Holger Schünemann, MD, PhD

Understanding and interpretation of patient reported outcomes

Cochrane Colloquium – Freiburg, October 3, 2008
Session Processes and Tools | Patients and Research

Disclosure

Relevant Financial Relationships

- Co-developer of CRQ-SAS. License fees from for-profit organizations to McMaster University supporting research projects

Off label medication

- None mentioned

Content

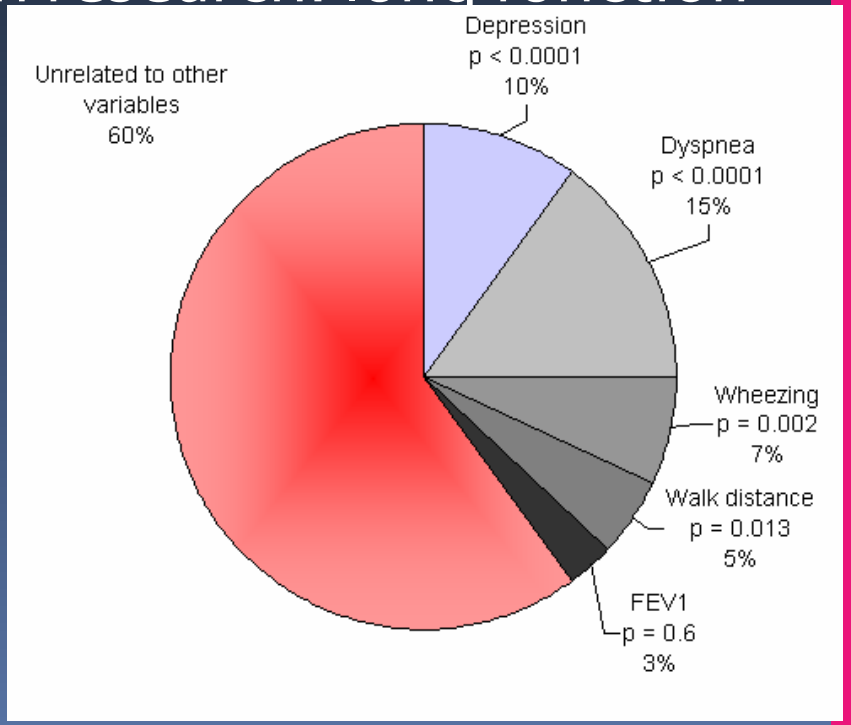
- Why “Patient Reported Outcomes (PROs)”?
- Types and principles of PROs?
- Interpreting results of PROs
  - Minimal important difference

Why PROs?

- Direct measurement of how people feel and function is replacing physiologic or laboratory tests as primary outcomes
- Primarily in chronic disease populations
- Shift is motivated by realization that changes in physiologic endpoints often bear a limited relation with changes in patient reported health

Example chronic obstructive lung disease (COPD)

- Fourth common cause of death worldwide
- No efficacious treatment to alter mortality
- Primary outcome measure in research: lung function
- Lung function correlates poorly with how patients feel
- How patients feel is a key patient-important outcome



Jones, ARRD 1992

Should we ask patients or clinicians?

Agreement between patients and clinician in assessment how patients feel

- COPD rehabilitation
- Feeling thermometer: 0 – 100 scale

Table 3 Baseline, follow-up and change scores of the patients' and physicians' estimation of health status on the FT.

	Baseline	Follow-up	Change scores
Patient FT	50.3 (14.0)*	63.0 (15.7)*	12.7 (range −14 to 40) (10.0; 15.3; $P<0.001$)‡
Physician FT	44.0 (14.4)*	66.2 (13.4)*	22.2 (range 0–55) (19.2; 25.1; $P<0.001$)‡
Difference between patients and physicians	6.3	−3.2	−9.5
	(1.2; 11.4 $P=0.015$)†	(−8.4; 2.1 $P=0.23$)†	(−5.6; −13.4; $P<0.001$)†

*Values are mean (standard deviation).
†Unpaired *t*-tests (95% confidence intervals; *P*-value).
‡Paired *t*-tests (95% confidence intervals; *P*-value).

