Coloplast

Statistical validation of scales for measuring health related quality of life

Henrik Andersen, Helle Rootzén and Søren Lophaven

Agenda

- Introduction
- Theory and methodology
- Application

Why measure quality of life at Coloplast

Throughout the world we wish, within our selected business areas, to be the preferred source of medical devices and associated services, contributing to a better quality of life

Introduction

- Health related quality of life cannot be measured directly
- Instead patients are asked how they feel with respect to ...
- A number of standardised instruments (scales) exist (generic ones as EQ-5D or SF-36, disease-specific as ostomy adjustment scale, StomaQOL)
- These scales consist of a number of questions (items). Based on the answers to these items the score is calculated.

EuroQol EQ-5D

- Mobility (MO) (I have no problems in walking about, I have some problems in walking about, I am confined to bed)
- Self-care (SC) (I have no problems with self-care, I have some problems washing and dressing myself, I am unable to was or dress myself)
- Usual activities (UA) (I have no problems with performing my usual activities, I have some problems with performing my usual activities, I am unable to perform my usual activities)
- Pain/discomfort (PD) (I have no pain or discomfort, I have moderate pain or discomfort, I have extreme pain or discomfort)
- Anxiety/Depression (AD) (I am not anxious or depressed, I am moderately anxious or depressed, I am extremely anxious or depressed)

Objective

Perform scale validation of EuroQol EQ-5D based on clinical data from patients with chronic wounds.

The goal is to obtain a valid and realiable scale for quality of life, which could be used for example for comparing treatment groups in clinical trials.

Coloplast

Theory and methodology

Validity

Does the scale measure what it intend to measure

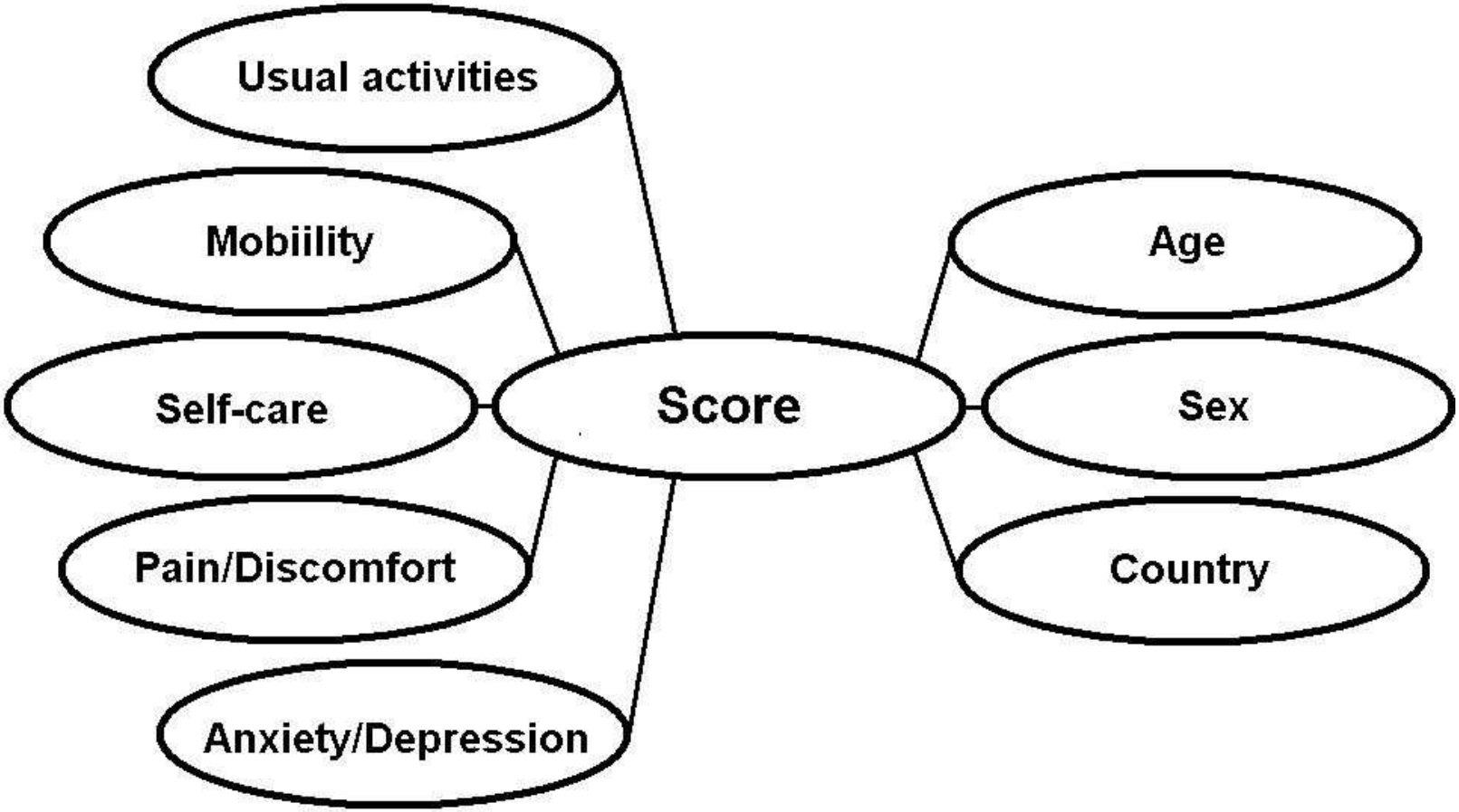
- Content validity (item coverage, the scale should include items relating to all relevant aspects of the latent variable)
- Criterion validity (the score must correlate with all variables known in advance to be correlated to the latent variable)
- Construct validity (item responses must not depend on anything but the latent variable)

