Concept of measurement: RA-defining decrements in physical functioning

Jasvinder Singh, MD, MPH
Associate Professor of Medicine and Epidemiology, University of Alabama at Birmingham
Staff Physician, Birmingham VA Medical Center

Toward Consensus Development: Qualifying Endpoint Measures for Rheumatoid Arthritis Clinical Trials 08/28/2012, Silver Spring, MD

Outline

- Physical Function domain
- Measures of Physical Function (PF)
  - Generic measure: Short-form 36
  - Disease-specific measure: Health assessment Questionnaire (HAQ), MACTAR, AIMS2 etc.
- Measurement properties of PF measure (HAQ)
- Evidence of Sensitivity to Change of PF measures

What is PHYSICAL FUNCTION?

- NHANES:
  - “An individual's level of physical performance is a reflection of their overall health, and the impact of several chronic diseases ... such as arthritis … on the ability to function without limitations in the course of daily life.
  - incorporates aspects of strength, mobility, freedom of movement, balance and coordination. “

NHANES.
http://www.cdc.gov/nchs/data/nhanes/nhanes3/cdrom/nchs/manuals/physical.pdf

Modified McMaster Toronto Arthritis patient preference questionnaire (MACTAR)[1]: or Patient Elicitation Technique (PET)

- Asks patients to rank their most affected activity
- "Do you think your arthritis limits your ability to carry out activities you did before you had arthritis?" No Yes
- Please tell me which activities are affected
  - _____
  - _____
- Which of these activities would you most like to be able to do without the pain or discomfort of your arthritis?
- How would you say your overall physical functioning has been? For example, over the last 2 weeks would you call your physical function... Good Good to fair Fair Fair to poor Poor

[1]Tugwell P et al. J Rheumatol. 1987 Jun;14:446-51

Problem Elicitation Technique: Selection of Top 5 Physical Activities: US301 [1]

	Frequency (n=482)	%
Do chores	204	43
Stand from Chair	203	42
Dressing Self	195	41
Get in/out of bed	163	34
Get down 5-lb bag	160	33
Open milk cartons	148	30
Take a tub bath	147	30
Open jars previously opened	145	25
Shampoo hair	118	23
Climb up 5 steps	112	23

[1]Tugwell P et al. Arthritis Rheum 2000; 43 (3): 506-14

Problem Elicitation Technique: Selection of Top 5 Physical Activities: US301

	Frequency (n=482)	%
Walk outdoors on flat ground	110	22
Get in/out of car	106	19
Run errands and shop	92	18
Turn faucets on/off	85	17
Open car doors	83	17
Cut meat	83	17
Bend to pick up clothing	82	17
Lift glass to mouth	64	13
Wash and dry body	62	8
Get on/off toilet	38	8

Arthritis Impact Measurement Scale (AIMS 2) [1-3]

- 67 items, burdensome for patients to complete; charge for use
- Measures limitations in and impact of arthritis on 5 scales: lower extremity function, upper extremity function, affect, pain, and social interactions
- 3 additional scales evaluate arm function, work, and social support
- Scored for last month: All, most, some, few or no days
- Asks patients to identify areas they would like to see improved
- Each score can be expressed in the range 0-10, with 0 representing good health and 10 representing poor status

[1] Meenan R. Arthritis Rheum 1980; 23(2):146-152
[2] Ren XS, Kazis L, Meenan RF. Arthritis Care Res. 1999 Jun; 12(3):163-71
[3] Meenan RF et al. Arthritis Rheum 1992; 35:1-10.

HAQ-DI: What is it? [1]

- 20 items in 8 categories (dressing, arising, eating, walking, hygiene, reaching, gripping, outside activities) over past week
- measured on a 4-point ordinal scale from 0 to 3:
  - 0 without any difficulty;
  - 1 with some difficulty;
  - 2 with much difficulty; [requires aids and/or help] and
  - 3 unable to do.
- Highest score in each category averaged into "disability index"
  - 0 = none to 3 = "disabled"
  - Score adjusted for devices/aids from 0 or 1 to 2