Puhan et al. Respir Med 2004

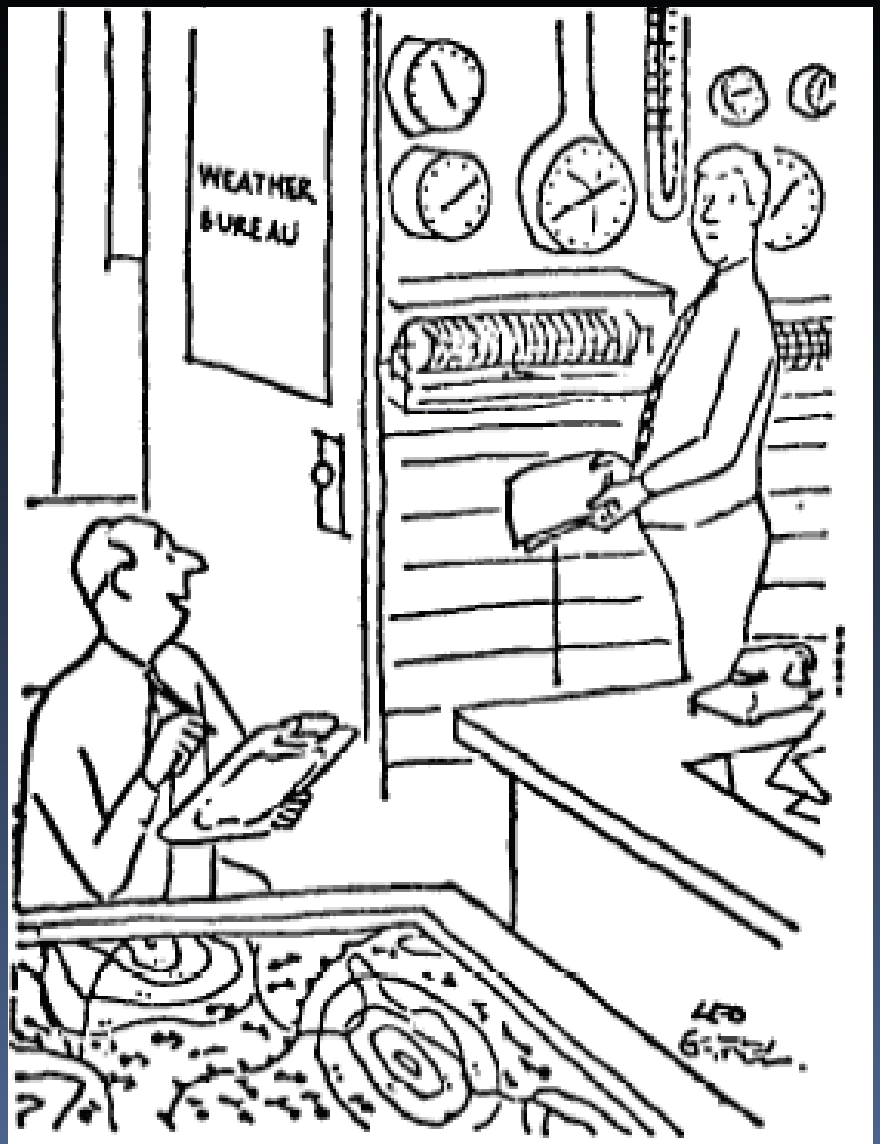
Validity of clinician assessment

Table 4 Correlation* between the FT change scores and the other outcome measures.

	Patient FT	Physician FT
Dyspnoea CRQ	0.43† (0.31–0.55)	0.04 (−0.09–0.17)
Fatigue CRQ	0.42† (0.30–0.54)	0.17 (0.04–0.30)
Emotional function CRQ	0.43† (0.31–0.55)	0.33‡ (0.21–0.45)
Mastery CRQ	0.51† (0.40–0.62)	0.16 (0.03–0.29)
SF-36 general health perception	0.30‡ (0.17–0.43)	0.04 (−0.12–0.14)
SF-36 vitality index	0.27‡ (0.14–0.40)	−0.05 (−0.19–0.09)
Borg dyspnoea scale	−0.31‡ (−0.19 to −0.43)	0.08 (−0.05–0.21)
6-min walking test	0.03 (−0.10–0.16)	0.36‡ (0.24–0.48)

*Pearson correlation coefficient (95% confidence interval).
†Significant at <0.01.
‡Significant at <0.05.

Puhan et al. Respir Med 2004



"I figure there's a 40% chance of rain, and a 10% chance we know what we are talking about."

Wall Street Journal

Why patients (PROs)?

