

# Responsiveness to Change and Minimally Important Difference

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

March 3, 2003

# Construct Validity

- Does measure relate to other measures in ways that are consistent with hypotheses?
- Responsiveness to change

# Relative Validity Analyses

- Form of "known groups" validity
- Relative sensitivity of measure to important clinical difference
- One-way between group ANOVA

# Relative Validity Example

## Severity of Heart Disease

	Severe	Mild	None	F-ratio	Relative Validity
Scale #1	87	90	91	2	---
Scale #2	74	78	88	10	5
Scale #3	77	87	95	20	10

# Responsiveness to Change

- HRQOL measures should be responsive to interventions that change HRQOL
- Evaluating responsiveness requires assessing HRQOL relative to an external indicator of change (anchor)

# Two Essential Elements

1. External (not HRQOL measure being evaluated) indicator of change (Anchor)
2. Amount of HRQOL change among those determined to have changed on anchor, relative to noise (variance).

# Responsiveness Indices

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

(2) Standardized Response Mean (SRM) =  $D/SD^\dagger$

(3) 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"

# Kinds of Anchors

- Self-report
- Clinician or other report
- Clinical parameter
- Clinical intervention

# Self-Report Anchor (A)

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*

# Self-Report Anchor (B)

Rate your overall condition. This rating should encompass factors such as social activities, performance at work or school, seizures, alertness, and functional capacity; that is, your overall quality of life.

7 response categories; ranging from no impairment to extremely severe impairment

# Examples of Other Anchors

## Clinician report

- How is Jan's physical health now compared to 4 weeks ago?

## Clinical parameter

- Change from CDC Stage A to B
- Became seizure free

## Clinical intervention

- Before and after Prozac

# Mean Scores of NEI RQL Scales Before and After Surgery (n = 185)

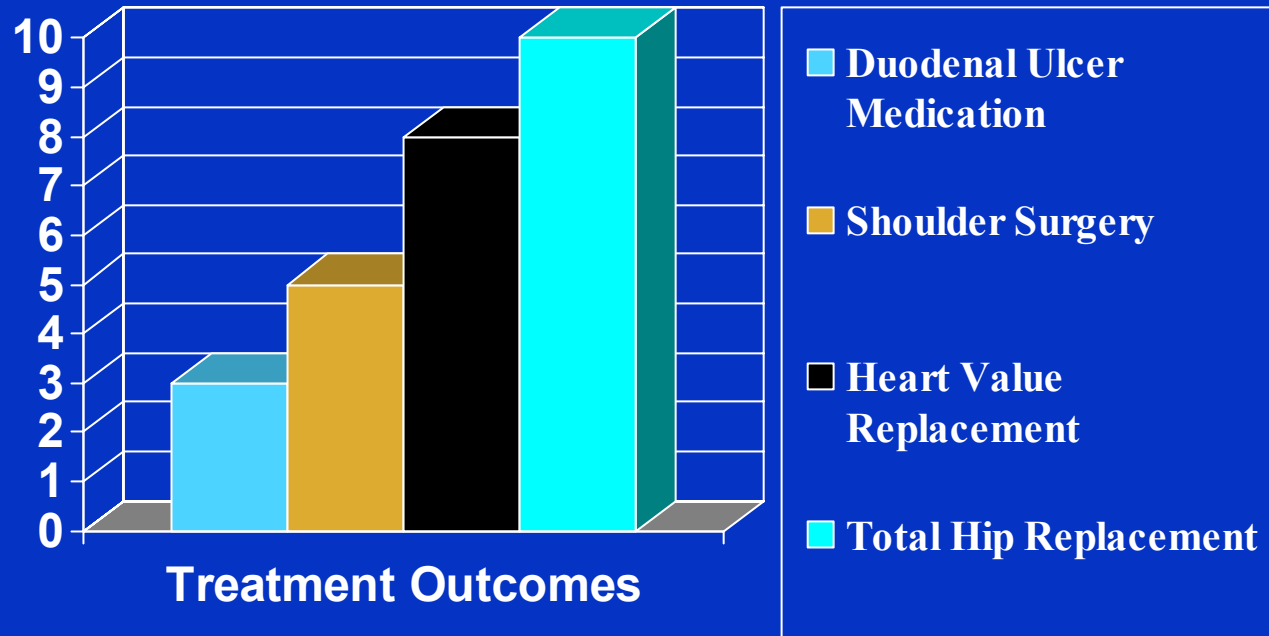
Scale	Before	After	Change	t-statistic	p-value	Ratio of F-statistics
Clarity of vision	83.31	84.95	1.64	1.17	0.2431	0.003
Expectations	14.05	55.81	41.76	13.4	<.0001	0.203
Near vision	78.40	87.72	9.32	5.88	<.0001	0.039
Far vision	81.92	89.38	7.46	6.75	<.0001	0.052
Diurnal fluctuations	72.21	76.62	4.41	2.18	0.0305	0.005
Activity limitations	64.28	93.46	29.18	14.68	<.0001	0.244
Glare scale	74.73	67.09	-7.64	-3.13	0.0020	0.011
Symptoms	78.53	84.76	6.23	4.70	<.0001	0.025
Dependence on correction	26.08	83.85	57.77	29.71	<.0001	1.000
Worry	64.53	77.64	13.11	7.90	<.0001	0.071
Suboptimal correction	86.21	96.55	10.34	5.90	<.0001	0.040
Appearance	64.28	91.79	27.52	10.80	<.0001	0.132
Satisfaction with correction	56.41	82.61	26.20	11.41	<.0001	0.147

