

## Guidelines to calculation by the free software and interpretation of the measures of disagreement applied to studies of change Elisabeth Svensson, Anders Avdic.

Assessments on rating scales generate ordered categorical data, also called *ordinal data*. This type of data represents a rank order only and not a numerical value in a mathematical sense, even when the assessments are numerically labelled. These non-numerical properties of data from scale assessments imply that calculation of sums and differences is meaningless.

Consequently evaluation of change in an outcome variable that is measured by assessments on a rating scale cannot be based on differences.

The statistical method by Svensson for evaluation of disagreement in paired ordered categorical assessments is also applicable to the evaluation of change since a change in outcome between two occasions means disagreement according to following. The detailed interpretations of the measures are given in the guidelines for reliability.

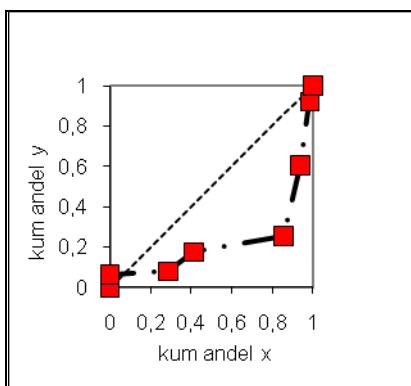
Evaluation of change in outcome assessed on a rating scale involves at least two assessments by each individual, for example before and after treatment. A complete *agreement* in the assessments made on the two occasions imply *unchanged* outcome. Observed *disagreement* is a sign of *change* in outcome for at least one individual. By the approach by Svensson it is possible to identify and measure the systematic part of a disagreement separately from addition occasional disagreement, see guidelines for reliability studies.

A systematic disagreement in change data is a sign of **common group changes**, while occasional variability indicate **individual variation in changes**, not explained by the common group change. Statistical evaluation of the homogeneity of group change is important, as it reflects the efficacy of a common intervention programme on the study group. Large individual changes are sign of heterogeneity in changes with means that individual interventions or treatments would be considered.

The interpretation of the measures of agreement and disagreement applied to evaluation of change will be demonstrated by a worked example regarding outcome in social functioning in 63 patients assessed on an eight-point scale by a psychiatrist, see figure 1. The categories are labelled from A (dead) to H (fully recovered).

Please note that the  $11 \times 11$  cross-table is used, but the  $8 \times 8$  lower cells were used only. According to the paired distribution, Figure 1, the 18 patients level C on the first occasion changed in the range between A and G, and 6 patients had unchanged outcome on the two assessments ( $PA = 6/63, 10\%$ ), see Table 1.

The Q-Q- curve (ROC Curve) and the marginal heterogeneity indicate presence of a systematic group change. The positive RP value (0.49) showed strong evidences of a systematic group change towards higher categories, and the rank-transformable pattern shows the pattern of group change that one can expect in the long run. The level C on the first occasion seems to be the most crucial level of outcome on the follow-up occasion (from A to F).



The paired distribution of data, Figure 1, differs from the expected pattern of group change, the RTP, which means that some patients were scored to higher and some to lower levels of outcome than expected (D, 13%). This heterogeneity in change is measured by the measures of disorder, D, and by the RV. This individual heterogeneity means that the recovery process is not in common for all patients even though the overall outcome for the group is improvement in outcome.

Table 1	X First occasion														
Y	A	B	C	D	E	F	G	H	I	J	K				
K												0			
J												0			
I												0			
H					1	2	1	1				5			
G			3	4	10	2	1					20			
F			4	2	15		1					22			
E			2	1	2							5			
D			4	1		1						6			
C			1									1			
B												0			
A			4									4			
	0	0	18	8	28	5	3	1	0	0	0	63			

Results		SE	95% confidence interval	
PA	10%			
RP	0,4873	0,0819	0,3268	0,6477
RC	-0,347	0,166	-0,67	-0,02
RV	0,1580	0,0641	0,0324	0,2837
D	0,1265			

RTP	X									
Y	A	B	C	D	E	F	G	H		
<b>H</b>						1	3	1		5
<b>G</b>					16	4				20
<b>F</b>			2	8	12					22
<b>E</b>			5							5
<b>D</b>			6							6
<b>C</b>			1							1
<b>B</b>		1								0
<b>A</b>			4							4
	0	0	18	8	28	5	3	1		63

**Read more:**

Sonn U, Svensson E. Measures of individual and group changes in ordered categorical data: application to the ADL Staircase. Scand J Rehab Med 1997;29:233-42

Svensson E. Ordinal invariant measures for individual and group changes in ordered categorical data. Statistics in Medicine 1998;17(24): 2923-36

Lagging LM, Westin J, Svensson E, Aires N, Dhillon AP, Lindh M, Wejstål, Norkrans G. Progression of fibrosis in untreated patients with Hepatitis C virus infection. Liver 2002; 22:136-144

Westin J, Lagging LM, Spak F, Aires N, Svensson E, Lindh M, Dhillon AP, Norkrans G, Wejstål R. Moderate alcohol intake increases fibrosis progression in untreated patients with hepatitis C virus infection. J of Viral Hepatitis 2002;9:235-41

Svensson E, Starmark JE. Evaluation of individual and group changes in social outcome after aneurysmal subarachnoid haemorrhage: a long-term follow-up study. J Rehab Med 2002;34:251-9

Svensson E. Statistical methods for repeated qualitative assessments on scales. Int J Audiol. 2003; 42 Suppl 1:13-22.

Engman E, Andersson-Roswall L, Svensson E, Malmgren K. Non-parametric evaluation of memory changes at group and individual level following temporal lobe resection for pharmaco-resistant partial epilepsy. Journal of Clinical and Experimental Neuropsychology 2004;26(7):943-54

Lund I, Lundeberg T, Carleson J, Sönerfors H, Uhrlin B, Svensson E. Corticotropin releasing factor in urine – a possible biochemical marker of fibromyalgia. Responses to massage and guided relaxation. Neuroscience Letters 2006;403:166-71

Lund I, Lundeberg T, Lönnberg L, Svensson E. Decrease of pregnant women's pelvic pain after acupuncture: a randomized controlled single-blind study. Acta obstet gynecol scand, 2006; 85:12-9