Clinicians don't always have a good sense of how patients are feeling

Best use of PROs

When the goal of treatment is to improve how the patient is feeling, rather than to prolong life

- Or incidence of "hard outcomes"

Even then it is important to capture the variability in patient's function and feelings

- Mild vs severe stroke
- Large vs small infarct
- Painful vs painless death

Bryant et al Med Prakt 2007

Patient Reported Outcome Measures

Symptoms

Quality of life

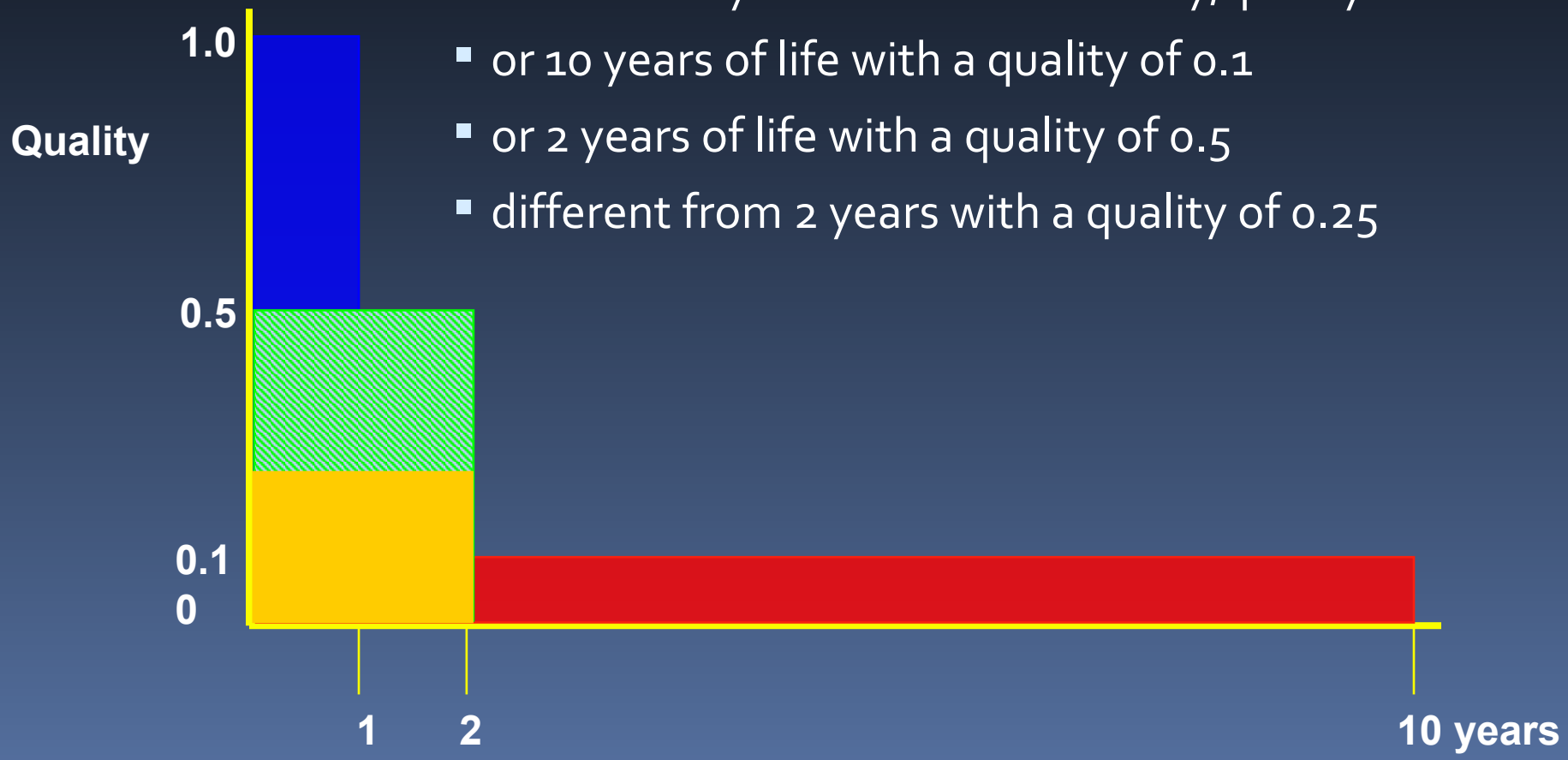
Patient Reported Outcome Measures

Symptoms

Quality of life

Quality adjusted life years

- 1 QALY = 1 year of life with a utility/quality of 1.0
- or 10 years of life with a quality of 0.1
- or 2 years of life with a quality of 0.5
- different from 2 years with a quality of 0.25



What is Quality of Life?

