

Responsiveness

Henrik Hein Lauridsen
Associate Professor
Research unit for Clinical Biomechanics
University of Southern Denmark

1

Definition

Responsiveness

"The ability of an HR-PRO instrument to detect change over time in the construct to be measured"

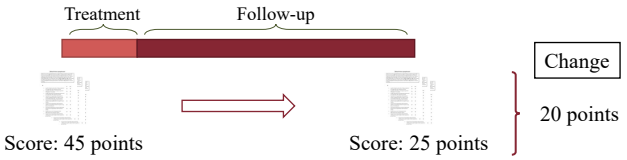
COSMIN, 2011

"Målbarhed af en reel ændring"

Bendix T., 2008

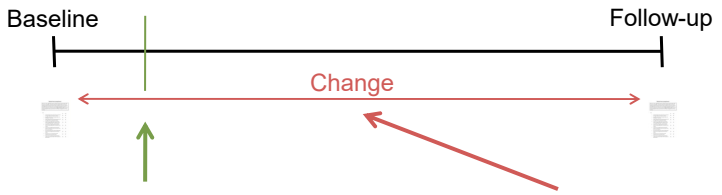
2

Measuring change



Is the measured change valid?

Responsivness versus validity



Validity
"Cross-sectional" validity
Refers to validity of a single score

Responsivness
"Longitudinal" validity
Refers to validity of a score change

Consequently, validity and responsiveness have similar standards

Hays & Hardon, Qual Life Res 1992

Distribution- versus anchor-based approaches

Distribution-based methods

- Express observed change to some form of variation in the obtained sample

$$Responsiveness = \frac{Observed\ change}{Variation}$$

Anchor-based methods

- Relationship between a questionnaire and an independent measure (or anchor)
- Elucidate the meaning of a particular degree of change

Distribution-based approaches

Effect size

- Individual change in SD of pretest

$$ES = \frac{Score_{pretest} - Score_{posttest}}{SD_{pretest}}$$

Standardized response mean

- Individual change in SD of change

$$SRM = \frac{Score_{pretest} - Score_{posttest}}{SD_{change}}$$

Guyatt's responsiveness statistics

- Accounts for spurious changes arising from measurement error

$$GRS = \frac{Minimal\ Important\ Change}{SD_{change,stable}}$$

$$GRS = \frac{(Score_{pretest} - Score_{posttest})_{Improved\ pts.}}{SD_{change,stable}}$$

Crosby et al. Defining clinically meaningful change in health-related quality of life. 2002

Problems with ES, SRM & GRS

Is not a reflection of the 'true change'

- Possible reasons:
  - Instrument could have a ceiling effect
  - Instrument lack relevant items

Highly dependent on variability (SD)

- High coefficients with small SD

Are measures of the ***magnitude*** and not the ***validity*** of change scores

Does not provide a good indication of the ***importance*** of the observed change

7

Two types of responsiveness

Construct responsiveness

The degree to which the change scores of a HR-PRO instrument are consistent with hypotheses

- Tests hypotheses about expected change

Criterion responsiveness

The degree to which the change scores of a HR-PRO instrument are an adequate reflection of a 'gold standard'

- Correlations
- Sensitivity/specificity
- Receiver Operating Characteristic (ROC) curves

Mokkink et al. (2010); Terwee et al. (2009)

8

Construct responsiveness

Specify the hypotheses before collecting data

Eg.:

- Expected correlations between score changes and changes in other variables
  - Example: function and satisfaction with treatment
- Expected differences in score changes between groups
  - Example: severe depression versus mild depression
- Expected score changes after a treatment with known effect
  - Example: treatment of tension headache with NSAID

Terwee et al. (2009)

Construct responsiveness

Hypothesis

- Should be specific
  - Direction of the expected correlation/difference
  - Size of the expected correlation/difference
- Should be challenging
- The more hypotheses the better

Terwee et al. (2009)

Example: Low Vision Quality of Life" questionnaire (LVQoL)

4 subscales

- Basic aspects
- Mobility
- Adjustment
- Reading

Compared to

- VF-14
- Acuity
- Global change →

- ❑ Much better
- ❑ Better
- ❑ A little better
- ❑ No change
- ❑ A little worse
- ❑ Worse
- ❑ Much worse

De Boer et al. Qual Life Res 2006

Example cont'd

Hypotheses example

Correlation between score changes on LVQoL and global change is least 0.10 higher compared to the correlation between the score changes on LVQoL and acuity

Results

LVQoL subscale	Global change		Change in acuity
Δ basic aspects	0.26	>	0.13
Δ mobility	0.26	>	0.08
Δ reading	0.18	>	0.07
Δ adjustment	0.20	>	0.04

4 out of 4 (100%) hypotheses confirmed

De Boer et al. Qual Life Res 2006

Criterion responsiveness

Requires a 'gold standard' measuring change in the domain

Eg.