[1] http://aramis.stanford.edu/HAQ.html

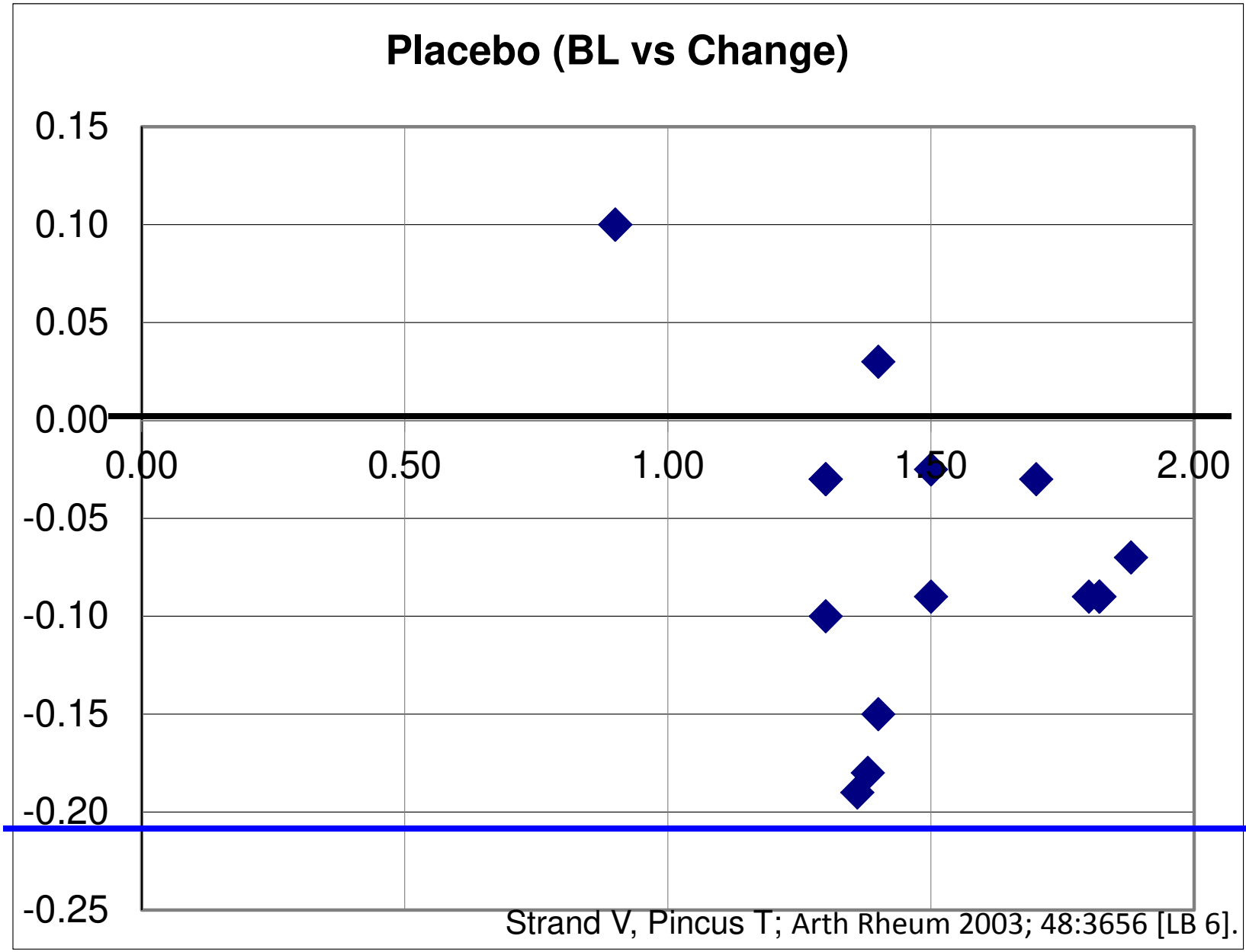
HAQ in Trials: The Success Story

Improvement in HAQ-DI in Trials: Comparison of Active vs. Placebo[1]

- In an analysis, HAQ-DI changes at 6- to 12-months in RCTs of approved agents in RA were compared
- It is difficult to compare HAQ improvements across RCTs
- Changes reported mean or median
- Comparison of mean change within treatment groups
  - Subtract final from baseline scores
  - To account for baseline scores

[1]Strand V and Pincus T. Arth Rheum 2003; 48:3656 [LB 6]

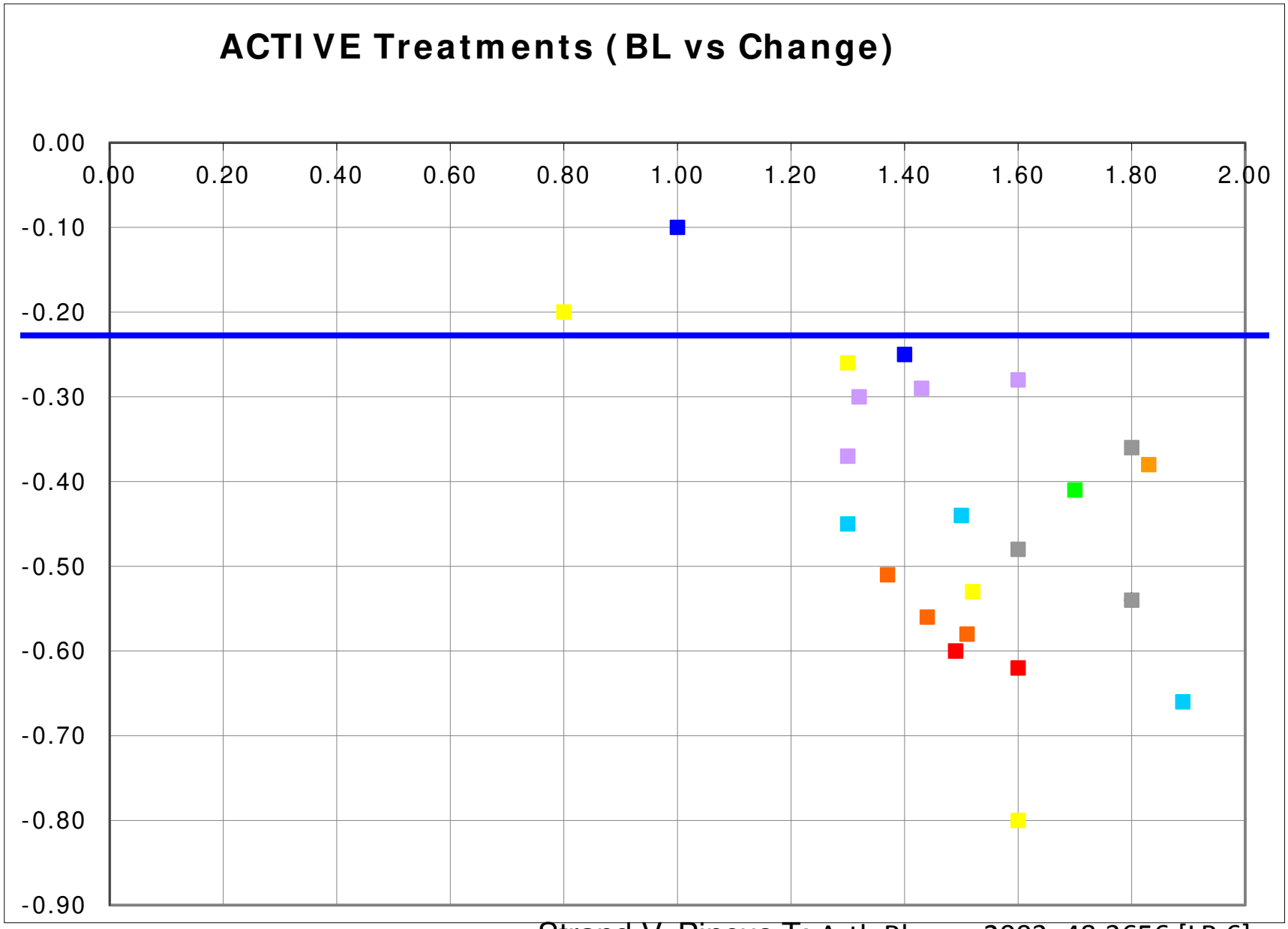
Mean Changes in HAQ DI with Placebo Rx



Strand V, Pincus T; Arth Rheum 2003; 48:3656 [LB 6].