- Quality of life is an individual's satisfaction or happiness with life in domains she or he considers important
- *Health-related* quality of life (HRQL) is an individual's satisfaction or happiness with domains of life insofar as they affect or are affected by "health"

Health Related Quality of Life Measures

Generic Instruments:

Attempt to measure all important aspects of HRQL (e.g., health profiles and preference measures)

- SF-36

Specific Instruments:

Measure aspects of HRQL that are specific for the area of interest (e.g., a disease, a population or a problem, some preference measures)

Types of Instruments

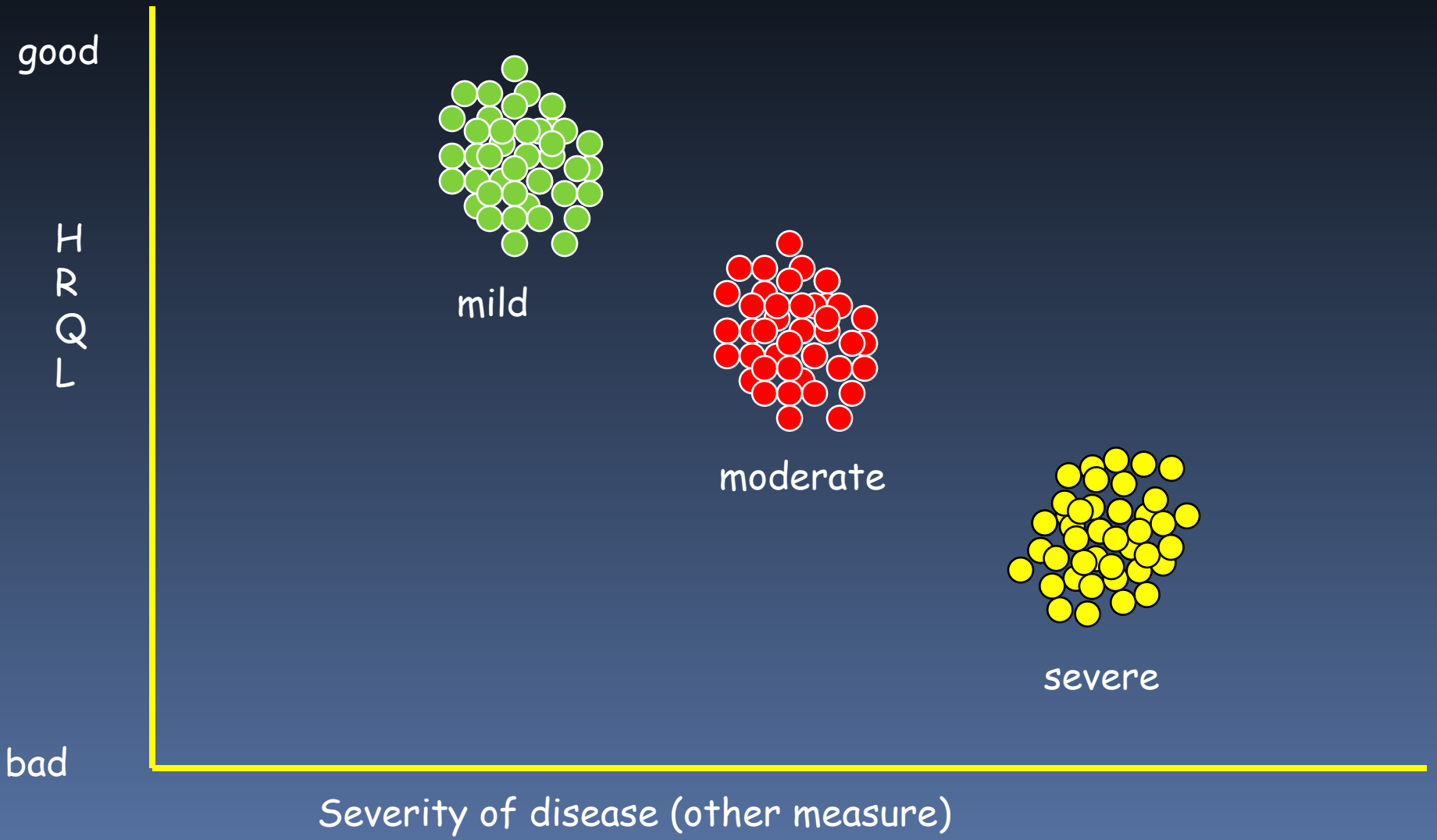
Discriminative Instruments:

- Differentiate between people who have better HRQL and those who have worse HRQL at one point in time

