Soz.- Präventivmed. 45 (2000) 239–246
0303-8408/00/060239-08 $ 1.50 + 0.20/0

David Wilson, Jacqueline Parsons, Graeme Tucker

Department of Human Services, Adelaide

The SF-36 summary scales: Problems and solutions

Summary

To determine the accuracy of the SF-36 summary mental and physical health scales in reflecting their underlying subscales using the traditional method of scoring based on factor coefficients derived through principle components analysis and orthogonal rotation. A representative Australian population survey containing the SF-36 was used to obtain factor coefficients from principle components analysis and orthogonal rotation for scoring the physical component summary (PCS) and the mental component summary (MCS) of the SF-36 in the traditional way. In addition two other methods were used to produce coefficients. The first method used maximum likelihood extraction and oblique rotation. The second method fit a structural equation model to the data in a confirmatory factor analysis. The coefficients derived by each of the methods were applied to the data of a second representative population survey. This survey also provided data on physical and mental health status which allowed comparison of the summary scores and underlying subscales according to various health states. Neither of the scoring methods based on the exploratory factor analyses methods (orthogonal and oblique) produced summary scale scores, by age group, that adequately reflected the underlying subscales. When coefficients derived using structural equation modeling were fit to the data in a confirmatory factor analysis the MCS and PCS accurately reflected their underlying subscale scores. They also produced MCS and PCS scores for the various health states as would be expected from the underlying subscales. The traditional methods of scoring the SF-36 summary scales produce results that would not be expected from the underlying subscales. The problem was only corrected by fitting a structural equation model to the data in a confirmatory factor analysis. The results advise caution in the use of the SF-36 summary scales and suggests that alternative methods of developing factor coefficients need to be employed in studies using the SF-36 summary scales.

The short form SF-36 health related quality-of-life questionnaire has been widely accepted as a generic summary of health status[1–3] and as an investigative tool in health assessment or monitoring[4–5]. Despite this a study by Simon et al. in 1998 recommended caution in the interpretation of the mental component summary (MCS) and the physical component summary (PCS) when the condition or treatment of interest had strong effects on scales with negative scoring coefficients[7]. The original subscales of the SF-36 were aggregated into the physical and mental summary components using factor analysis of subscale scores from a United States population sample[8]. Principal components extraction and orthogonal rotation produced factor score coefficients, which were then used to compute the summary scores. The findings of Simon's study showed that the baseline physical component summary, in a prospective study of patients initiating antidepressant treatment, indicated no impairment based on a population norm of 50, although patients had reported modest impairment on the physical functioning, role-physical, bodily pain and general health perceptions, all of which contribute to the physical component summary in the hypothesized structure. In the three-month follow up stage of this study the four subscales showed moderate but statistically significant improvements, however the physical component summary was unchanged. The reason for this anomaly was attributed to the assumptions and methods used in computing the summary scores based on orthogonal factor rotation and the generation of negative scoring coefficients. In the algorithms used to score the summary components the mental health and role-emotional scales make modest negative scoring contributions to the PCS. In addition, the physical functioning, role physical and bodily pain subscales also make modest negative contributions to the MCS. The authors concluded that these negative scoring coefficients could produce unexpected results on either summary component scale.

In this study we looked again at the effects of scoring coefficients on the summary scales, how they compared with their underlying subscales and verified Simon's findings in the 1998 representative population South Australian Health Omnibus Survey (SAHOS). We then investigated other approaches of deriving scoring coefficients from the 1995 Australian National Health Survey, using i) maximum likelihood extraction and oblique rotation, and ii) structural equation modeling rather than Ware's[8] original method. These coefficients were then used to score the 1998 South Australian Health Omnibus Survey (SAHOS) and assess face validity of summary scores in comparison to subscale scores for a number of age groups (15–29 years, 30–49 years. 50–59 years and 70+ years). The SAHOS also provided information on health status, which allowed us to compare the subscales and summary scores according to physical, and mental health states reported. These summary scores were again calculated using i) principle components analysis and orthogonal rotation, ii) maximum likelihood extraction and oblique rotation, and, iii) structural equation modeling. These analyses provided an assessment of the external validity of the subscales and summary scales using the different methods of obtaining coefficients[9].

