Int J Public Health 52 (2007) 60–61
1661-8556/07/010060–2
DOI 10.1007/s00038-006-5091-9

Application of the German SF-36 in hospitals: Overestimations in the psycho-social scales

Horst Müller, Annegret Franke, Karl-Ludwig Resch

Forschungsinstitut für Balneologie und Kurortwissenschaft, FBK

Submitted: 12 July 2005; Revised: 16 March 2006, 6 June 2006; Accepted: 28 June 2006

The well-known SF-36 is the most widely used generic instrument to quantify health-related quality of life (Ware 2000; Ware et al. 1993). Although it was originally developed for epidemiological research, it is increasingly applied in in-patient settings. However, results from in-patients who are requested to fill out the SF-36 at the end of their hospital stay show high non-response rates for the items of three scales (Zwingmann et al. 1998). Since some of the terms in the German SF-36 (Bullinger & Kirchberger 1998) such as "at home" or "at work" do not make sense in a hospital setting, this is hardly surprising. To eliminate this problem we developed a slightly modified version of the SF-36. The psychometric evaluation of this modified SF-36 yielded favorable results (Müller et al. 2001). However, we also observed a small but statistically significant response shift bias yielding higher SF-36 scores in the hospital environment. It was not established whether this is unique to the original SF-36 or whether the modified instrument version is equally sensitive to this response shift bias. The present paper examines this issue.

Participants, methods, and results

To adapt the German SF-36 (Bullinger & Kirchberger 1998) for use in hospital settings we simply deleted 23 inappropriate terms like "at home" (zu Hause) and "at work" (bei der Arbeit) in item blocks four, five and eight. Both the resulting questionnaire – the SF-36m – as well as the original SF-36 were presented to 116 consecutive patients of a rehabilitation hospital with orthopedic, cardiologic, gynecologic and nephrologic units. These patients first responded to one of the two questionnaire versions while they were still at home, approximately 13 days (on average) before checking in at the hospital. The different versions of the questionnaire were assigned at random. Immediately upon their arrival at the hospital (beginning of in-patient rehabilitation), the patients answered the version of the SF-36 that had not been administered to them previously. Thus, each patient answered both versions in randomized order. More details on other aspects of this study have already been published in this journal (Müller et al. 2001).

The hospital scores for the original version of the SF-36 (but not for the modified version) were systematically higher than the home scores, i.e. the self-rated health scores assessed with the SF-36 at the hospital yielded higher subjective health scores compared to those assessed at home (Fig. 1a). As these differences were most pronounced in the four psycho-social scales, we tested the hypothesis of a response shift bias in the psycho-social scales of the original SF-36 via structural equation modeling. This was done using EffectLite 2.0 in combination with LISREL 8.7, adjusting for gender as a covariate. The observed bias for the psycho-social scales of the original SF-36 was strong (effect size = .56) and statistically significant (post hoc $p < .05$). The SF-36m did not show such systematic differences between home and hospital assessments (figure 1 b). Its scores correspond approximately to the original SF-36 in hospital settings.

Comment

The German SF-36 overestimated psycho-social scores if it was applied in a hospital environment compared to assessments at home. It was possible to eliminate this problem by deleting terms like "at home" and "at work" . The authors of the German SF-36 have already acknowledged the problems with the German SF-36 in hospital settings (Bullinger & Kirchberger 1998) and recommend our modified form for hospital contexts (Bullinger et al. 2003).

Interestingly, completely identical items (vitality, mental health, social function) yielded clearly different responses in the original versus the modified SF-36. The most reasonable explanation for this difference seems that these items are sensitive to preceding