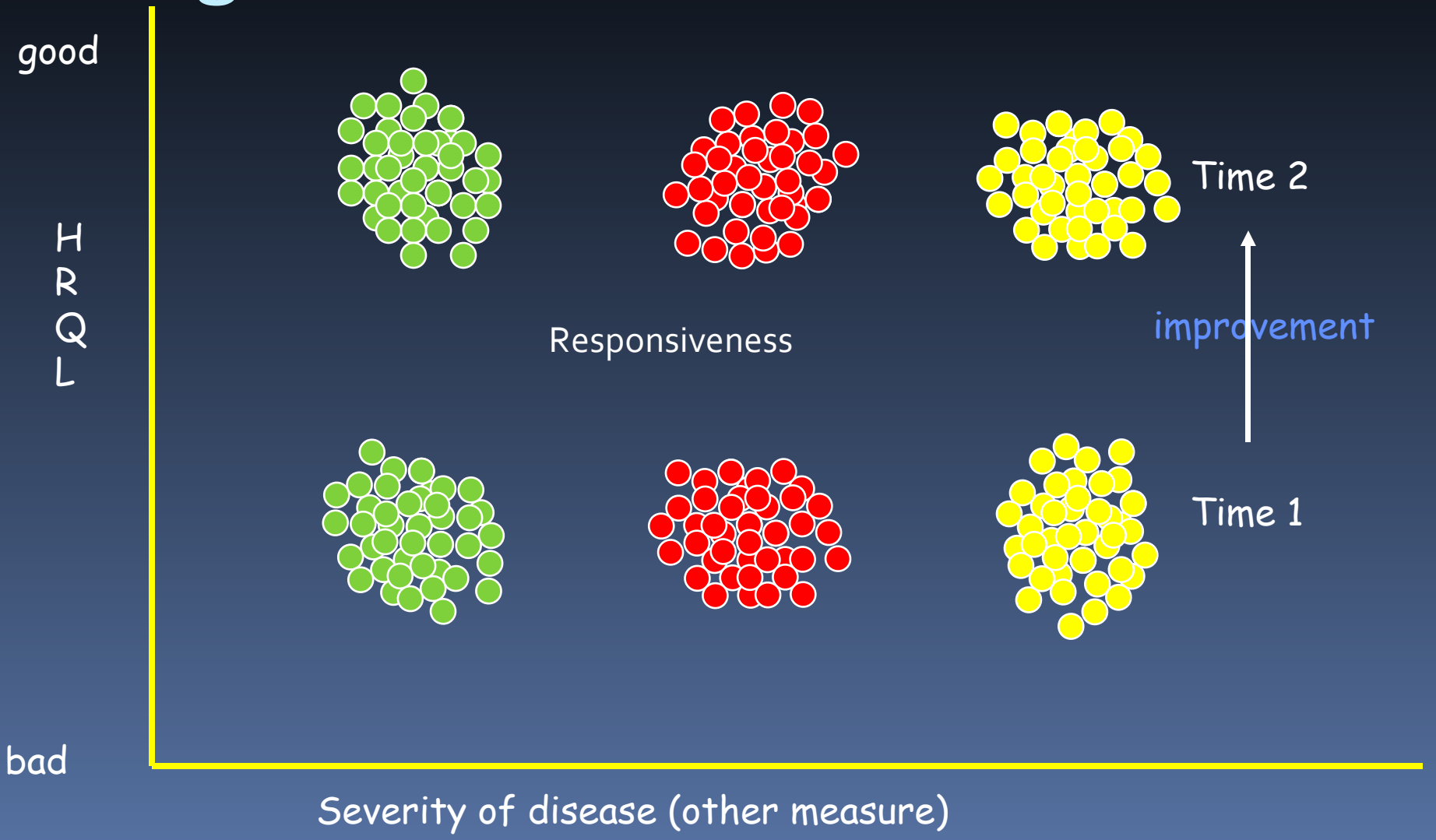
Evaluative Instruments:

- Focus on measuring how much HRQL has changed in repeated measurements

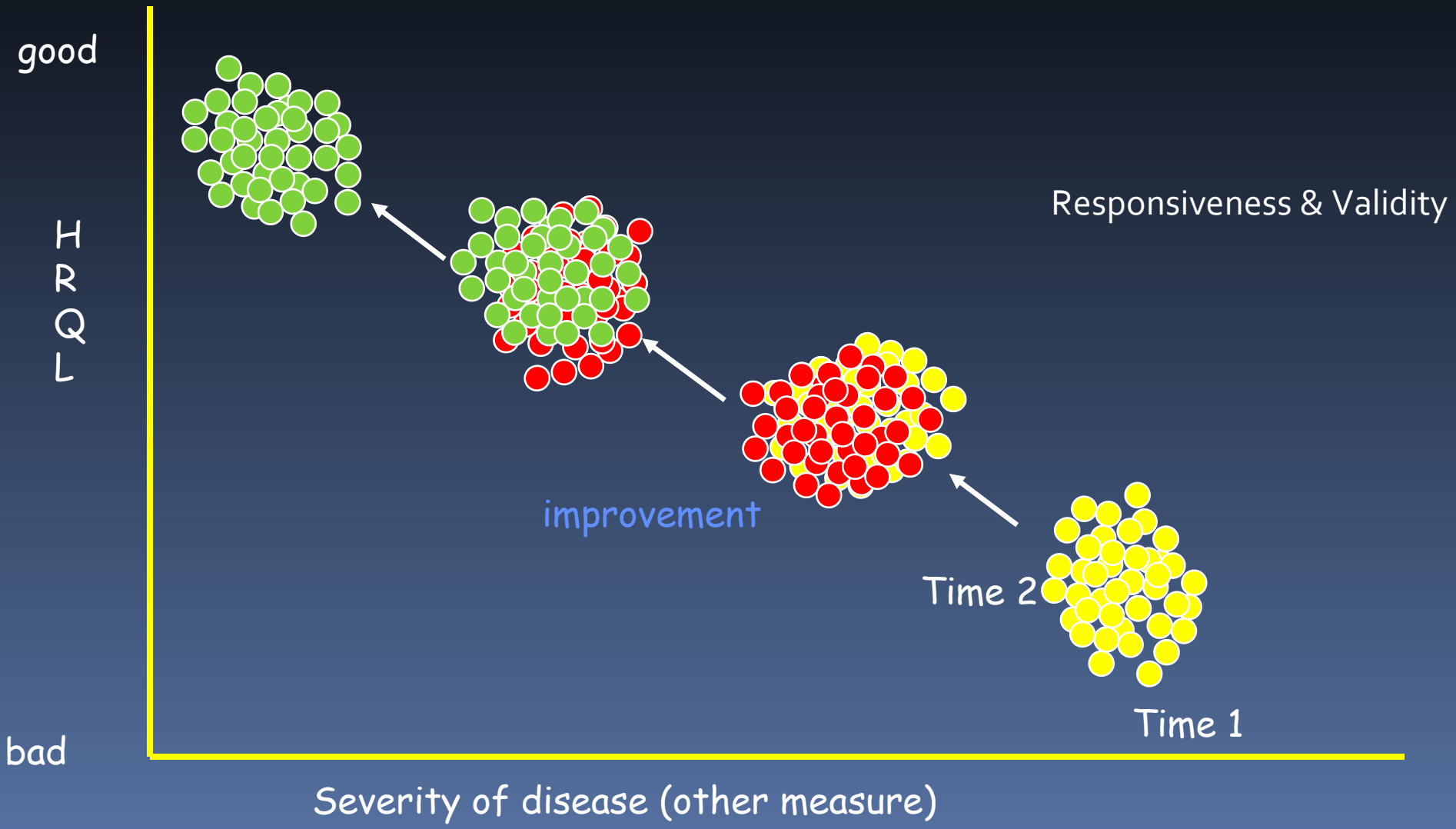
Discriminative properties
Cross-sectional measurement



Evaluative properties
Longitudinal measurement



Discriminative and evaluative properties: Longitudinal measurement

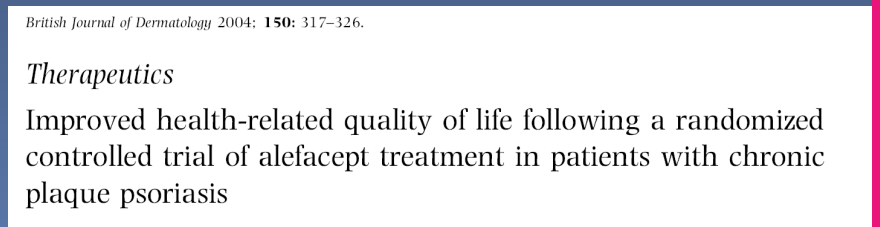


Content

- Why "Patient Reported Outcomes (PROs)"?
- Types and principles of PROs?
- Interpreting results of PROs
  - Minimal important difference

A (typical) presentation of QOL information

- Effect of alefacept on **quality** of **life** in 553 patients with psoriasis
- Alefacept significantly reduced (improved) mean Dermatology Life Quality Index compared with placebo:
  - 4.4 vs. 1.8 at 2 weeks (P<0.0001)
  - 3.4 vs. 1.4 at 12 weeks (P<0.001)



British Journal of Dermatology 2004; **150:** 317–326.

Therapeutics

Improved health-related quality of life following a randomized controlled trial of alefacept treatment in patients with chronic plaque psoriasis

Clinical research design

- We need to know how large a change is large enough?

Minimal Important Difference

- "the MID is the smallest difference in score in the outcome of interest that informed patients or informed proxies perceive as important, either beneficial or harmful, and which would lead the patient or clinician to consider a change in the management"

EFFECT MEASURE	EXPLANATORY	PRAGMATIC
	Therapy efficacious - Pragmatic trial needed!	Therapy effective - Implement!
	Therapy not efficacious – Abandon! Or non-inferior – Need pragmatic trial!	Therapy not effective - Consider explanatory trial! Or non-inferior – Use if other advantages!
	Indeterminate – Further research warranted!	Indeterminate – Further research warranted!