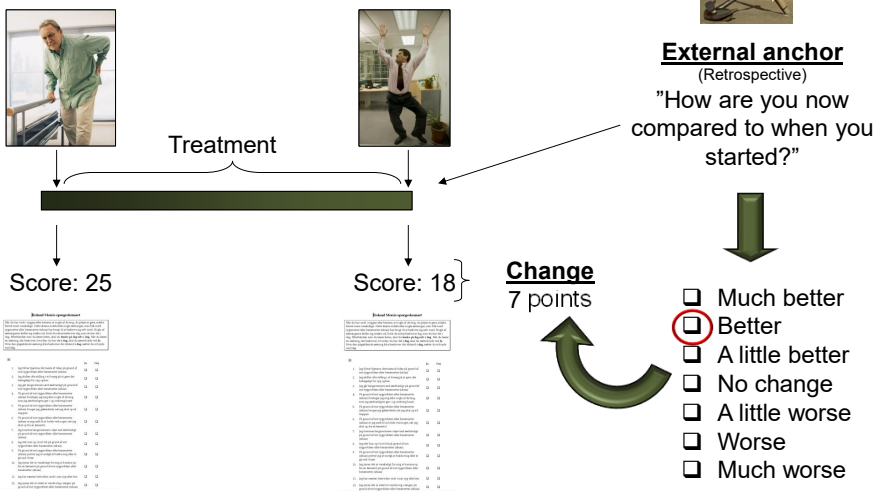
- Global change
- Change in a clinical variable

Methods

- Correlation between a 'gold standard' and score change
- Sensitivity/specificity – dichotomous variable
- Receiver Operating Characteristic (ROC) curves

13

ROC analysis



14

The external anchor (transition question)

”How are you now compared to when you started?”

Recall bias	Present state bias	Motivational bias	Contami-nation bias	Pleasing bias
Problems remembering prior healths status	Relate to how things are now and *not* to the change	Cumbersome treatment → Overestimation of the treatment response	Comorbidity contaminates the answer Eg. LBP patient with headache	Patients wants to please the clinician

Norman et al., 1997; Hägg et al., 2002; Middle et al., 2006; Aseltine et al. 1995; Herrmann et al., 1995, Lauridsen et al. 2007

SDU

15

Bias depends on time frame

Follow-up < approx. 6 wks.

- Retrospective anchor ≈ OK

Follow-up > approx. 6 wks.

- Retrospective anchor biases

Solution: ‘The Punum Ladder’

SDU

16

The Punum Ladder

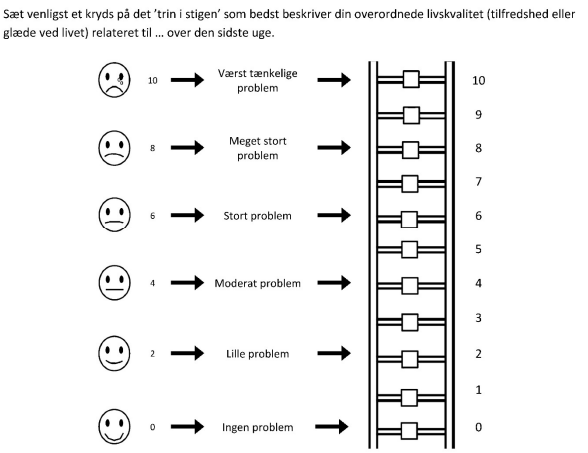
Prospective anchor which avoids biases

Score range: [0-10]

Applied to each dimension

Well validated:

- Validity
- Reproducibility
  - Reliability
  - Internal consistency
- Responsiveness

