

Do SF-36 summary scores work as outcome measures for chronic functional disorders?



Schröder A, Ørnbøl E, & Fink P

The Research Clinic for Functional Disorders and Psychosomatics
Aarhus University Hospital, Denmark

ansch@as.aaa.dk

Background

The SF-36 is a widely used instrument to assess health related quality of life. Physical and Mental Component Summaries (PCS and MCS) are increasingly used as main outcome measures in clinical trials¹. Yet, studies in different patient populations have demonstrated serious problems in interpreting these summary scores²⁻⁴. We are conducting an RCT on cognitive behavioural therapy (CBT) in patients with severe and chronic functional disorders. Self-rated physical health is primary outcome. Based on a pilot study and baseline data, we have assessed the performance of the summary scores.

Aim

- To demonstrate problems in the orthogonal factor solution for physical and mental health summary scores of the SF-36.
- To assess other methods of summarising the SF-36 subscales.

Methods

We examined three methods of summarising physical and mental health subscales of the SF-36:

- The original summary scores PCS and MCS, which are based on an orthogonal factor solution of all 8 subscales, assuming that physical and mental health are uncorrelated¹.
- An alternative summary scoring based on an oblique factor solution of all 8 subscales, assuming positive correlation of physical and mental health.
- The summary scores (PHC and MHC) of the RAND-36 HSI⁵. This instrument uses the same 36 items, but scoring of subscales is based on item response theory⁶. Summary scores are based only on 4 subscales.

All three methods agree on two distinct factors, which are interpreted as dimensions of physical and mental health (Fig. 1). Yet, the construction of summary scores differ in weighting of subscales (Fig. 1).

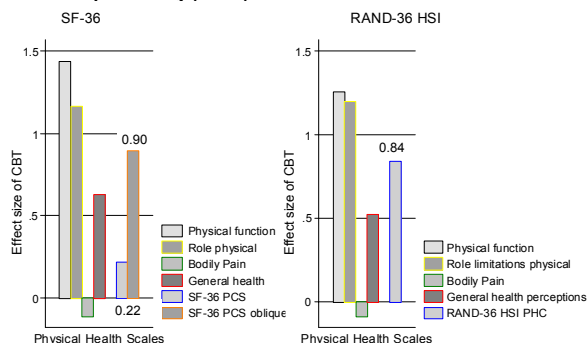
Fig. 1. SF-36 subscales and summary score coefficients (weights)

Summary measures	Subscales	Summary score coefficients (weights)					
		PCS	MCS	PCS o	MCS o	PHC	MHC
Physical health	Physical functioning	0.42	-0.23	0.35	-0.13	0.27	-
	Role limitations physical	0.35	-0.12	0.31	-0.03	0.35	-
	Bodily pain	0.32	-0.10	0.28	-0.00	0.28	-
	General health	0.25	-0.02	0.22	0.09	0.29	-
Mental health	Vitality	0.03	0.24	0.09	0.26	-	0.31
	Social functioning	-0.01	0.27	0.11	0.19	-	0.31
	Role limitations emotional	-0.19	0.43	-0.05	0.36	-	0.16
	Mental health	-0.22	0.49	-0.11	0.44	-	0.43

Results I

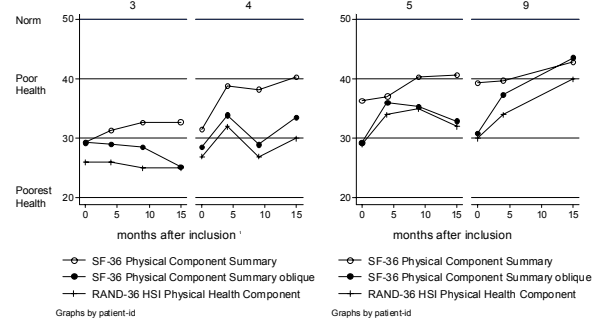
Pilot study: Effect sizes of CBT on physical health are shown in Fig. 2. The positive changes on the subscales were not reflected by the original PCS. Both the oblique PCS and the PHC summarised the effect sizes of the subscales more accurately. There were comparable effect sizes on mental health (MCS 1.99, MCS oblique 1.94, MHC 1.78). Examples of individual changes over time illustrate the substantial differences between the three summary scores (Fig. 3).