Karanicolas et al, JCE , 2007

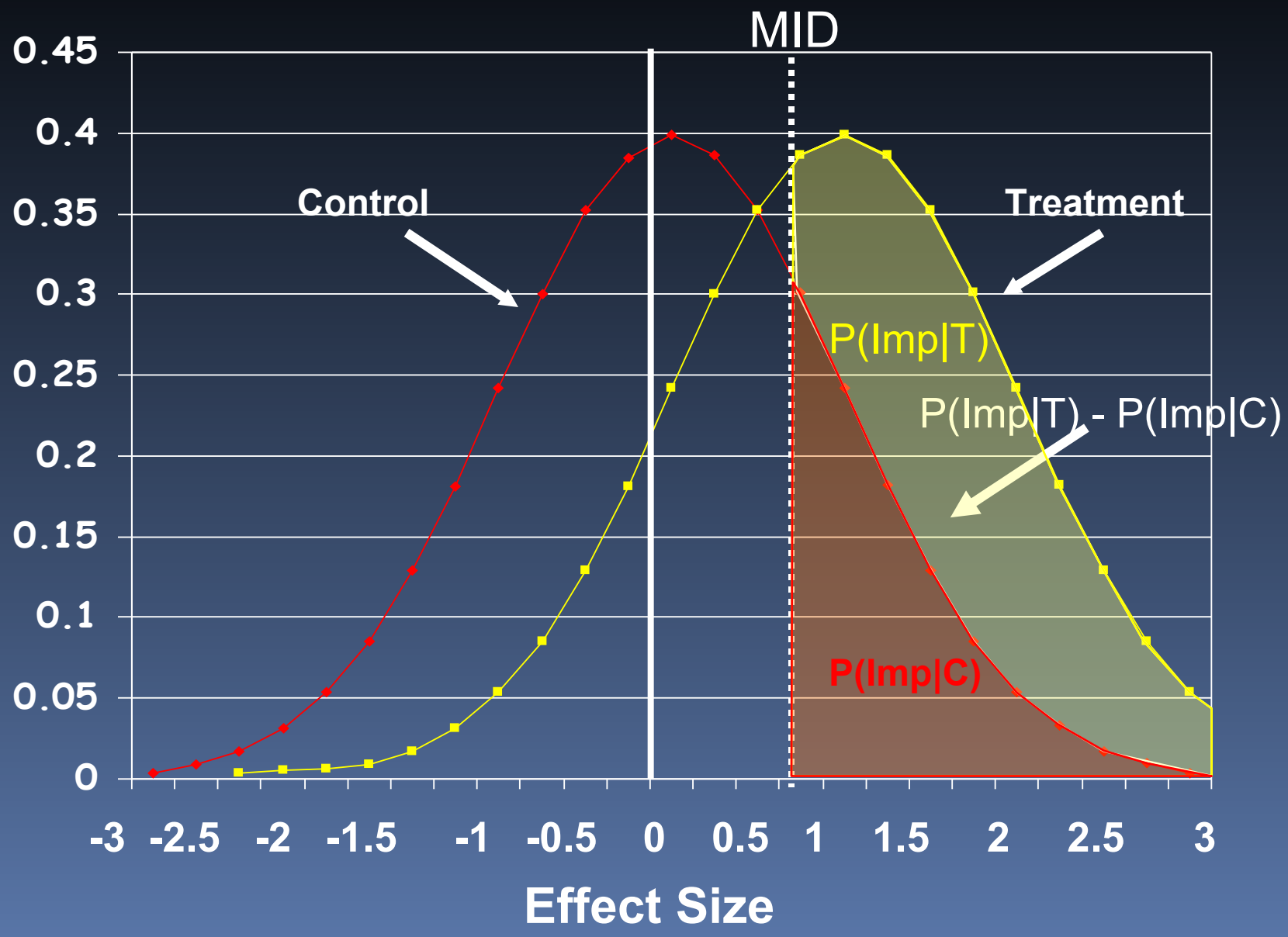
Minimally important difference

- Methods
  - scenarios
    - patients, clinicians
  - global ratings of change
- MID (CRQ/CHQ/AQLQ = 0.5)
- Mean change 0.6

  $\Rightarrow$ everyone benefits?
- Mean change 0.3

  $\Rightarrow$ no one benefits?