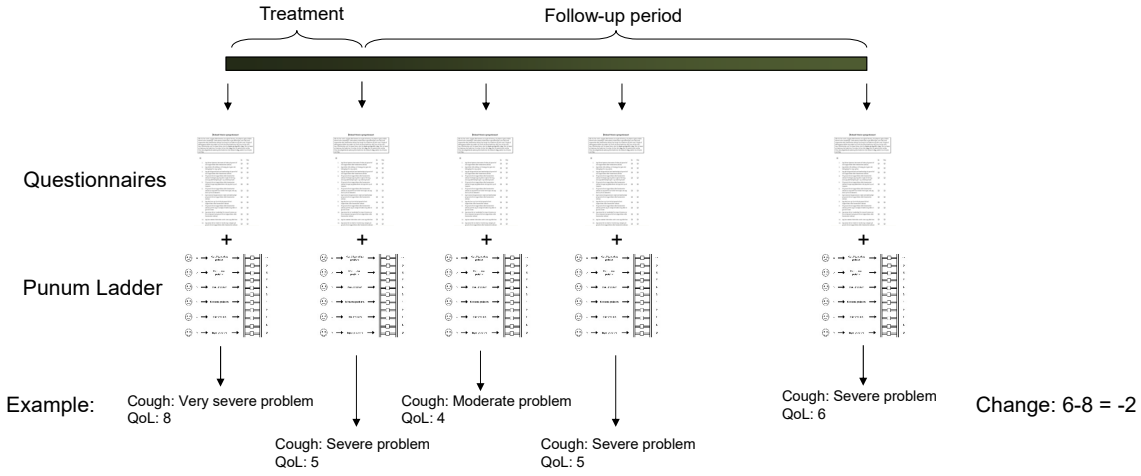


NB: Danish version has not been tested for cross-cultural validity

Fletcher et al, 2010

17

Study design with Punum Ladder



Fletcher et al, 2010

18

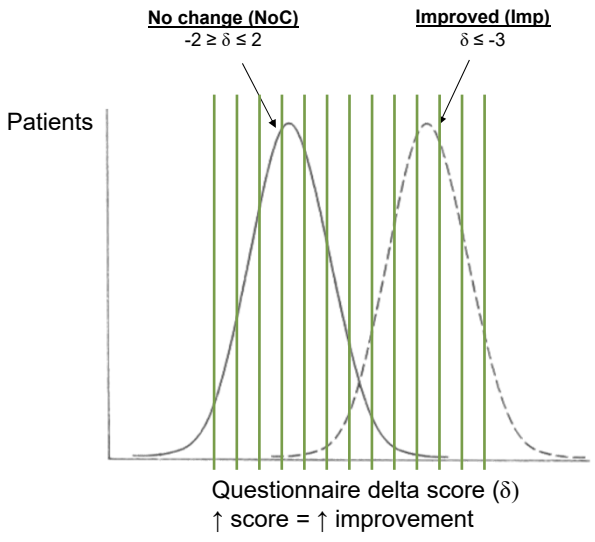
Trichotomisation

Change score ratings	Interpretation
-2, -1, 0, +1, +2	No change
≤ -3	Improvement
≥ +3	Deterioration

Fletcher et al, 2010

19

Sensitivity and specificity



		Anchor	
		Imp	NoC
Score	≥ δ	A	B
	< δ	C	D

Sensitivity: A/(A+C)

True positive rate

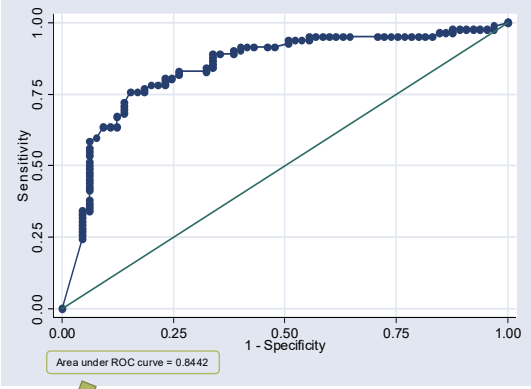
Specificity: D/(B+D)

True negative rate

20

The ROC curve

Plots the true positive rate (sensitivity) against the false positive rate (1-specificitet)



Area under the curve = proportion of correctly identified patients, who truly have improved = responsiveness

ROC_{auc} = 0.50 → unresponsive

Questions?