Mean Changes in HAQ DI with Active Rx

MCID



Strand V, Pincus T; Arth Rheum 2003; 48:3656 [LB 6].

HAQ-DI: Face and Content Validity

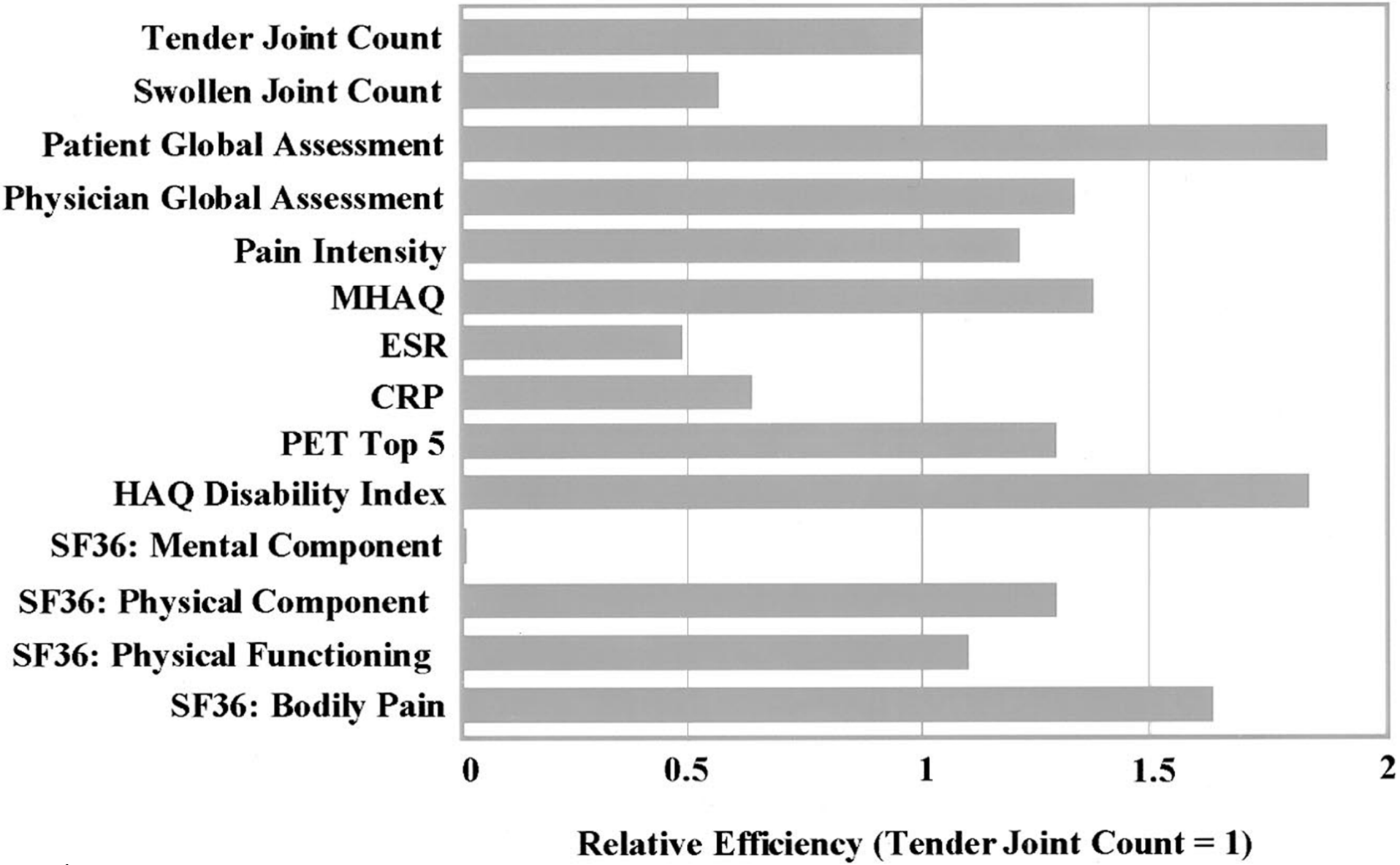
Comparison of Function Instruments

Activities Assessed	MHAQ	HAQ-DI	SF-36 PF
Walking	x	x	x
Climbing Steps	x	x	
Reaching	x	x	
Getting in and out of a car	x	x	
Arising	x	x	
Reaching over head		x	
Gripping	x	x	
Eating	x	x	x
Self care ADLs			
Hygiene	x	x	
Dressing, grooming	x	x	x
Instrumental Activities			x
Discretionary activities			
Walking > 1 mile			x
Climbing several sets of stairs			x
Moderate activities			
x			

HAQ: Post-Hoc Face, content validity

- 12-month, RCT of leflunomide 20 mg vs. placebo vs. methotrexate 15 mg/week in US 301
- N=482
- Analysis of 438 who completed baseline and at least 1 FU visit
- 73% female, age, 54 years,
- Disease duration, 7 years, 65% RF positive
- Relative efficiency compared to tender joint count: standardized effect size differences

Relative efficiency: HAQ and other measures



[1]Tugwell P et al. Arthritis Rheum 2000; 43 (3): 506-14

HAQ-DI: Floor and Ceiling Effects

HAQ: floor and Ceiling Effects

- Leflunomide 301 study [1]:
  - Ceiling effect:
    - Role emotional SF-36 (41%)
    - HAQ-DI (2%)
- Pincus MHAQ validation study [2]:
  - 23 of 144 patients (16%) reported normal HAQ scores of 0.
  - 20/23 (87%) reported some abnormality on "advanced" ADL.
- Fries et al [3]:
  - 10-15% at ceiling with HAQ-DI
- Floor Effect: Less likely an issue with early RA treatment

