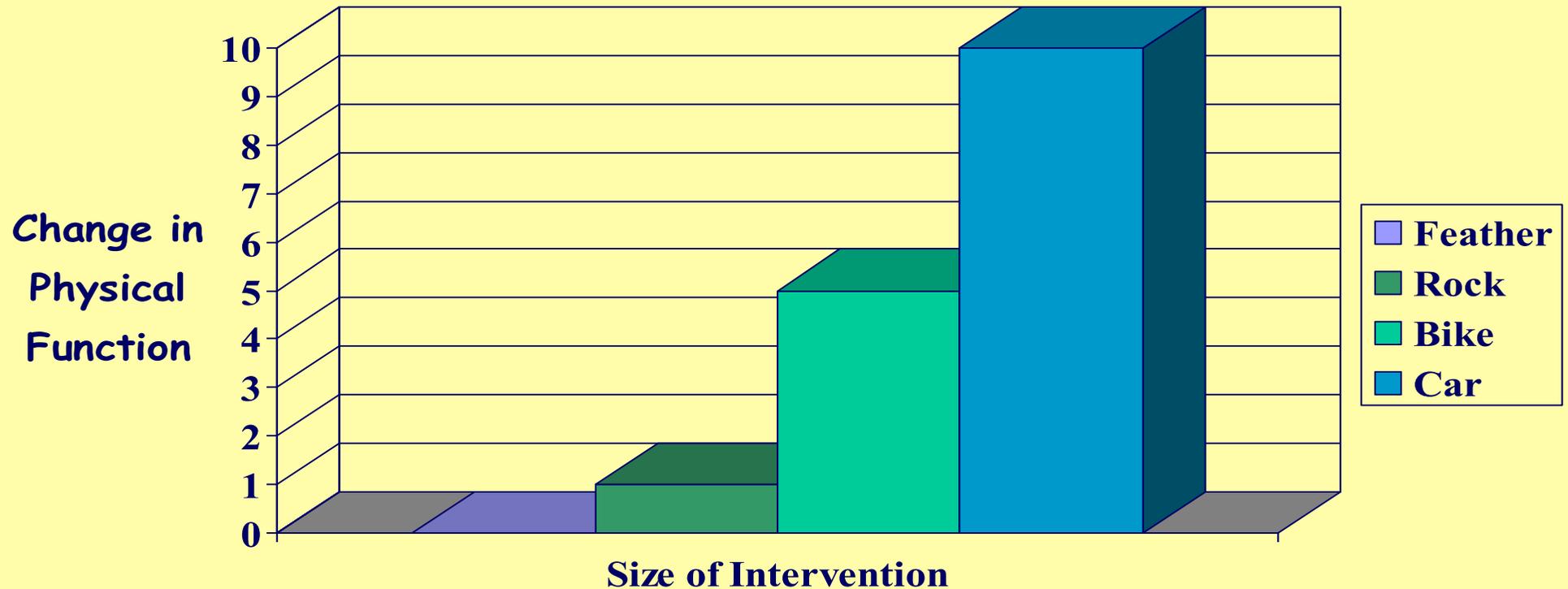
Norman, Sloan, Wyrwich (2003)

- ◆ “Interpretation of Changes in Health-related Quality of Life: The remarkable universality of half a standard deviation”
- ◆ Table 1 reports estimates of MIDs for 33 published articles. “For all but 6 studies, the MID estimates were close to one half a SD (mean = 0.495, SD = 0.155)” (p. 582).

