A comparison of two multidimensional measures of health status

Van der Zee, K.I.; Sanderman, Robbert; Heyink, J

A comparison of two multidimensional measures of health status: The Nottingham Health Profile and the RAND 36-Item Health Survey 1.0

K. I. VanderZee,* R. Sanderman and J. Heyink

Department of Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712
TS Groningen, The Netherlands (K. I. VanderZee); Northern Centre for Health
Care Research (R. Sanderman) and Department of Health Sciences (J. Heyink),
University of Groningen

In this study the applicability of two multidimensional instruments, the NHP and the RAND 36-Item Health Survey 1.0, for measuring health status in population surveys was examined. A population sample of 1,063 persons aged over 17 years participated in the study. It was shown that, as compared with the NHP, the RAND 36-Item Health Survey 1.0 is a more reliable measure of health status. Second, within a group of subjects who scored 'zero' on the NHP, considerable dispersion in RAND 36-Item Health Survey 1.0 scores was found. For the whole group, no significant differences were found in the amount of variance explained by the corresponding scales from both instruments in the prevalence of chronic diseases. However, among subjects with a zero score on the NHP, the RAND 36-Item Health Survey 1.0 scores were still predictive of the occurrence of chronic diseases. it was concluded that, compared with the NHP, the RAND 36-Item Health Survey 1.0 seems to be a more sensitive instrument for the use in population samples.

Key words: Health Status measurement; NHP; population surveys; RAND 36-Item Health Survey 1.0.

Introduction

Several instruments have been constructed to survey the health status of certain populations and to evaluate the effect of health care programmes. Most of these instruments are focused on a global evaluation of health status, or they are aimed at diagnosing the impact of specific diseases. Global measures make a comprehensive indication of health impossible, while at the same time instruments aimed at specific diseases are of limited application when there is a need to compare patients suffering from different diseases. Therefore there is a growing need for multidimensional measures to survey general health status. In this study the applicability of two multidimensional instruments: The Nottingham Health Profile (NHP)[1] and the RAND 36-Item Health Survey 1.0[2,3,4] for measuring health status in population surveys will be examined.

The Nottingham Health Profile is a widely used scale which contains subscales for physical functioning, social isolation, pain, sleep, emotional reactions and energy. The NHP has important advantages because it is short and easily administered. In addition, studies have shown that the Nottingham Health Profile is a valid and sensitive measure of subjective health.[5] However, critics claim that it is an insensitive instrument for use in population surveys, because its modal response is zero, indicating no health problems, and the dimensions measured by the instrument are insufficiently distinct.[6,7] However, its use is still widespread due probably to the applicability and the high content validity and face validity of the NHP.

In this study we will examine how a recently developed short form health survey, the RAND 36-Item Health Survey 1.0[2,3,4] compares to the NHP in terms of its sensitivity for use in general population samples. This 36-item questionnaire was developed from longer instruments developed for the Medical Outcomes Study,[8] which in turn were based, in part, on items included in the RAND Health Insurance Study.[9] The health definition of the World Health Organization[10] was taken as a basis for the construction of these instruments. Three dimensions of health are central to this definition: physical, mental and social health. The RAND 36-Item Health Survey 1.0

* To whom correspondence should be addressed at Department of Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands.

contains subscales for physical functioning, social functioning, role limitations (physical problem), role limitations (emotional problem), mental health, vitality, pain and general health perception. The reliability and validity of the instrument both for the use in population samples[11,12,13] and in patient samples,[14,15,16,17] have been evaluated favourably in preliminary studies in England and in the United States.

In this study the reliability and validity of the RAND 36-Item Health Survey 1.0 in general population surveys will be compared with the psychometric qualities of the NHP. First, the reliability of the scales and the dispersion of the scores will be examined. Second, it will be investigated if the RAND 36-Item Health Survey 1.0 is capable of discriminating between the large group of zero responders on the NHP reported in previous research, as was noted already.[6,7] Third, the sensitivity of the RAND 36-Item Health Survey 1.0 and the NHP in predicting chronic illness will be examined. An instrument aimed at measuring dimensions of health should in the first place discriminate between subjects suffering from chronic diseases and healthy subjects. Finally, to evaluate construct validity, the relationship between scales from both instruments and measures of mental and physical health will be examined.