[1]Tugwell P et al. Arthritis Rheum 2000; 43 (3): 506-14
[2]Pincus T et al. Arthritis Rheum 1999; 42(10):2220–2230
[3]Fries J et al. J Rheumatol 2011;38;1759-1764

HAQ-DI: Construct validity

HAQ: Four Nation Study[1]

- 4 Nation Study: French, Dutch, Irish, Norwegian
- RA by ACR 1987, duration 0-4 years
- Significant Predictors of HAQ disability:
  - Ritchie index, sex, ESR, age, and disease duration
- Explained 44% adjusted variance in HAQ score
- No correlation of RF with HAQ

[1]Smedstad LM et al. Br J Rheumatol 1996;35:746-51

HAQ: Construct validity in a population-based sample[1]

- Random sample of 1,530 Finnish adults in the Central Finland District
- Mean, 0.25 (95% CI, 0.22–0.28)
- 32% of respondents had at least some disability
- Disability
  - Increased with age
  - Lower in those with more education, lower BMI, and increasing frequency of physical exercise
- HAQ-DI correlated with
  - Pain (r=0.58; 95% CI, 0.55–0.62)
  - Patient global self-assessment (r=0.61; 95% CI 0.58–0.65)

[1]Krishnan E et al. Arthritis Rheum 2004;**50**: 953-960.

HAQ-DI: Discrimination and Efficiency

Effect Sizes with DMARDs in RA[1]

Publication	Study Arms	Effect Size
Proudman 2000	Sulfasalazine (SSZ)	1.00
	MTX/CSA/Steroids	0.75
Mottonen 1999	DMARD combination	0.66
	DMARD monotherapy	0.66
Ziedler 1998	Cyclosporin	0.66
	IM Gold	0.66
Haegsma 1997	Sulfasalazine	0.33
	Methotrexate	0.63
	Combination	0.63
Tiley 1995	Minocylcine	0.72
	Placebo	0.55
HERA study 1995	Hydroxychloroquine	0.56
	Placebo	0.41

[1]Scott and Strand. Rheumatology 2002; 41: 899-909 .

HAQ MCID with Biologics in RA[1]

Publication	Disease Duration	Study Arms	% with HAQ MCID
EARLY RA			
Emery 2008, Kekow 2009 (COMET)	0.8 years	ETN + MTX MTX	55% 39%
St. Clair 2004 (ERA)	0.9 years	IFN 3 mg/kg + MTX IFN 6 mg/kg + MTX MTX	76% 76% 65%
Breedveld 2006; Kimel 2008 (PREMIER)	0.7 years	ADA + MTX ADA MTX	72% 58% 63%
Bejarano 2008 (PROWD)	0.8 years	ADA + MTX MTX	-0.7 (HAQ change) -0.4
ESTABLISHED RA			
Schiff 2006	>3 years	ETN + MTX ETN	71% 58%
Keystone 2009 (GO-FORWARD)	6 years	GOL 50 mg + MTX GOL 100 mg MTX	68% 45% 39%

[1]Strand and Khanna. Clin Exp Rheumatol 2010; 28 (Suppl 59): S32-40.

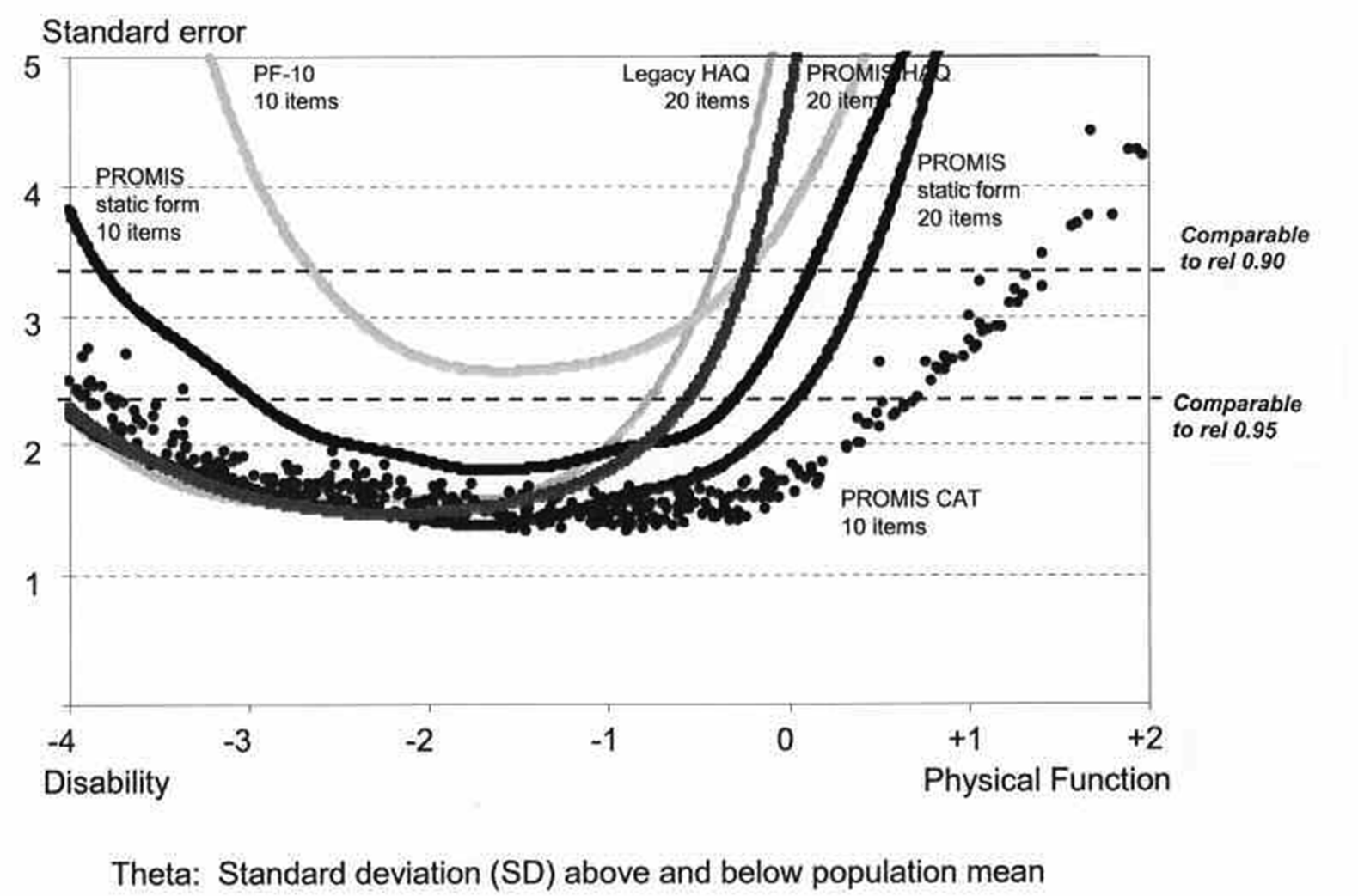
HAQ-DI: Feasibility
A PLUS- 5 minutes for completion;
Millions have completed it in clinics and trials