Method

Two population data sources were used for this study. The first was the 1995 Australian National Health Survey (NHS)[10] which was the second in a series of five yearly population surveys designed to obtain national benchmark information on a range of health issues. The sampling method of the NHS is a self-weighting multistage clustered area sample based on Australian Bureau of Statistics (ABS) collector districts in which households are selected with equal probability. In this survey n = 23800 households were selected and all adults aged 15 years or older were interviewed. A subset of n = 19785 were asked to complete the SF-36 quality-of-life questionnaire. Of those interviewed, n = 18492 provided some data on the SF-36 and were included in these analyses. These data were used by the ABS to calculate subscale scores using principle components analysis and orthogonal factor rotation after the method of Ware et al.[11] and produce coefficients that could be used to calculate summary scores for Australia[10].

Using the NHS data set two additional methods were used in the present study to compute factor coefficients for scoring the PCS and MCS which were then checked using the 1998 SAHOS. The factor coefficients were developed on the basis of the following logic.

First, exploratory factor analysis using maximum likelihood extraction and oblique (oblimin) rotation. Maximum likelihood extraction estimates the hypothetical factor structure, which assumes that one or more latent factors account for the correlations observed between manifest variables. The oblique rotation allows for the real world fact that physical and mental health is correlated, by allowing for this to be reflected in the rotation.

An obvious alternative to an exploratory factor analysis is to use a structural equation modelling approach to fit a confirmatory factor analysis to the implied structural model. This was undertaken using the eight subscale scores as manifest variables. Measurement error was incorporated into the model based on the reliability of the subscale measures (using Cronbach's alpha for the NHS data set as the reliability measure), by constricting the error variances of the manifest variables representing the subscale scores. This model provided a poor fit to the NHS data (RSMEA = 0.55) and the approach of calculating scores based on the subscales was therefore rejected.

A hypothetical factor structure has already been documented for the SF-36 (see Figure 1). It was therefore possible to fit a structural equation model to the data in a confirmatory factor analysis (as above). The model fit was the full measurement model, using items re-coded as detailed in the SF-36 scoring manual. As this model is based on the hypothetical factor structure, coefficients are only produced for the relevant variables related to each summary score. The other paths do not appear in the model and are therefore not estimated.

The above model was fit on the covariance matrix of 18141 observations with no missing data for all eight SF-36 subscales following imputation of data items by mean substitution, where more than half the data items in a subscale were not missing, as set out in the SF-36 scoring manual.