Construct validity

- Items: $Y = (Y_1, ..., Y_k)$
- The total score: $S = \sum_i Y_i$
- The latent variable: Θ
- Covariates: $X = (X_1, ..., X_m)$

- Unidimensionality: Separation of items into several item bundles. Difficult to distinguish between multidimensionality and local dependence.
- Local independence ($Y_i \perp Y_j | \Theta$)
- No item bias/Differential Item Functioning (DIF) ($Y_i \perp X_j | \Theta$)

Construct validity



Consequences of construct validity

- The items must be positively correlated
- Items must be positively correlated with rest scores
- If the score correlates with a covariate, X, then all items must correlate with X in the same way

Coloplast

Technical requirements

- `Reliability:` The correlation between test and retest results performed for the same person in such a way that the test and retest - results are conditionally independent given the latent variable, (Test $\perp$ Retest | Θ) (Retesting not possible in practice, Chronbachs α gives the lower bound of reliability)
- `Sufficiency:` The score is a sufficient statistic for the person parameter in the conditional distribution of items given the latent variable.
- Ability to discriminate, simplicity

The Rasch model

The Rasch model is the only model meeting the technical requirements as well as requirements regarding validity.

The problem that the test of the Rasch models is supposed to solve is not a problem a defining a model that describes the variation of items in a given sample.

The problem is a question of the quality of the scale. That is, an evaluation of the degree to which it makes sense to attempt to measure anything at all (validity) and the problem of some technical properties of the scale.

The classical Rasch model

$$P(Y_i = 1 | \Theta = \theta) = \frac{\exp(\alpha_i + \theta)}{1 + \exp(\alpha_i + \theta)}$$

The graphical log-linear Rasch model

$$\ln(P(Y=\mathbf{y}|\theta_i,X)) = \alpha_0 + \sum_j (\alpha_{jy_j} + \beta_i y_j) + \lambda^{r,t}_{y_r y_t} + \kappa^{u,w}_{y_u x_w},$$

where $\alpha_{jy} = \ln(\gamma_{jy})$ and $\beta_i = \ln\theta_i$. λ determines local dependence between items y_r and y_t. κ is the item bias/DIF between the covariate x_w and item y_u.