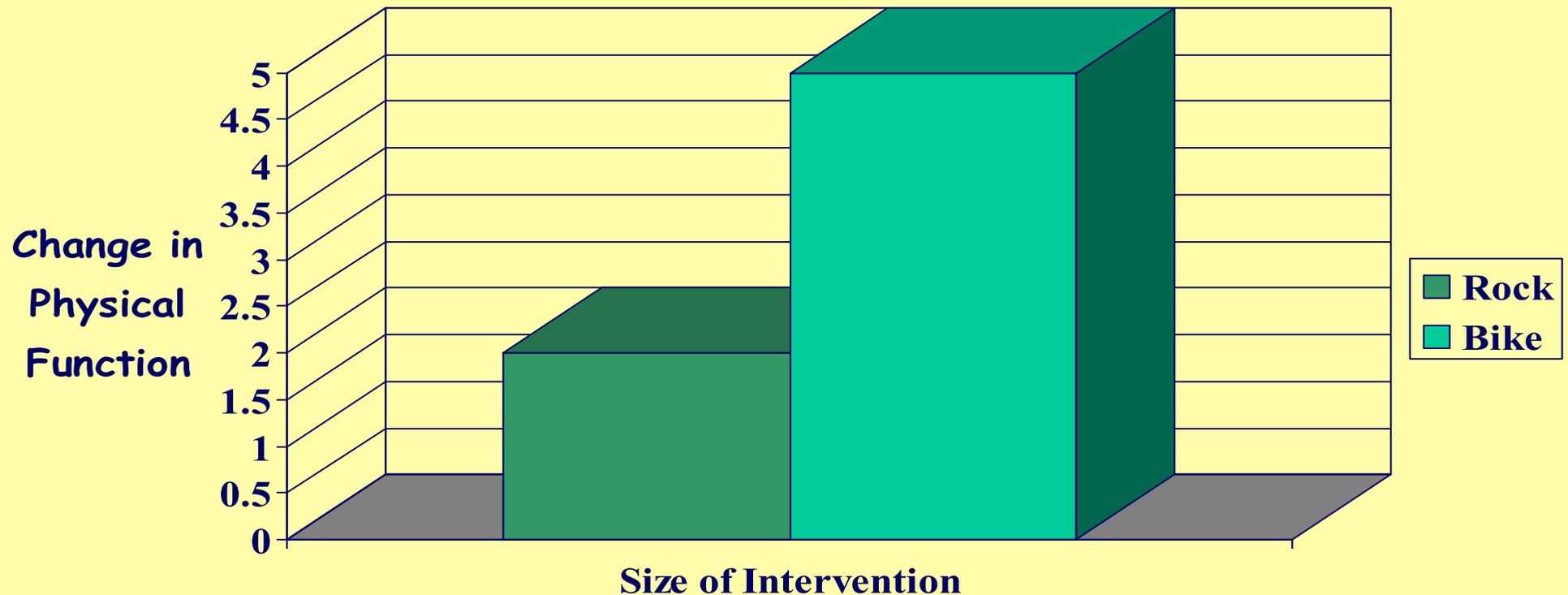
Why not accept 0.50 SD as MID?

- ◆ Based on 33 published articles.
 - While 33 may seem like a large number of studies, not really a very large sample size.
- ◆ Problems with Norman et al. paper
 - Selective reporting of HRQOL results
 - Included an article based on a 6-minute walk test
 - Included articles with anchors that did not necessarily represent minimal change
 - Included articles with no estimates of MID
- ◆ Wide variation in estimates of MID

Change in Physical Function by Intervention



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



ES derived from assumed MID differences

- ◆ Wyrwich et al. (1999) studied 605 CAD/CHF patients and Wyrwich et al. (1999) evaluated 417 COPD patients.
- ◆ No anchors were used in these studies. ES of 0.36 and 0.35 for the CHQ and CRQ were based on previously reported MID recommendations.
- ◆ $ES = 0.35$ for CRQ is simply the ratio of the previously reported MID of 0.5 per item divided by the standard deviations observed in sample of 417 COPD patients.

Wide variation in MID estimates

- ◆ Median of the mean ES for studies was 0.42.
- ◆ Range = 0.11 to 2.31
- ◆ SD of mean ES = 0.31
- ◆ Coefficient of variation = 64%

Recommendations for Estimating the MID

- ◆ Estimating the MID is challenging--it is easier to conclude that a difference is clearly or obviously important than it is to say one is always unimportant.
- ◆ No one 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.

Value of “Control Group” in Estimating MID

	Change #1 MID = ?	Change #2 MID = ?	Change #3 MID = 4
No Change on Anchor	- 4	+ 2	+ 2
Minimal Change on Anchor	0	+ 2	+ 4