Methods

Procedure

A sample of 3,000 inhabitants of a Dutch township aged over 17 years was selected randomly from the population register. Three subsamples of 1,000 persons received different versions of a questionnaire. Because of the large number of scales that we wanted to include in the study it was impossible to include all scales in all questionnaires. For each group, the RAND 36-Item Health Survey 1.0 was combined with the NHP, the List of Chronic Diseases and some additional measures of physical and mental health which differed for each version. In addition, every subject received an accompanying letter in which the purpose and procedure of the study were explained. The respondents could return their questionnaire in a stamped addressed envelope. At the end of the questionnaire, the respondents were asked if they were willing to participate in a follow-up study. After two weeks the whole sample received a reminder. To guarantee anonymity, this reminder was sent to both respondents and non-respondents. For the former group, the letter was used to thank people for their

cooperation. Two months later a subgroup of 200 subjects selected randomly from those of the original sample who agreed to participate in a follow-up study received a second questionnaire. This questionnaire contained the NHP, the RAND 36-Item Health Survey 1.0, the List of Chronic Diseases and some additional personal characteristics (age, sex and marital status).

Measurements

The RAND 36-Item Health Survey 1.0[*,2,3,4] is a self-administered questionnaire which contains 36 questions measuring both positive and negative health states. The items of the RAND 36-Item Health Survey 1.0 are identical to the items of the SF-36.[14,18] The RAND 36-Item Health Survey 1.0 differs from the SF-36 in the recommended scoring method.[3] For the purpose of this study, the American version of the RAND 36-Item Health Survey 1.0 was translated into the Dutch language and backtranslated for comparison with the original version. To reach optimal comparability between the American and the Dutch version, an attempt was made to remain as close to the original version as possible. The layout of the questionnaire was chosen in accordance with the original American version. Eight scales are included in the RAND 36-Item Health Survey 1.0. The scale for *physical functioning* considers limitations performing daily activities as climbing stairs, bathing or dressing oneself, or carrying groceries, as a result of health problems. The *social functioning* scale considers limitations in social activities, such as visiting friends or relatives. The scale *role limitations (physical problem)* measures problems with work or other daily activities as a result of physical health problems during the last four weeks, while the scale *role limitations (emotional problem)* considers role limitations due to emotional problems. The scale for *mental health* contains questions about feelings of depression and nervousness. The items of the *vitality* scale consider feelings of energy and tiredness. The questions about *pain* consider the amount of pain and limitations due to bodily pain. The scale for *general health perception* measures the subjective evaluation of general health status. Finally, one item is added considering *health change*. This item considers general health compared to one year ago. The item is not included in one of the eight dimensions. For six of the scales the respon-

---

*The version of the survey we used in this study differed slightly from the final version (VanderZee and Sanderman, 1994). The Dutch version of the RAND 36-Item Health Survey 1.0 is developed in cooperation with Dr JCM de Haes.

dents are asked to give their ratings on 3–6 point Likert scales. In computing scale scores, item scores were coded and summed into scale scores. The scores were transformed into a scale from 0 (worse health) to 100 (best health).

The *Nottingham Health Profile (NHP)*[1] is also a self-administered questionnaire containing 38 questions. The respondents are asked to give their answers by responding 'yes' or 'no' to each question. For example: 'I sleep badly at night'. The NHP contains scales for physical mobility, social isolation, sleep, pain, emotional reactions and energy. The scale scores are computed by summing the weighted or unweighted item scores. In this case, we calculated scale scores from the unweighted item scores. The scales scores