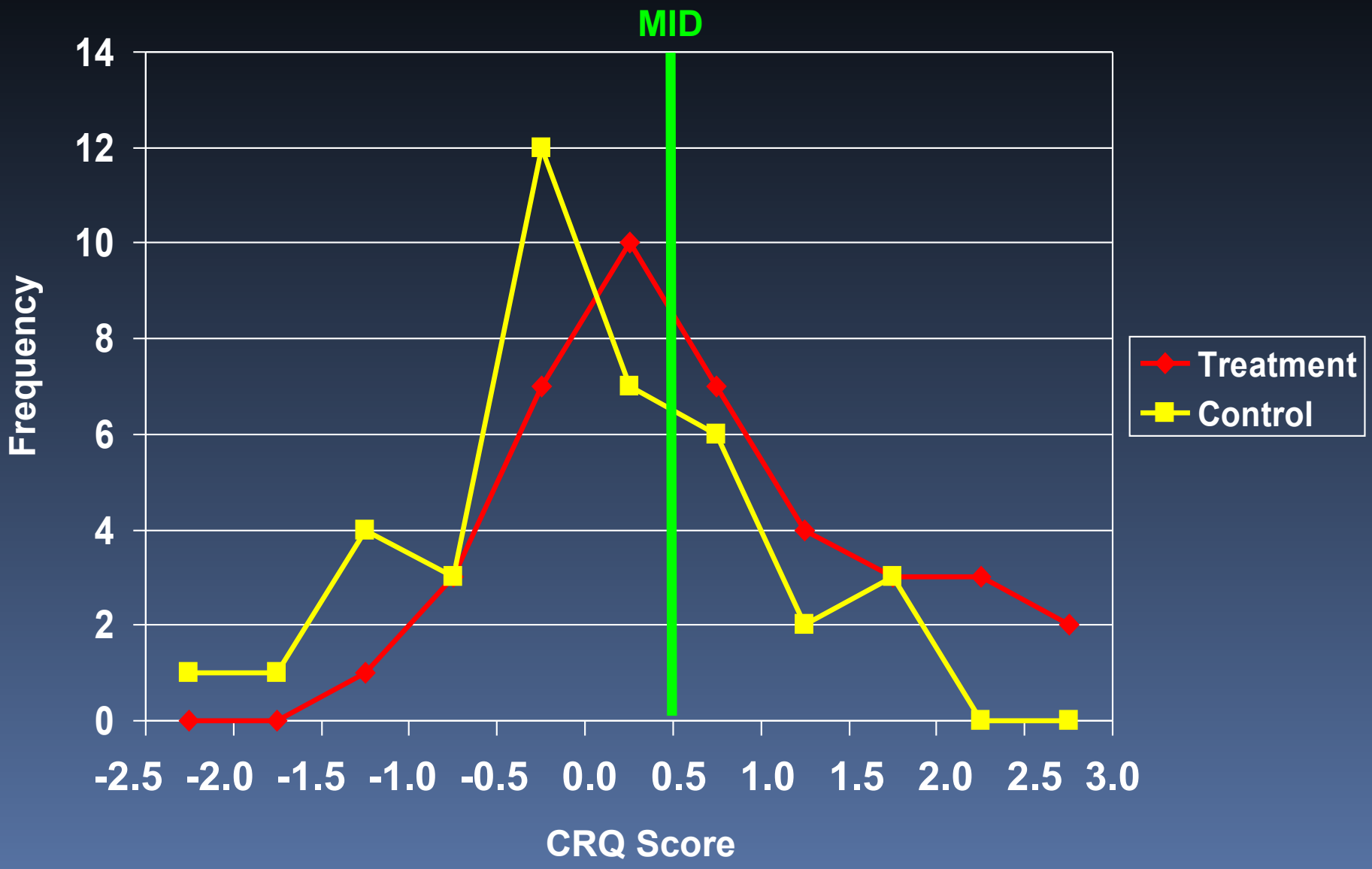
Change score difference



Effect Size and NNT

Effect Size	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
NNT	20	13	9	7	5	5	4	4

CRQ Dyspnea Change Scores



CRQ Dyspnea Change Scores

Difference between groups: mean=0.60, 95% CI=(0.18, 1.03)

Cut-point	Proportion better on rehab	Proportion better on conventional care	Proportion benefiting from rehab	NNT for a single patient to benefit
0.0	0.72	0.46	0.26	3.8
0.5	0.47	0.28	0.19	5.2
1.0	0.30	0.13	0.17	5.8
1.5	0.20	0.08	0.12	8.1
2.0	0.12	0.00	0.12	8.0
2.5	0.05	0.00	0.05	20.0

Steps to making QOL data from RCT compelling to clinicians

- Determine the MID
- Report mean differences
  - in terms of MID
- Choose threshold
  - absolute
  - change related to MID
- Calculate proportion benefit, NNT, natural frequencies

Conclusion

- Clinicians not always good surrogates for patients
- Science of PRO is well advanced with standard methods to assess instruments' measurement properties
- However, guidance for interpretability is key
  - MID as useful concept

Thank you