# Responsiveness to Change (n=185)

Scale	Effect Size (ES)	Standardized Response Mean (SRM)	Responsiveness Statistic (RS)	t-statistic
Clarity of vision	0.11	0.09	0.15	1.17
Expectations	1.77	0.99	1.66	13.40
Near vision	0.45	0.43	0.86	5.88
Far vision	0.53	0.50	0.85	6.75
Diurnal fluctuation	0.19	0.16	0.26	2.18
Activity limitation	1.18	1.08	2.23	14.68
Glare	-0.29	-0.23	-0.38	-3.13
Symptoms	0.36	0.35	0.74	4.70
Dependence on correction	2.29	2.18	2.92	29.71
Worry	0.62	0.58	0.93	7.90
Suboptimal correction	0.46	0.45	0.67	5.90
Appearance	0.90	0.80	1.25	10.80
Satisfaction	1.10	0.84	1.68	11.41

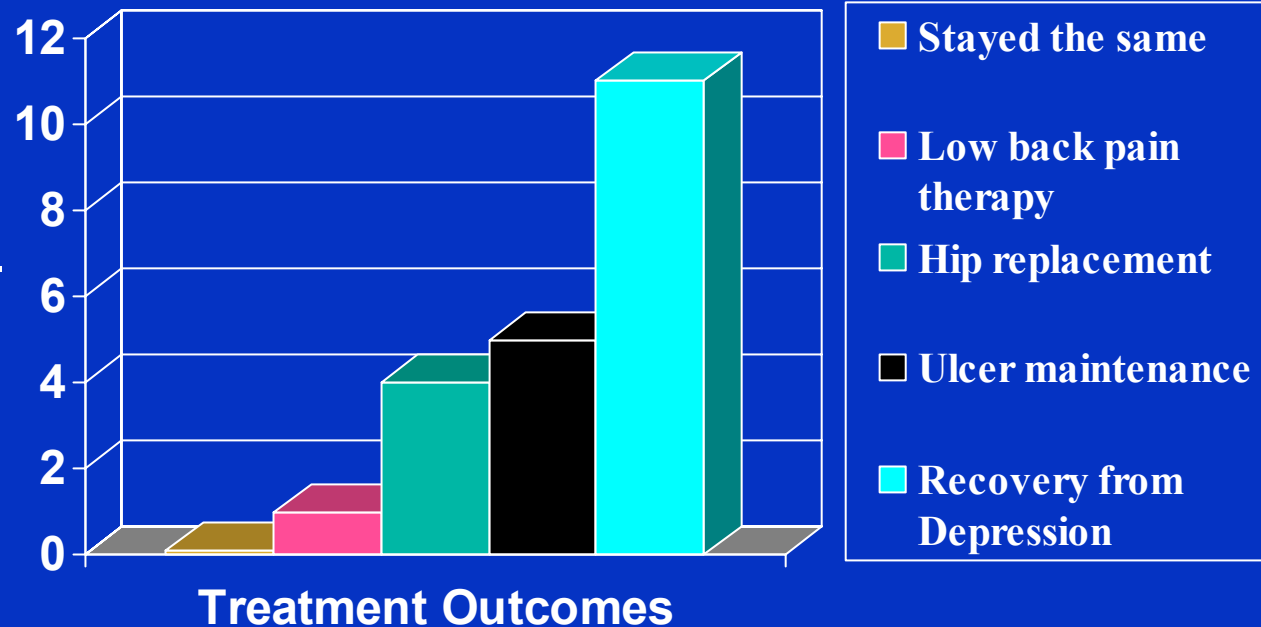
# Change and Responsiveness in PCS Depends on Treatment

Impact on SF-36 PCS



# Change and Responsiveness in MCS Depends on Treatment

Impact on SF-36 MCS



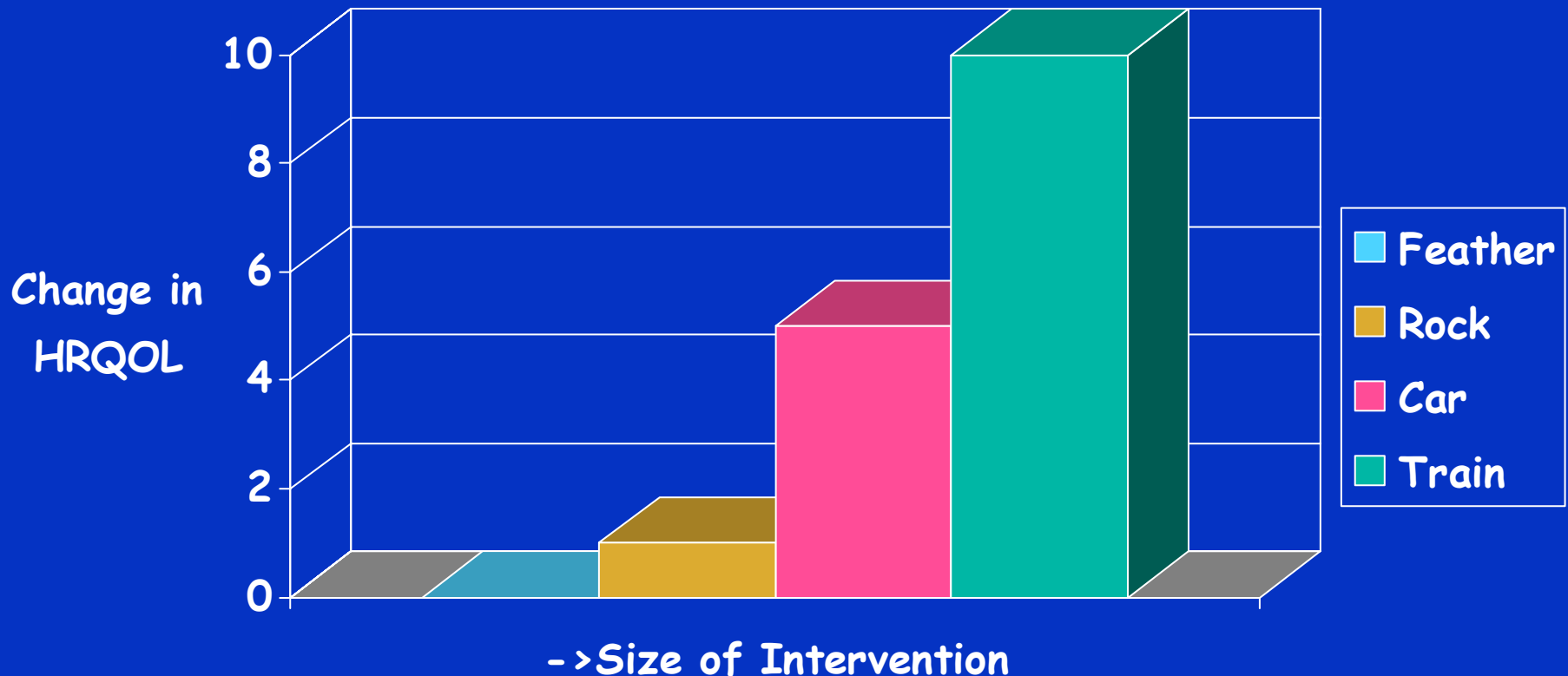
# So How Big Are Different Changes?