HAQ-DI: Summary

- Advantages:
  - Patient buy-in and history of success in trials
  - Valid
  - Sensitive to change; discriminates within and between groups
  - Feasible
  - Trialists and patients familiar
- Shortcomings:
  - Patient feedback after development
  - Ceiling effect in 10-15% patients
  - May not capture instrumental activities of daily living

Other Measures of Physical Function

PROMIS tools



Left to Right: PROMIS 10-item static form; the 10-item PF-10; the 20-item Legacy HAQ; the 20-item PROMIS HAQ; the 20-item PROMIS static form; and the PROMIS CAT 10-item simulation

Thank you for your attention

EXTRA SLIDES

HAQ: Face, content Validity[1, 2]

- Questions obtained from a variety of sources: Uniform database for rheumatic diseases, Convery patient status measure and ADL scales (Barthel index, Katz index)
- 100 questions developed
- Nurse assessor interviewed clinic patients to identify problems and modification for areas of imprecision and redundancy
- Numerous revisions (patients' feedback per Dr. Fries)
- Assessor and self-administered questionnaires tested with new samples of patients
- Study I: 20 volunteers at Stanford: interview then assessor; assessor then interview
- Study II: 28 surveyed patients were contacted for home evaluation and direct observation of fifteen tasks representing majority of the questions was performed by nurse-assessor

[1]Fries JF et al. Arthritis Rheum 1980;23 (2):137-145;

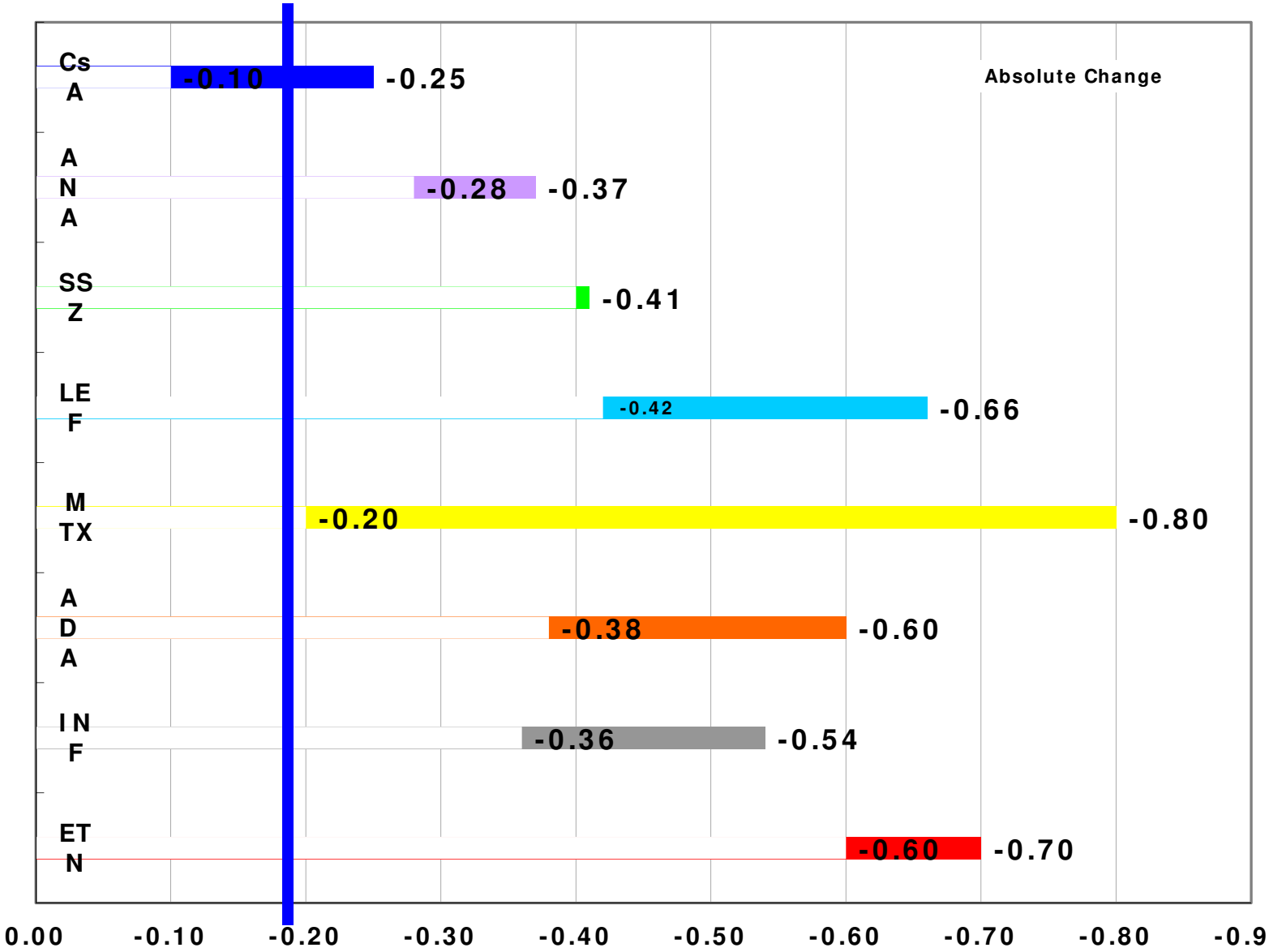
[2]Fries JF et al. J rheumatol 1982; 9(5):789-793