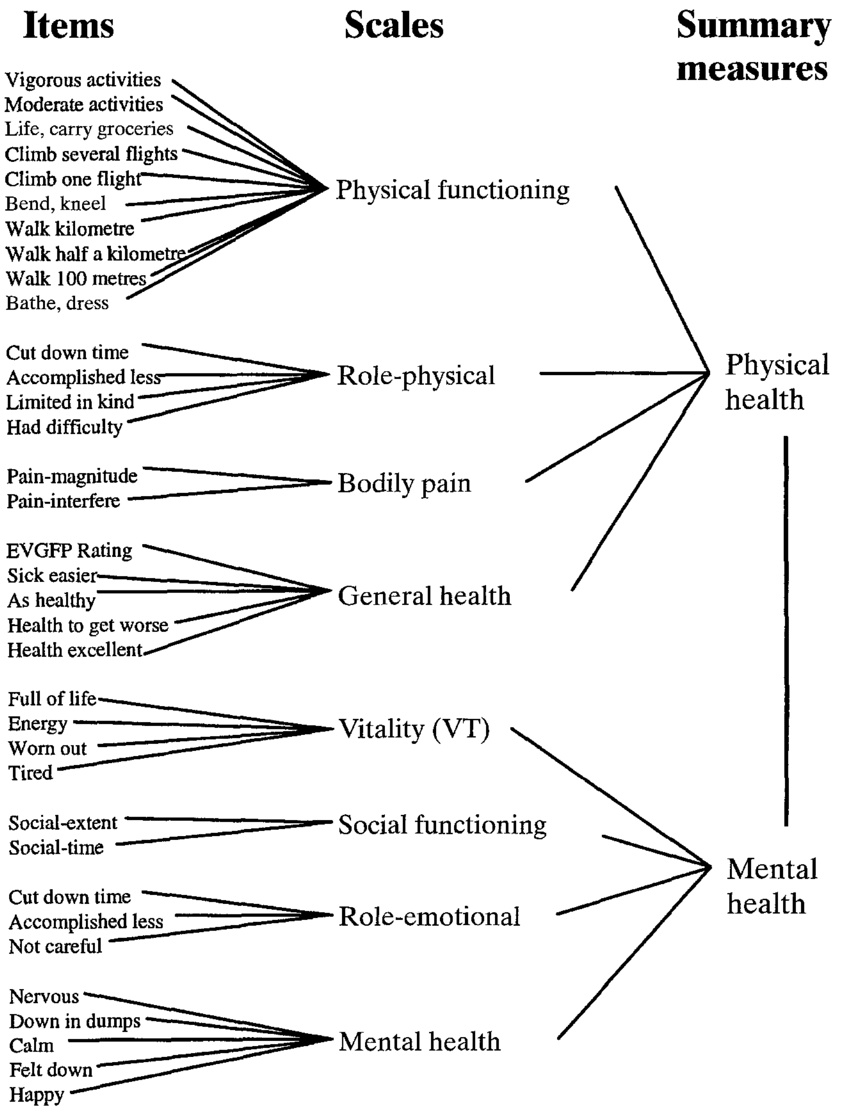

Figure 1. *Hypothesized structure of SF-36 health dimensions and the summary mental (MCS) and physical (PCS) health measures.*

Amos includes an option for handling missing data without imputation known as full information maximum likelihood (FIML) estimation. The full measurement model was also fit to the complete NHS data file of 18492 complete or partial returns, with no imputation of missing values but using FIML estimation. This model does not provide fit measures, but does provide parameters. When these parameters were compared to the model above (using imputation by mean substitution), the parameters were very similar with minor differences appearing generally at the third decimal place of the parameter. It was therefore concluded that this analysis was not biased by any problems relating to missing data and it was therefore accepted as the preferred approach.

All coefficients derived from the NHS data set were then used to score the summary scales of the SF-36 data obtained in the independent 1998 SAHOS (n = 3001, 70% response rate). The SAHOS is a statewide South Australian survey designed to obtain state benchmark information on a range of health issues. The SAHOS is also a self-weighting multistage clustered area sample selected from ABS collector districts of people aged 15 years or more who live in metropolitan Adelaide or country towns with populations over 1000. The survey is conducted annually and the method has been extensively published[3]. Households are selected with equal probability of selection within each collector district and then one adult in each household, aged 15 years or older is selected for interview according to the most recent birthday.

In addition to the SF-36, a range of other health and demographic questions were asked in the 1998 SAHOS. The survey has been running since 1990, and response rates average 73% with approximately 3000 people interviewed. In the 1998 survey the response rate was 70%. Data from the survey were weighted so they accurately represented the age, sex, household size, and geographic area of the South Australian population. The SF-36 data collected in the survey of 1998 were used to test the results of the different extraction and rotation methods.

SAHOS respondents were also asked whether or not in the previous twelve months they had used any medication for a chronic physical condition: that is one that has lasted for, or is likely to last for, six months or more. They were also asked the same question for depression (medication such as tranquilisers or anti-depressants) or a diagnosed mental illness (such as schizophrenia). From these data the study population were divided into four groups according to no medication, physical health medication only, mental health medication only or both physical and mental health medication. These implied health groups were then used to assess the external validity of the subscales and the PCS and

MCS. Mean scores for the SF-36 subscales and the PCS and MCS summary scales were calculated for each of the four health groups and compared with the overall subscale means to assess whether or not the scoring of the SF-36 subscale and summary dimension scores were consistent with the different health states as would be predicted. The external validity of the PCS and MCS subscale scores of the SAHOS data were assessed using the three methods identified above.
Data were analysed using SPSS version 9 and AMOS version 3.6.2.[12]

Results

The data in Table 1 illustrates the problem of using orthogonal or oblique rotation methods to calculate the PCS and MCS subscales. It can be seen from Table 1 that the orthogonal method produces some age-specific MCS scores that vary around the overall MCS score. This in itself is not a problem. It does, however, become a problem when the age specific MCS is higher than would be expected given the underlying subscale scores. Scores for three subscales that make up the orthogonal MCS are significantly lower for the over 70 age group (vitality, social functioning and role emotional) than for the 15–29 and 30–49 age group, yet the orthogonal MCS score for the over 70 age group is significantly higher (Tukey's honestly significant difference (HSD) <0.001) than that for the younger age groups. In the same way the oblique method produced MCS summary scores that are not significantly different across all age groups despite significantly lower subscales (vitality, social functioning

