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 trials. Yet, studies in different patient populations have demonstrated serious problems in interpreting these summary scores. 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 uncorrelated.
- 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 HSI. This instrument uses the same 36 items, but scoring of subscales is based on item response theory. 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).

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).

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).

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.