HAQ: Face, content Validity

- History of Patient Feedback (Jim Fries, personal communication):
  - "ARAMIS logged all patient comments received and would periodically review and incorporate them in essentially a modified Delphi approach. There were many thousands of comments over the years from over half a million administrations."
  - "Changes made included dropping the sexual function item which people didn't feel comfortable with, getting rid of the bathtub item, and changing wording on the push-button car door item as well as format and item stem changes. Thus, patient input was very seriously considered at all times but was indirect compared with current approaches."

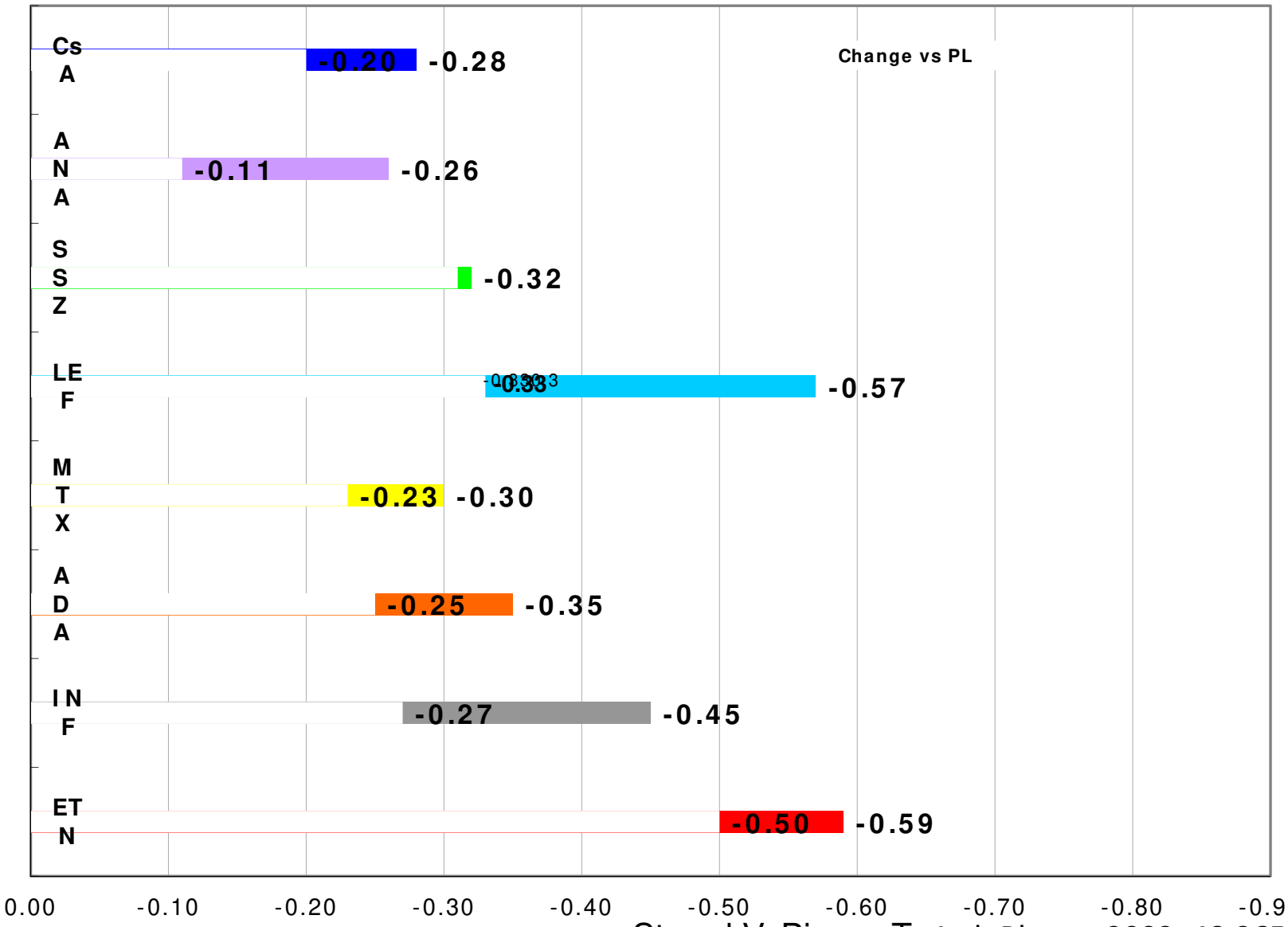
- Used as a part of ACR20/50/70 for >15 years
  - >10 new approvals for RA treatments
  - RCTs with Manageable sample sizes