Uniform DIF and local dependence: The association between items and association between items and covariates does not depend on the latent parameter.

Coloplast

- Agenda
- Why measure quality of life at Coloplast
- Introduction
- EuroQol EQ-5D
- Objective

Theory and methodology

Application

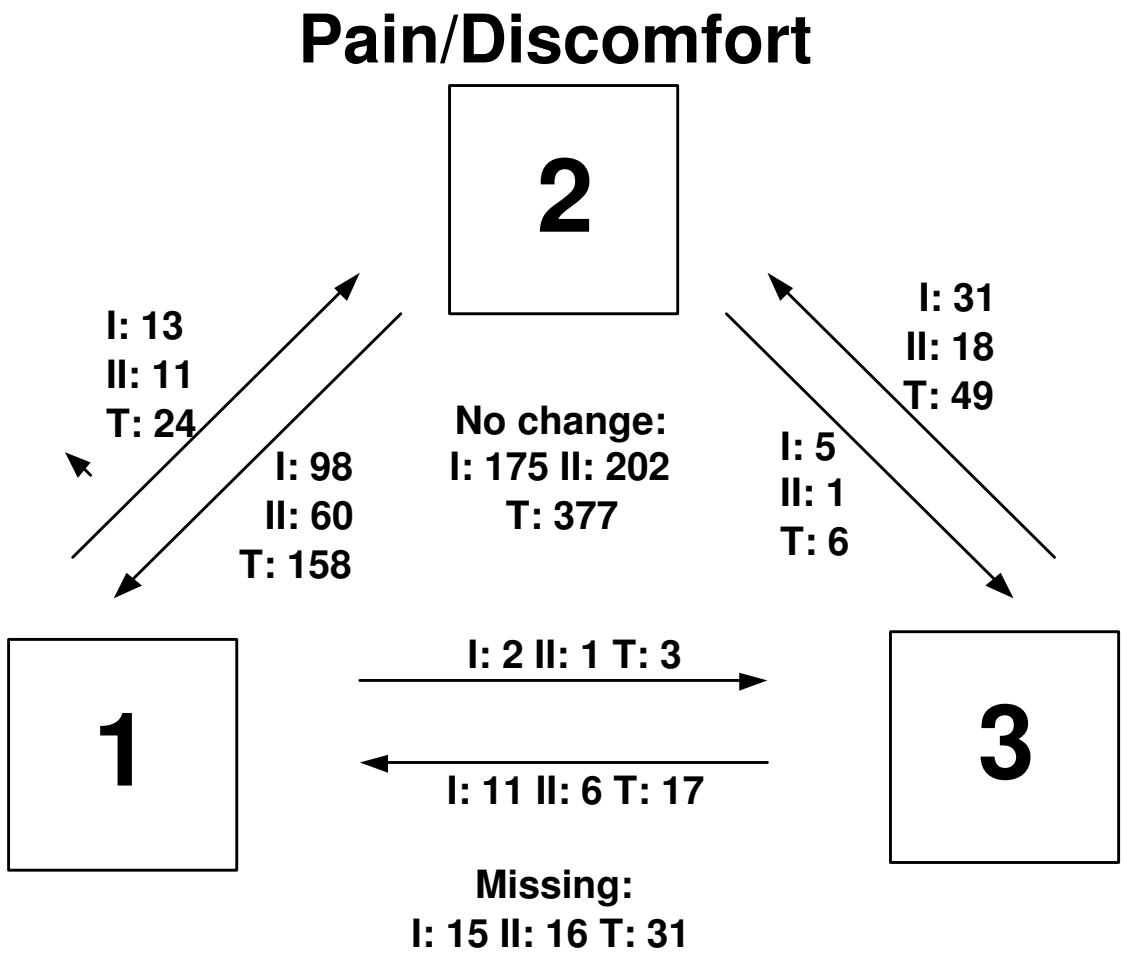
- Dataset
- Pain/discomfort before and after treatment
- Dimensionality/local dependence
- Results for one of the dimensions
- Differential item functioning
- The log-linear Rasch models
- DIF corrected scores
- Comparisons
- Some conclusions