Table 1. Incidence of chronic diseases[19]

a) asthma, bronchitis, chronic non-specific lung disease (CNSLD)	11%[1]
b) infection of the nasal cavity, the sinus cavity or the maxillary sinus	10%
c) heart condition/coronary diseases	4%
d) hypertension	12%
e) stroke	1%
f) stomach ulcer	2%
g) abdominal disorders (> three months)	3%
h) gallstones or gallbladder infection	1%
i) hepatitis or cirrhosis of the liver	<1%
j) kidney stones	1%
k) kidney disease	<1%
i) chronic bladder infection	2%
m) diabetes	2%
n) thyroid gland defect	3%
o) serious backtroubles (> three months) or hernia	14%
p) abrasion of the joints	14%
q) rheumatoid arthritis (hands or feet)	6%
r) other forms of rheumatoid arthritis (> three months)	3%
s) epilepsy	1%
t) dizziness with falling	3%
u) migraine	10%
v) skin diseases	2%
w) malignant diseases or cancer	2%
x) irreversible injury as a result of an accident	4%
ii) psychological problems, e.g. anxiety, depression, overstrain	12%

Note: These figures represent the percentage of subjects that suffered from a specific disease.

were computed by summing the affirmative responses per individual, with a lower total indicating a higher health status.

The *List of chronic diseases*[19] consists of 25 chronic diseases. For each chronic disease, the respondent has to answer three questions: (1) Do you suffer from this disease? (2) Did you see a doctor for this disease? and (3) Did you take any medicines for this disease in the last 12 months? The version used in this study was a revised version of the list frequently used by the Central Bureau of Statistics in population surveys. The incidences of the chronic diseases that were considered are listed in Table 1.

The *Groninger Activity Restriction Scale (GARS)*[20,21] was used as a measure of physical health. The Groninger Activity Restriction Scale (GARS) measures functional limitations and considers the extent of limitation in performing 18 specific tasks. The GARS has subscales for Activities of Daily Living (ADL: 11 items) and Instrumental Activities of Daily Living (IADL: 7 items).

The *Center for Epidemiological Studies Depression Scale (CES-D)*[22] was used as a measure of mental health. The CES-D is a 20-item instrument designed to measure an individual's current level of depressive symptomatology, with emphasis on depressed mood.

The *State and Trait Anxiety Scale (STAI)*[23] was used as a second measure of mental health. This twenty item anxiety scale considers feelings like tension, nervousness, confusion.

The final measure of mental health we used was the *General Health Questionnaire (GHQ)*[24] (Goldberg, Williams, 1988) which measures psychological distress and concerns itself with two major classes of phenomena: inability to continue to carry out one's normal healthy functions and the appearance of new phenomena of a distressing nature. The presently used measure is a short 12-item version of the GHQ.[25] The GHQ was included in two versions of the questionnaire, the STAI, the CES-D and the GARS were included in one version of the questionnaire.

Statistical analyses

With respect to sample characteristics, χ^2 tests were performed to examine differences between the three groups (receiving different questionnaires) with respect to age, sex and education. To examine the

representativeness of the sample, χ^2 tests were also performed for differences between the study sample and the population of the township with respect to the same characteristics.

To examine the internal consistency of the scales, Cronbach's α was computed. A test statistic recommended by Feldt[26] was used to test whether the scale reliabilities for both instruments differed significantly:

$$t = \frac{(\zeta_1 - \zeta_2)\,(n-2)^{1/2}}{[\,4\,(1-\zeta_1)\,(1-\zeta_2)\,(1-\rho^2)\,]^{1/2}} \quad (Df = n-2)$$

whereby ζ_1 and ζ_2 refer to the reliability coefficients of the corresponding scales from both instruments, n to the sample size, and ρ^2 to the squared correlation between the scores on the two corresponding subscales.

As an indicator of the stability of the scales, correlations between test and retest scores were computed. However, correlations between test and retest scores give no indication of the direction of the association between these scores. Even when scores on the retest are consistently lower, a high positive association will be found. Therefore, paired samples *t*-tests were performed to test if the mean scale scores on both measurement points were significantly different. The reliabilities for both instruments were compared by performing Z-tests on the Fisher transformed values of the correlations (and alpha's) for the scales from both instruments.