Mean Absolute Changes in HAQ DI across RCTs



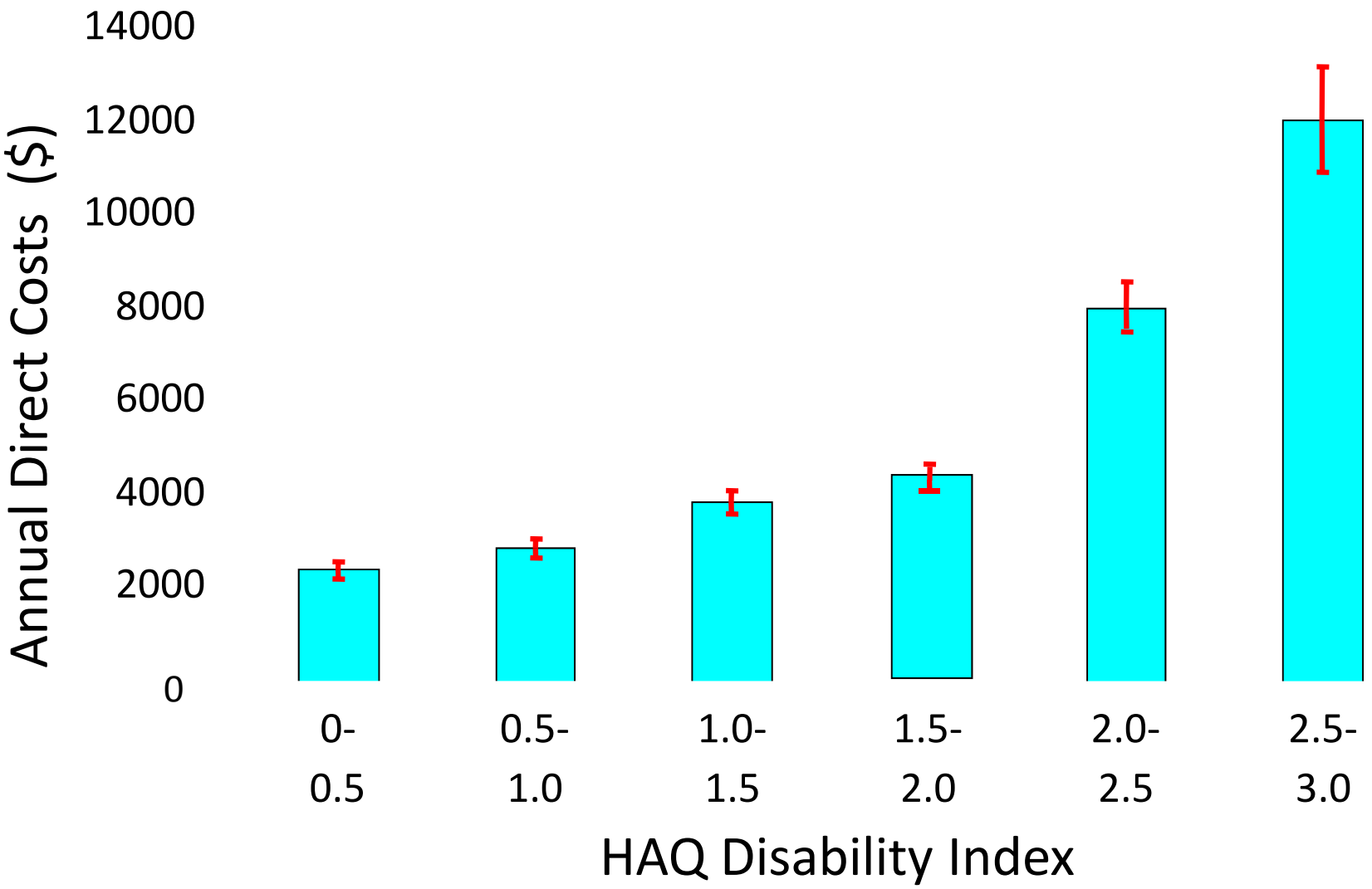
Strand V, Pincus T; Arth Rheum 2003; 48:3656 [LB 6].