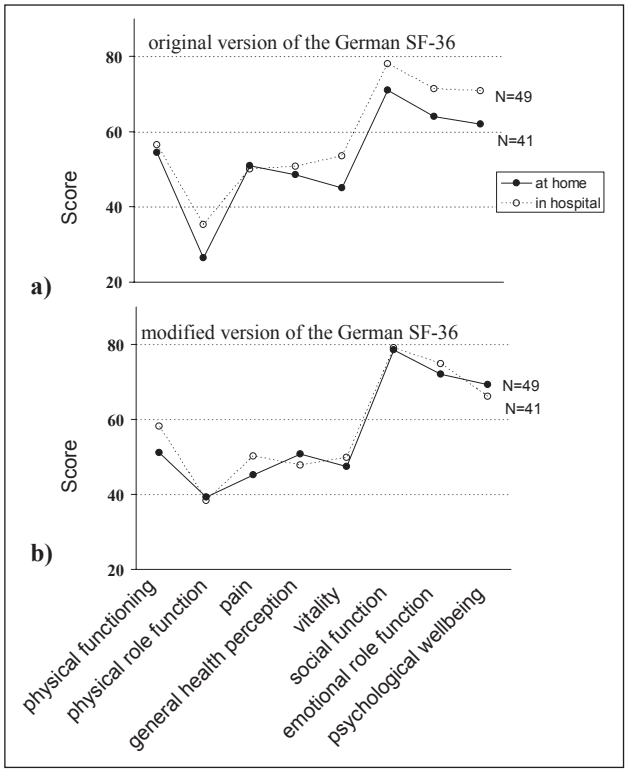

Figure 1 Mean Scores of the eight SF-36 scales for home and hospital settings a) for the original version and b) with minimal modification of some items.

questionnaire items. Such questionnaire context effects are not at all unusual and point towards one important reason why it is imperative to adhere to the questionnaire format and order unless there is an extremely good reason to depart from the standard.

Focusing on the original SF-36 only, one might speculate that the observed pre-post differences may actually reflect some kind of learning effect (the SF-36 in the hospital setting was always assessed after the SF-36 in the home setting). However, given that our minimally modified SF-36 (SF-36m) yielded stable scores at home and at the hospital, this seems to be rather unlikely. Nevertheless, this observation could have important implications for the application of the original SF-36: In typical quality assurance designs, in which the SF-36 is increasingly often applied, therapeutic effects may be underestimated if the SF-36 is first assessed at the beginning of a hospital stay, and then again as a follow-up at home. This could explain why researchers sometimes report difficulties in obtaining clear and positive long-term results with the SF-36. Our minor modifications of the SF-36 seem to eliminate this problem and should be taken into account in future versions of the German SF-36.

A further implication is that many published norm values for the SF-36, which were obtained in home settings only, may not be applicable for hospital settings. Psycho-social quality of life seems to be clearly overestimated at the beginning of a hospital stay, and in the worst of cases this could even hamper therapeutic decisions – however, the SF-36 is seldom applied as a diagnostic instrument.

Of course, our observations and conclusions are preliminary and restricted to the German version of the SF-36. The English version seems to be less problematic because of a slightly different wording. Moreover, we report data of a single study with a small sample size so that there is always the risk that the results are artificially significant. This threat to internal validity is even more serious since our hypothesis test was clearly post hoc. The SF-36m was developed in order to pragmatically solve a specific application problem that had been found with the original SF-36, and this study was not designed to deal with a potential response shift bias. Therefore, before final conclusions can be drawn, it is necessary to replicate these results for the German SF-36. Moreover, as the response shift bias is a rather general phenomenon, it seems recommendable to examine such problems in other language versions of the SF-36, and possibly for other questionnaires as well.

References

Bullinger M, Kirchberger I (1998). SF-36 Fragebogen zum Gesundheitszustand. Handanweisung. Göttingen: Hogrefe.

Bullinger M, Morfeld M, Kohlmann T et al. (2003). Der SF-36 in der rehabilitationswissenschaftlichen Forschung – Ergebnisse aus dem Norddeutschen Verbund für Rehabilitationsforschung (NVRF) im Förderschwerpunkt Rehabilitationswissenschaften. Rehabilitation (Stuttg), 42(4): 218–25.

Müller H, Franke A, Schuck P, Resch KL (2001). Eine kliniktaugliche Version des deutschsprachigen SF-36 und ihr psychometrischer Vergleich mit dem Originalfragebogen. Soz Präventivmed, 46, 96–105.

Ware JE (2000). SF-36 Health Survey Update. Spine, 25(24): 3130–9.

Ware JE, Snow KK, Kosinski M, Gandek B (1993). SF-36 Health Survey manual and interpretation guide. Boston, M.A.: New England Medical Center, The Health Institute.

Zwingmann C, Metzger D, Jäckel WH (1998). Short Form-36 Health Survey (SF-36): Psychometrische Analysen der deutschen Version bei Rehabilitanden mit chronischen Rückenschmerzen. Diagnostica 44(4): 209–19.

Address for correspondence

Dr. Horst Müller
FBK
Lindenstr. 5
D-08645 Bad Elster
Germany
Tel.: +49 37437 55 7 22
Fax: +49 37437 557 77
e-mail: horst.mueller@fbk.sms.sachsen.de