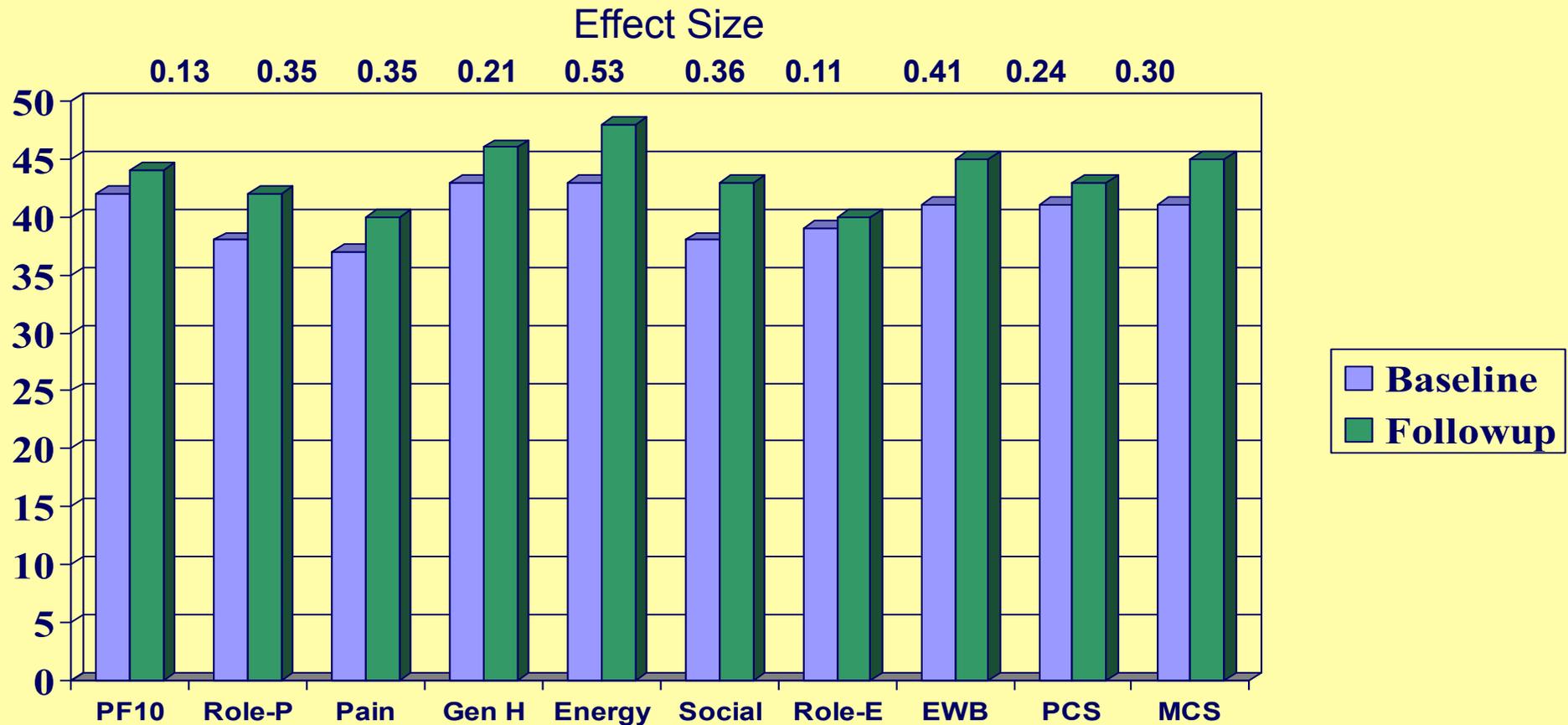
Norman, Sloan, & Wyrwich (2004)

- ◆ “The size of difference that is important for individual patient change exceeds the size for group differences because of the larger error associated with individual assessment” (Farivar et al., 2004)

Norman, Sloan, & Wyrwich (2004)

- ◆ “We seriously question this point, and hope that other health services outcome researchers will also re-examine this conclusion. We agree that there is more error in an individual estimate than a group estimate or mean. However, if an individual wants their HRQOL score to improve by a certain amount, much like setting a goal of losing 5 lbs on a diet, it is irrelevant how much their weight (or scale) varies from day to day. Likewise, if we calculate change in HRQOL across many patient, the group difference is only the average of the individual differences, and hence it is not necessarily larger or smaller than each individual’s goals” (p. 583-584).

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



Effect Size for Significant Individual Change

PF-10	0.67
RP-4	0.72
BP-2	1.01
GH-5	1.13
EN-4	1.33
SF-2	1.07
RE-3	0.71
EWB-5	1.26
PCS	0.62
MCS	0.73

Norman, Sloan, & Wyrwich (2004)

- ◆ “Finally, it is important to note that the examination of the MID in health services research has focused on group level comparisons. In contrast, parallel work in psychology has emphasized differences for individual patients that are clinically significant. **The size of difference that is important (MID) for individual patient change exceeds the size for group differences because of the larger error associated with individual assessment**” (Farivar et al., 2004)

Bibliography

- ◆ Farivar, S. S., Liu, H., & Hays, R. D. (2004). Half standard deviation estimate of the minimally important difference in HRQOL scores?. Expert Review of Pharmacoeconomics and Outcomes Research, 4 (5), 515-523.
- ◆ Hays, R. D., Farivar, S. S., & Liu, H. (in press). Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. Journal of COPD.
- ◆ Hays, R. D., & Woolley, J. M. (2000). The concept of clinically meaningful difference in health-related quality-of-life research: How meaningful is it? PharmacoEconomics, 18, 419-423.

Useful URLs

- ◆ <http://gim.med.ucla.edu/FacultyPages/Hays/>
- ◆ <http://www.rand.org/health/surveys.html>
- ◆ <http://www.qolid.org/>
- ◆ www.sf-36.com