Dimension	Age group (years)				
	15–29	30–49	50–69	70+	Overall population
Physical functioning	91.2	87.5	77.6	61.4	83.0
Role physical	87.6	82.7	74.6	63.8	79.8
Bodily pain	81.0	77.7	72.2	71.0	76.5
General health	76.4	76.8	71.2	64.7	73.9
Vitality	67.3	64.5	63.7	58.9	64.3
Social functioning	88.6	88.5	88.2	84.0	87.9
Role emotional	88.3	87.1	89.0	86.5	87.8
Mental health	80.1	79.0	80.7	81.6	80.0
Summary PCS (ORTHOGONAL extraction and rotation: using ABS weights based on NHS)	52.5	51.0	46.4	41.1	49.1
Summary MCS (ORTHOGONAL extraction and rotation: using ABS weights based on NHS)	51.4	51.3	53.3	54.2	52.1
Summary PCS (maximum likelihood extraction, OBLIQUE rotation)	52.4	51.2	48.5	44.7	50.1
Summary MCS (maximum likelihood extraction, OBLIQUE rotation)	47.7	48.4	48.0	48.4	48.1
Summary PCS (using regression coefficients from STRUCTURAL EQUATION MODELING)	52.8	51.8	48.3	44.4	50.3
Summary MCS (using regression coefficients from STRUCTURAL EQUATION MODELING)	52.1	51.1	51.3	49.9	51.2

Table 1. *SF-36 dimension and summary scores for adult age groups using various summary approaches.*

and role emotional) for the over 70 age group.
These results are due to the mathematics that converts the subscales to summary scores. Each subscale score is converted to a Z score before being multiplied by the factor score coefficient. The summary score is computed by adding the Z score, multiplied by the coefficient, for each subscale (not only the four that make up the dimension). This creates a problem when a negative Z score is multiplied by a negative coefficient, as it results in a positive number inflating the opposing scale. In this case the scores on the physical subscales for people over 70 years were lower than the population norm, resulting in a negative Z score. This was then multiplied by the negative coefficient and summed to give an artificially high MCS using the orthogonal method. Table 1 also shows the MCS and PCS scores calculated using structural equation modeling. The model used to derive the coefficients for the PCS and MCS scores was found to provide an adequate fit to the data (RMSEA 95% CI = 0.068, 0.069) based on 18141 input records. Most fit indices were about 0.82. It can be seen from Table 1 that both the PCS and MCS summary scales, using structural equation modeling, produce scores that are consistent with their underlying subscales as hypothesised. The over 70

Dimension	MEDICATION TAKEN				
	No medication	Physical medication only	Mental medication only	Both physical & mental	Overall population
Physical functioning	88.7	68.4	79.0	56.6	83.0
Role physical	86.9	63.3	69.0	39.1	79.8
Bodily pain	81.2	65.2	71.6	49.6	76.5
General health	79.2	61.8	66.1	41.7	73.9
Vitality	67.6	58.3	49.2	41.0	64.3
Social functioning	90.9	83.0	77.1	58.8	87.9
Role emotional	90.6	85.5	63.0	56.5	87.8
Mental health	81.9	78.6	61.3	59.2	80.0
Summary PCS (ORTHOGONAL extraction and rotation: using ABS weights based on NHS)	51.8	41.9	50.0	37.5	49.1
Summary MCS (ORTHOGONAL extraction and rotation: using ABS weights based on NHS)	52.8	52.5	41.5	40.9	52.1
Summary PCS (maximum likelihood extraction, OBLIQUE rotation)	52.5	44.7	46.4	36.0	50.1
Summary MCS (maximum likelihood extraction, OBLIQUE rotation)	46.7	49.9	58.2	61.6	48.1
Summary PCS (using regression coefficients from STRUCTURAL EQUATION MODELING)	53.2	43.6	47.1	34.6	50.3
Summary MCS (using regression coefficients from STRUCTURAL EQUATION MODELING)	52.8	48.9	41.6	37.3	51.2

Table 2. *Subscale and summary dimension scores for four health conditions.*

age group produced an MCS score that was significantly lower than the youngest age group (Tukey HSD = 0.003) and as would be expected given the underlying subscales. The 30–49 and 50–69 age groups produced scores differing in the right direction according to their subscales, but did not achieve statistical significance. The PCS using structural equation modeling produced PCS scores as would be expected given the underlying subscales.