## Effect size benchmarks

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



# Magnitude of HRQOL Change Should Parallel Underlying Change



# Minimal Clinically Important Difference (MCID)

Smallest difference in score that is worth caring about (important).

- Some differences between groups or over time may be so small in magnitude that they are not important.

Change large enough for a clinician to base treatment decisions upon it.

# Two Essential Elements

1. Indicator (not HRQOL measure being evaluated) of “minimal” change (Anchor)
2. Amount of HRQOL change among those determined to have changed on anchor.

# Example Anchor (1)

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*

# MID for SF-36 is:

“typically in the range of 3 to 5 points” (p. 149).  
{.09-→0.28 ES}

Samsa, G., Edelman, D., Rothman, M. L., Williams, G. R., Lipscomb, J., & Matchar, D. Pharmacoeconomics, 15, 141-155: 1999.

Does this mean that 1-2 points on the SF-36 scales are unimportant?

# MID Varies by Anchor

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.

# Changes in SF-36 Scores Associated with Minimal Change in Anchors

Scale	Self-R	Clin.-R	Pain	Swell	Tender	Mean
PF	8	8	8	6	8	8
Role-P	21	20	11	13	13	16
Pain	15	12	8	12	7	11
GH	4	2	2	3	1	2
EWB	7	5	5	3	1	4
Role-E	18	12	8	16	11	16
SF	12	9	8	8	10	9
EF	11	10	5	5	8	8
PCS	4	4	3	3	3	4
MCS	5	3	2	3	2	3

# MID Varies by Starting Position and Direction of Change

Same retrospective report of change associated with bigger prospective change for those with more room to change

- Among those who said their physical health was *somewhat worse*, change ranged from -26 points to +3 points for people with high (81-100) versus low (0-20) baseline physical health (Baker et al., 1997, Medical Care).

# Group Average is Different from Individual Change

MID “may differ across groups of patients defined by diseases, conditions, levels of severity, socioeconomic status, nationality and various other factors as well” (Samsa et al., 1999, p. 142).

Average change collapses across individual responses.

Is it reasonable to infer the minimum amount of change that is important for individuals based on a group average?

What if SF-36 scores improved by 4 points for half the people and 0 points for the other half?

# MID Determination Complicated By Cumulative Change Over time

Baseline 42

Year 4 36

Note: 4-year decline in PCS among US seniors, 1990-94.

-> 1.5 points per year (0.15 SD)



# Value Depends on Cost



A small positive change has greater value if it costs less.

Importance of HRQOL change depends on what it costs to produce it.

# Summary

Identification of MID aids interpretation by providing familiar anchors to unfamiliar units.

Trying to give a single point estimate is too simplistic.

Bounded estimates are necessary given the uncertainty.