Mean Changes in HAQ DI: Active - Placebo

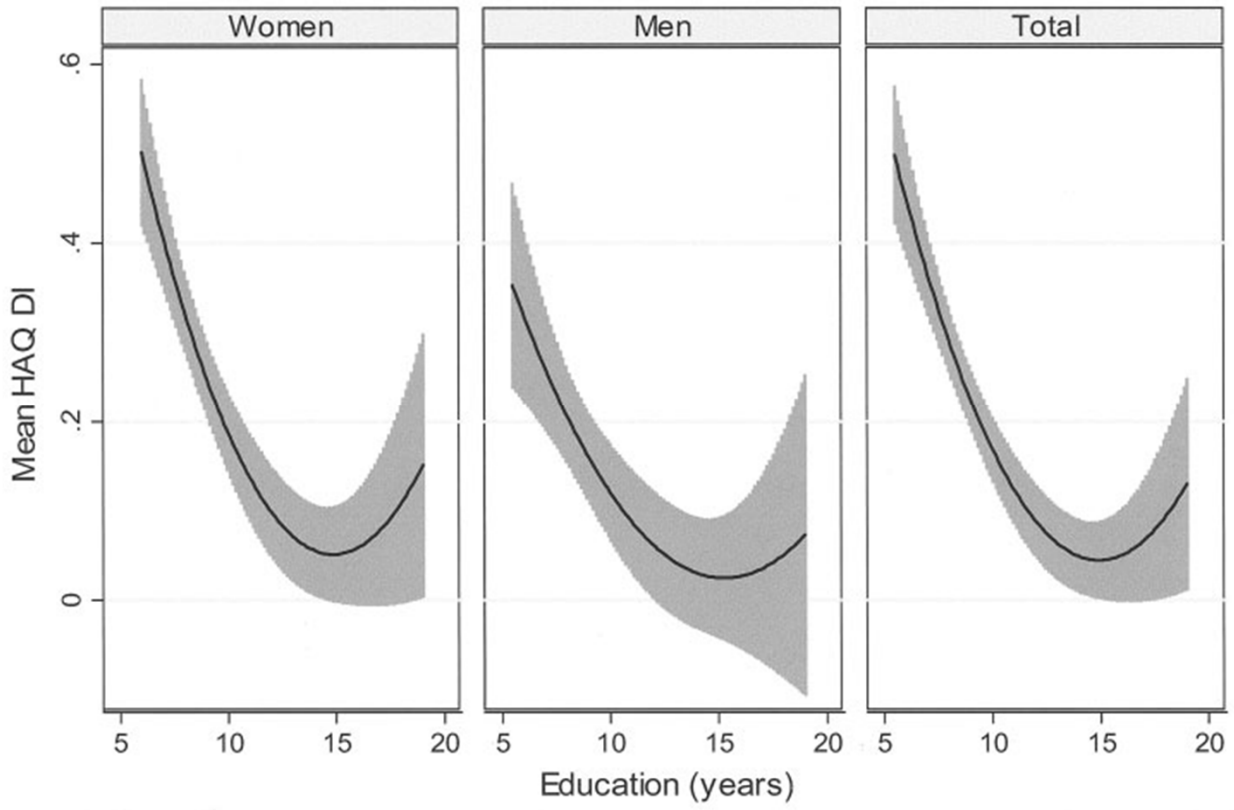


Strand V, Pincus T; Arth Rheum 2003; 48:3656 [LB 6].

Disability and Annual Costs



HAQ: Construct validity in a population-based sample[1]



[1]Krishnan E et al. Arthritis Rheum 2004;**50**(3): 953-960.

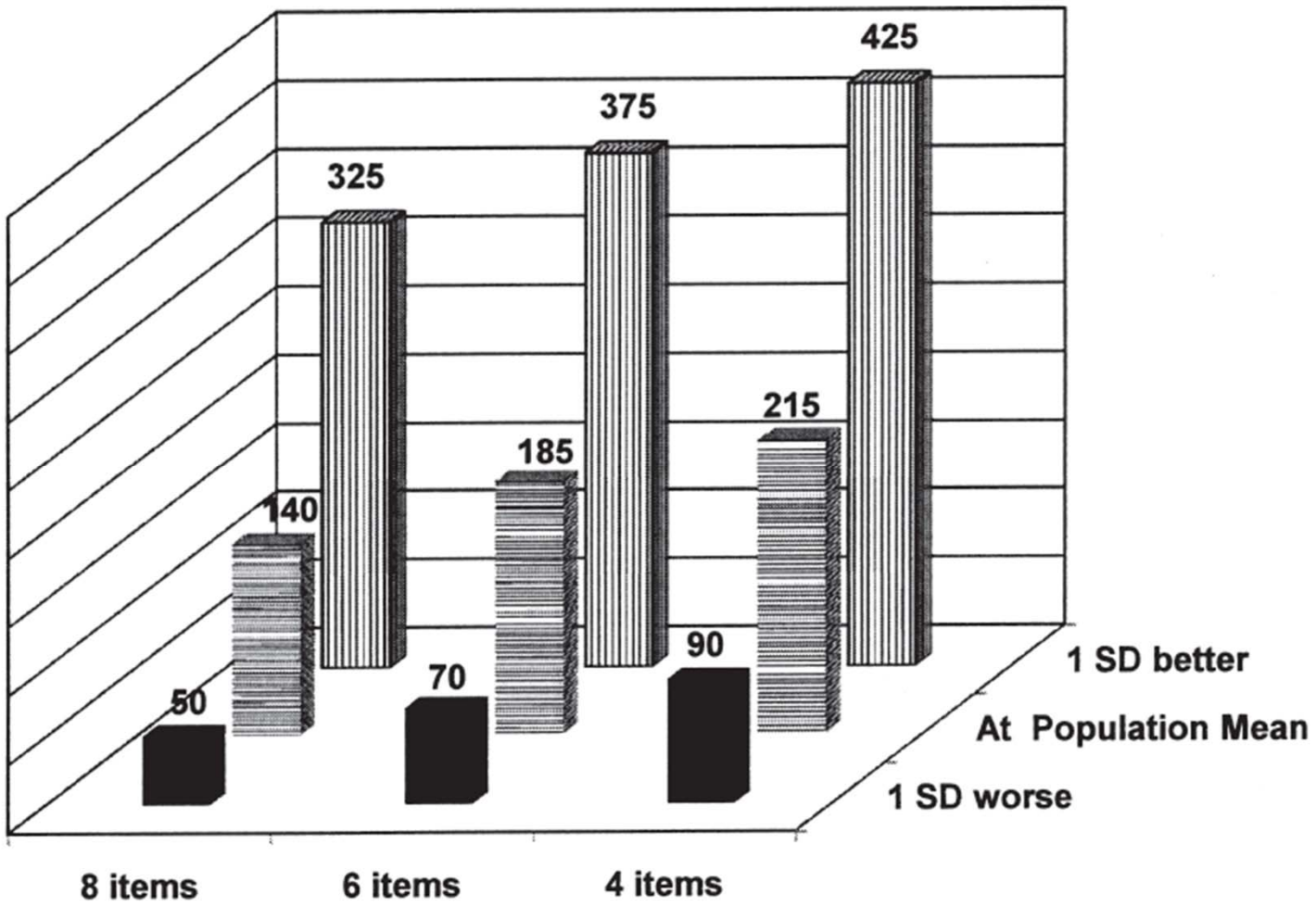
HAQ: Construct Validity, Comparison with WOMAC[1]

- Comparison of content
  - Structure and context different; stems differ
  - Wording, ordering of items differ
- Spearman's correlation between HAQ and
  - WOMAC PF: 0.71 (initial), 0.79 (follow-up)
  - WOMAC Stiffness: 0.61 (initial), 0.69 (follow-up)

	HAQ-DI	WOMAC PF
Total items	20	17
Items assessing both upper and lower extremity function	8	10
Only lower extremity function	3	5
Only upper extremity function	9	-

[1]Bruce B &? Fries J. Arthritis Care Res 2004; 51(5):730–737

Implications for sample size: New PROMIS Physical function questionnaires[1]



[1]Fries J et al. J Rheumatol 2011;38;1759-1764