Table 2 shows the subscale and summary scores for the four groups with varying medical conditions according to the medication used. The PCS and MCS summary scores have again been computed using coefficients obtained by all three methods reported above. Using the logic of "known groups" validity according to Ware et al.[11] it can be seen from Table 2 that there are external validity problems with the exploratory factor analysis methods. The orthogonal model produced a higher PCS score for people taking mental health medication (Tukey HSD = 0.05) and a higher MCS score for those taking physical health medication (Tukey HSD = 0.05) when compared to the overall scores, despite lower subscale scores than the overall scores. The oblique method produced the lowest MCS score for those taking no medication and the highest MCS scores those taking both physical and mental medications. Thus the ranking was reversed from that expected. The structural equation coefficients again behaved appropriately producing higher and lower summary scores than overall where expected.

In assessing external validity, the PCS and MCS means are ordered as would be expected using the structural equation coefficients. The group reporting using no physical or mental health medication had higher MCS and PCS scores than overall. The group taking physical health medication had a lower PCS score than the overall score and also lower than the score for the group reporting mental health medication only. This latter group, in turn, had a lower MCS score than overall and those reporting physical medication only.

Discussion

The first question in this study was to identify whether or not the findings of Simon et al.[7] were verified in producing summary scores for the SF-36 that did not reflect their underlying health subscales. We would conclude from the population survey data used in this study that Simon was correct in identifying problems with the summary scales. While Simon's problem related to the PCS in assessing change in health status over time this study found problems with the MCS when comparing summary scores across age groups. We would, however, express the caution that given the large sample size in this study statistically significant differences are relatively easy to achieve. This means that we must also be careful not to over interpret small but significant differences between groups.

The second question asked was whether an alternative approach to deriving the scoring coefficients would produce summary scores with improved face validity. Structural equation modeling to derive coefficients used in producing the summary PCS and MCS scores has produced scores that are consistent with underlying subscale scores across age groups. On face value this has corrected the problems that were identified by Simon et al.[7]. The fact that structural equation modeling uses only those dimensions from the hypothesized structure that make up the PCS and the MCS also makes logical sense. When a person responds to a question related to any subscale they have the opportunity to respond to both physical and mental health issues in a way they judge to reflect their current health status. If they have a physical health problem, which also affects their mental health, or vice versa, they have the opportunity to identify the level of this effect when they answer each of the mental health questions for each subscale. The impact of mental health issues on physical health, is therefore already accounted for in the individuals direct answer to each physical health question, without accounting for it again by virtue of the statistical analysis method used (factor analysis and orthogonal rotation) which includes a weight for the MCS subscales. Similarly, the impact of physical health is reflected in the physical health questions without adding a weight for the MCS subscales. It does not make logical sense that there is a weight from subscales making up the PCS included in the MCS, or vice versa. This would appear to be double counting if the respondent has already accurately answered questions that contribute to the subscales, as occurs in the exploratory factor analysis methods. As a result, and as Simon et al.[7] point out, it produces incorrect results that are based on the factor methods used.

McCallum[13,14] in his earlier validation of the SF-36 for Australia, has suggested that structural equation modeling to calculate summary scores is inappropriate because the SF-36 items are highly inter-correlated. In his analyses the raw data did not produce a positive definite matrix but rather a singular variance/covariance matrix. In the present study, where a larger and more representative data set was used the matrix was non-singular. It should also be pointed out that McCallum did not have access to the NHS data when he conducted his validation studies.

The data shown in this study would also lead us to agree with Ware's conclusion that oblique rotation is not an alternative scoring strategy[11]. It would appear initially that Ware did not consider the further alternative of structural equation modeling, even though he has created a hypothesized structure that would suggest such an approach. In more recent work, however, structural equation modeling was used in assessing the construct validity of the SF-36 across ten countries[15]. In this study the authors concluded that the study results confirmed the hypothesized relationships between the SF-36 items and scales.