Application

Dataset

- Post marketing study comparing a silver dressing with local best practice
- Two parallel groups, 650 patients, 10 countries
- The treatment period is 4 weeks, and health related quality of life is assessed before and after treatment using EuroQol EQ-5D
- Covariates: Treatment group, age, sex, wound type, region, dummy variable (identifying before and after treatment)

Pain/discomfort before and after treatment



Dimensionality/local dependence

Partial γ coefficients with Generalised Tjur conditions - Row item was deleted from the score:

	MO	SC	UA	PD	AD
MO		0.71	0.59	-0.28	-0.65
SC	0.64		0.67	-0.63	-0.32
UA	0.34	0.58		- 0.49	-0.27
PD	0.21	-0.27	0.19		-0.05
AD	-0.36	0.09	0.45	-0.13	

Physical imension: Mobility, Self Care, Usual Activities

Mental dimension: Pain/Discomfort, Anxiety/Depression

Results for one of the dimensions

		Mobility	Self-care	Usual activities
Mobility	γ		0.184	-0.184
	p		0.357	0.356
Self-care	γ	-0.110		0.110
	p	0.475		0.463
Usual activities	γ	-0.301	0.301	
	p	0.057	0.068	

Differential item functioning

		Mobility	Self-care	Usual activities
Group	γ	0.149	0.193	-0.295
	p	0.189	0.152	0.010
Age	χ^2	28.7	41.3	28.4
	p	0.017	0.000	0.016
Sex	γ	-0.089	0.178	-0.046
	p	0.423	0.202	0.694
Type	χ^2	42.9	28.3	45.0
	p	0.083	0.350	0.022
Region	γ	-0.697	-0.066	0.633
	p	0.001	0.759	0.000
Index	γ	-0.140	0.144	0.026
	p	0.210	0.294	0.824

Benjamini & Hochberg rejects at 0.00125 to control of α=0.01

The log-linear Rasch models

Physical dimension: Mobility, Self Care, Usual Activities and Region

$$\begin{aligned}\ln(P(Y=\mathbf{y}|\theta_i,X)) &= \alpha_0+\sum_j(\alpha_{jy_j}+\beta_i y_j)+\\ &= \kappa^{SC,Region}_{y_{SC}x_{Region}}+\kappa^{UA,Region}_{y_{UA}x_{Region}}+\\ &= \kappa^{MO,Age}_{y_{MO}x_{Age}}+\kappa^{UA,Age}_{y_{UA}x_{Age}}\end{aligned}$$

Mental dimension: Pain/Discomfort, Anxiety/Depression and Region, Type

$$\ln(P(Y=\mathbf{y}|\theta_i,X))=\alpha_0+\sum_j(\alpha_{jy_j}+\beta_i y_j)+\kappa^{PD,Region}_{y_{PD}x_{Region}},$$

The mental dimension has low reliability

DIF corrected scores

Region/age group (11∼Europe/-59)

Score	11	21	12	22	13	23
1	1.00	0.57	1.17	0.81	1.18	0.81
2	2.00	1.13	2.16	1.49	2.15	1.48
3	3.00	1.73	3.50	2.08	3.51	2.06
4	4.00	2.49	4.82	2.76	4.87	2.75
5	5.00	3.61	5.47	4.23	5.51	4.35

Comparisons

		μ	P	μ	P
Group	1	0.444	0.046/0.019	1.032	0.025/0.006
	2	0.152		0.711	

Some conclusions

- Not a unidimensional scale
- Items are separated into two dimensions - a physical and a mental scale. Data analysis should be done for both scales
- Only one of the dimensions had a satisfactory reliability
- The score is corrected according to the patient population
- It is the person parameters based on a valid scale which should be analysed and used for comparison