Fig. 2. Effect sizes of CBT on physical health at 12 months' follow-up. Results of pilot study (N=10)



Abbreviations: PCS: Physical Component Summary PHC: Physical Health Component
Effect sizes are calculated as $d = \frac{(\text{Mean}_{\text{CBT}} - \text{Mean}_{\text{Support group}})}{\text{SD}}$. SD derives from a large primary care sample (N=1500).

Fig. 3. Examples of individual changes in physical health summary scores in patients participating in support group (patient 3 and 4) and patients receiving CBT (patient 5 and 9).



Results II

Baseline data: Surprisingly, SF-36 PCS and MCS showed moderate negative correlation, which is contradictory to the orthogonal construction of the summary scores. The oblique summary scores were not correlated, and only the PHC and MHC showed positive correlation (Fig. 4). These differences in correlation are a result of the differences in weighting of subscales (Fig. 1) and were reflected in significant differences between the original MCS and both MCS oblique and MHC (Fig. 5). However, summary scores of physical health showed no significant differences (95% CI in brackets): PCS 31.6 (29.6-33.7), PCS oblique 29.9 (28.2-31.7), PHC 30.0 (28.5-31.5).

Fig. 4 Pearson's Correlation of summary scores (N=102).

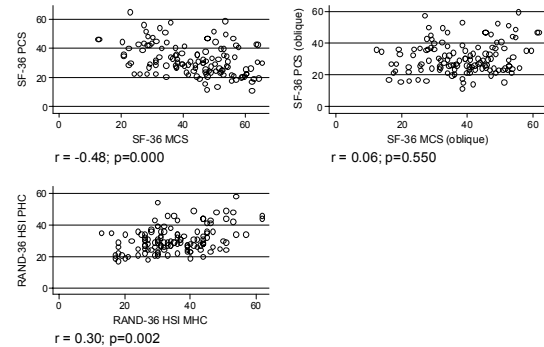
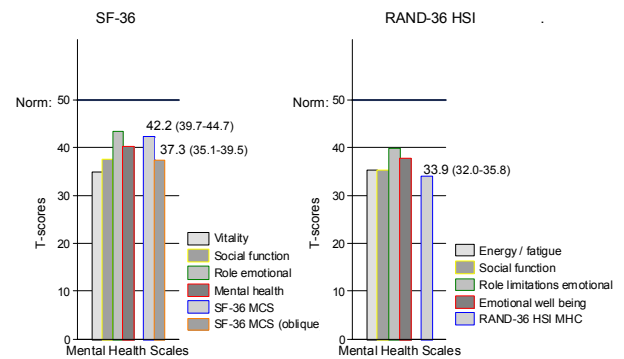


Fig. 5 Mental health in patients with severe chronic functional disorders (N=102) at time of referral to the Research Clinic.



Abbreviations: MCS: Mental Component Summary MHC: Mental Health Component
All T-scores are based on US norms; N=102 patients with chronic functional disorders

Conclusion

The negative weights of the original orthogonal factor solution distorted the results in both study samples. The SF-36 summary scores summarised the subscales inconsistently in both cross-sectional and longitudinal analyses of these samples. As outcome measures, PCS and MCS should be interpreted with caution in patients with functional disorders and in other patient groups with interacting physical and mental health problems. One may benefit from using other scoring methods such as the RAND-36 HSI to validate the results.

1. Ware JE, Kosinski M. SF-36 Physical and Mental Health Summary Scales: A Manual for Users of Version 1, Second Edition. Lincoln, RI: Quality Metric Inc.; 2001.
2. Nordvold MW, Rise T, Myhr KM, Nyland HL. Performance of the SF-36, SF-12, and RAND-36 summary scales in a multiple sclerosis population. *Med Care* 2000;38:1022-3.
3. Wilson D, Parsons J, Tucker G. The SF-36 summary scales: problems and solutions. *Soz Präventivmed*. 2000;45:239-46.

4. Tait C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? *Qual Life Res* 2001;10:395-404.
5. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med* 2001;33:350-7.
6. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000;38:1428-1442.