The ten-country study mentioned above also raised the possibility that summary scale scores may vary across different cultural groups. In this study non-significant differences were observed in the PCS and MCS with greatest variation occurring in the MCS[15]. In another study conducted in Japan by some of the same investigators caution was expressed regarding some of the scales that comprise the MCS[16].

The third question asked in the present study was whether or not the PCS and MCS summary scores truly measured health in varying and understandable medical health states. Four mutually exclusive groups known to differ in the type and severity of self-reported physical and mental conditions (no conditions, physical health conditions, mental health conditions and both physical and mental conditions) were associated with the PCS and MCS as predicted according to type and severity of condition[11]. The summary scores which best represented the underlying subscales were produced from structural equation modeling.

The conclusion we would draw is that structural equation modeling corrects the effects of negative scoring coefficients produced in scoring methods based on exploratory factor analyses. It should be stated, however, that this is the first Australian assessment using structural equation modeling and further population studies should be conducted to assess reliability of the method over time.

Zusammenfassung

Die Summenskalen des SF-36: Probleme und Lösungen

Die zur Berechnung der Summenskalen der physischen (PCS) und psychischen Dimension (MCS) des SF-36 nötigen Faktorwerte wurden in einer repräsentativen australischen Bevölkerungsbefragung nach dem Standardverfahren durch eine Hauptkomponentenanalyse mit orthogonaler Rotation ermittelt. Zusätzlich wurden zwei weitere Verfahren zur Berechnung der Koeffizienten angewendet: eine Faktorenextraktion nach Maximum-Likelihood mit anschliessender schiefwinkliger Rotation und die Anpassung eines Strukturgleichungsmodells an die Daten in einer konfirmatorischen Faktoranalyse. Die so berechneten Faktorwerte wurden in einer zweiten repräsentativen Bevölkerungsbefragung verwendet. In dieser Erhebung wurden zusätzlich verschiedene Masse zur physischen und psychischen Gesundheit erhoben, die einen Vergleich der Summenskalen und der zugrunde liegenden Subskalen in Gruppen mit unterschiedlichem Gesundheitsstatus erlaubt. Keine der auf Basis explorativer Faktoranalysen (orthogonale oder schiefwinklige Rotation) berechneten Summenskalen bildet die zugrunde liegenden Subskalen in verschiedenen Altersgruppen adäquat ab. Werden die Faktorwerte in einer konfirmatorischen Faktoranalyse mit einem Strukturgleichungsmodell ermittelt, entsprechen die Summenskalen MCS und PCS den zugrunde liegenden Subskalen besser. Auch die aufgrund der Subskalen erwarteten Unterschiede in Gruppen mit unterschiedlichem Gesundheitsstatus konnten reproduziert werden. Die Standardverfahren zur Berechnung der Summenskalen des SF-36 zeigen Ergebnisse, die aufgrund der zugrunde liegenden Subskalen nicht zu erwarten sind. Eine bessere Entsprechung konnte in einer konfirmatorischen Faktoranalyse durch die Anpassung eines Strukturgleichungsmodell an die Daten erzielt werden. Die Ergebnisse weisen darauf hin, dass die Interpretation der Summenskalen des SF-36 mit Vorsicht zu erfolgen hat und dass alternative Verfahren zur Berechnung der Faktorwerte angewendet werden sollten.

References

1 *Anderson J ST C, Sullivan F, Usherwood TP.* The Medical Outcomes Study instrument – use of a new health status measure in Britain. Fam Pract 1990; *7*: 205–18.

2 *Brazier JE, Harper R, Jones NMB,* et al. Validating the SF-36 health survey questionnaire: a new outcome measure for primary care. BMJ 1992; *305*: 162–4.

3 *Wilson D, Parsons J, Wakefield M.* The health related quality-of-life of never smokers, ex-smokers and light, moderate, and heavy smokers. Prev Med 1999; *29*: 139–44.

4 *Lyons RA, Lo SV, Littlepage NC.* Perception of health amongst ever-smokers and never-smokers: a comparison using the SF-36 health survey questionnaire. Tob Control 1994; *3*: 213–5.

5 *Ziebland S.* The short form 36 health status questionnaire: clues from the Oxford region's normative

Résumé

Les scores synthétiques du SF-36: problèmes et solutions

Pour déterminer si les échelles synthétiques de santé mentale et physique du SF-36 reflètent correctement les huit échelles sous-jacentes lorsqu'on utilise la méthode traditionnelle de scorage basée sur des coefficients dérivés de l'analyse des composantes principales suivie d'une rotation orthogonale. Une enquête représentative de la population australienne utilisant le SF-36 a été mise à profit pour calculer les scores synthétiques physique (PCS) et mental (MCS) du SF-36 selon la méthode traditionnelle, à partir des coefficients issus d'une analyse factorielle exploratoire de composantes principales suivie de rotation orthogonale des huit scores SF-36 initiaux. De plus, deux autres méthodes furent utilisées pour générer les coefficients. La première méthode utilisait la maximisation de la vraisemblance et la rotation oblique. La seconde méthode appliquait un modèle d'équations structurale aux données dans une analyse factorielle confirmatoire. Les coefficients dérivés dans chacune des méthodes furent appliqués aux données d'une deuxième enquête représentative de population. Cette enquête fournit également des données sur la santé physique et mentale qui permettent de comparer les scores synthétiques aux échelles sous-jacentes selon différents états de santé. Aucune des méthodes de scorage basée sur les analyses factorielles exploratoires (orthogonale et oblique) n'a produit des scores synthétiques, par groupes d'âge qui reflétait de façon adéquate ces échelles sous-jacentes. Lorsque les coefficients dérivés des équations structurales furent appliqués aux données dans une analyse factorielle confirmatoire, le MCS et PCS reflétaient correctement les échelles sous-jacentes. Ils produisirent aussi des scores de MCS et PCS pour les différents états de santé que l'on aurait pu attendre avec les échelles sous-jacentes. Les méthodes traditionnelles de scorage des échelles synthétiques du SF-36 produisent des résultats qui n'auraient pas été attendus avec les échelles sous-jacentes. Ce problème peut être corrigé par l'application d'un modèle d'équations structurale aux données dans une analyse factorielle confirmatoire. Ces résultats suggèrent que les échelles synthétiques du SF-36 devraient utilisés avec prudence et que les méthodes alternatives pour développer les coefficients factoriels devraient utilisés dans ce genre d'étude.

data about its usefulness in measuring health gain in population surveys. J Epidemiol Community Health 1995; *49*: 102–5.

6 *Garrat AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT.* The SF-36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? BMJ 1993; *306*: 1440–4.

7 *Simon GE, Revicki DA, Grothaus L, Von Korff M.* SF-36 summary scores. Are physical and mental health truly distinct. Med Care 1998; *36*: 567–72.

8 *Ware JE, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A.* Comparison of methods for the scoring and statistical analysis of the SF-36 Health Profile and Summary Measures: summary of results from the Medical Outcomes Study. Med Care 1995; *33*: AS264–AS272.

9 *McHorney CA, Ware JE, Raczek AE.* The MOS 36-item short form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1981; *19*: 247–63.

10 Australian Bureau of Statistics. National Health Survey. SF-36 Population Norms Australia. Canberra: Australian Bureau of Statistics, 1995. (Catalogue Number 4399.0).

11 *Ware JE, Kosinski M, Keller SD.* SF-36 Physical and Mental Health Summary Scales: a users manual. Boston, MA: The Health Institute, New England Medical Centre, 1994.

12 *Arbuckle JL.* Amos Users Guide Version 3.6. Small Waters Corporation. Chicago: SPSS Inc. 1997.

13 *McCallum J.* The SF-36 Physical and Mental Health Summary Scales: Australian validation. Proceedings of Health Outcomes and Quality-of-Life Measurement Conference 1995. Canberra: Australian Institute of Health and Welfare 1995.

14 *McCallum J.* The SF-36 in an Australian sample: validating a new, generic health status measure. Aust J Public Health 1995; *19*: 160–6.

15 *Keller SD, Ware JE, Bentler PM,* et al. Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: results from the IQOLA Project. J Clin Epidemiol 1998; *51*: 1179–88.

16 *Fukuhara S, Ware JE, Kosinski M, Wada S, Gandek B.* Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol 1998; *51*: 1045–53.

Address for correspondence

David Wilson
Department of Human Services
PO Box 6, Rundle Mall PO
Adelaide
South Australia 5001