In order to examine the sensitivity of both instruments for use in general population surveys the distributions of RAND 36-Item Health Survey 1.0 scores for subjects with a zero response ('no' answers to *all* 38 questions) on the NHP were considered. The dispersion of the RAND 36-Item Health Survey 1.0 scale scores among this group of non-responders was examined.

Moreover, the sensitivity of both instruments in predicting chronic diseases was examined. First, for both instruments, hierarchical linear regression analyses were performed for the number of chronic diseases on the five common scales. The comparable scales were entered in the regression in the same order. Second, based on the distribution of the number of chronic diseases, three categories were formed of subjects who: (1) did not suffer from a chronic disease; (2) suffered from one chronic disease or (3) suffered from more than one chronic disease. Respectively 40%, 27% and 33% of the subjects fell into these categories. RAND 36-Item Health Survey 1.0 and NHP scores were computed for each category. Uni-

variate analyses of variance were performed for the effects of category on the scale scores obtained with both instruments. Third, the sensitivity of the RAND 36-Item Health Survey 1.0 in predicting chronic diseases among those subjects who scored zero on the NHP was considered. For these subjects the scores on the different RAND scales were divided into high ('healthy') vs. low ('unhealthy') according to the median split. Univariate analyses of variance were performed to examine the effect of health status on the different dimensions on the number of chronic diseases. Finally, for the purpose of construct validation, correlations were considered between the common RAND 36-Item Health Survey 1.0 scales and NHP scales and corresponding scales from different instruments aimed at measuring physical health (GARS) and mental health (CES-D, STAI and GHQ). Support for the construct validity of an instrument is obtained when higher correlations are found with corresponding scales than with non-corresponding scales from different instruments.

Results

Sample characteristics

From the total sample, 35.4% (n =1,063) responded. In terms of age-sex characteristics the sample was approximately 35% male and 65% female, aged between 18–89 years (mean = 44.1 years; standard deviation = 17.5 years), with 22.3% aged over 60 and no more than 2.8% aged over 80 years. In addition, 66% of the sample was married, 7% cohabited, 5% had a partner but lived apart, 14% reported to be unmarried, 3% was divorced, and 5% was widowed. About 42% received lower education (primary school or lower professional training), 40% middle education (secondary school or middle professional training) and 18% received higher education or higher professional training. No significant differences were found between the three groups (receiving different questionnaires) for age, sex and educational characteristics.

It must be noted that the overall response was low, in particular among male subjects. It is possible that our sample is not totally representative of the total population. It seems that women and the younger ages are overrepresented in the sample. Indeed, significant differences between the sample and the population of the township were found with respect to age (p <0.01) and sex (p <0.01) distribution. No significant differences were found between the sample and the population with respect to educa-

tional level ($p = 0.12$). Because we are not primarily interested in the prevalence of illness but rather in correlations between the RAND 36-Item Health Survey 1.0 scales and comparable scales from other instruments, we assume that this bias will not seriously affect our results.

Reliability and validity

In Table 2 the means, standard deviations and reliability coefficients that were found for the NHP scales are presented. As can be seen, the dispersion of the scale scores was not very high. On the whole, 381 subjects scored zero (best health). The internal consistency reliability coefficients for the scales were moderately high (between 0.65–0.88). The internal consistency reliabilities for the RAND 36-Item Health Survey 1.0 scales assessing physical functioning ($t(1016)=2.65$; $p < 0.05$), social functioning ($t(1013)=3.30$; $p < 0.05$), mental health ($t(994)=6.79$; $p < 0.01$) and vitality ($t(992)=6.79$; $p < 0.01$) were significantly larger than for the corresponding NHP scales. Reliability estimates for the corresponding pain scales did not differ significantly.

In the lower half of Table 2, the corresponding results that were found for the RAND 36-Item Health Survey 1.0 scale are shown. Although the scale means

all have values towards the end of the scale, there is still considerable dispersion for all scales (SD's varying from 18.4–35.5). Compared with the NHP scales, the internal consistency of the scales was very high. Alpha coefficients ranged between 0.71–0.92.

To examine the sensitivity of the RAND 36-Item Health Survey 1.0 and the NHP to changes in health status, after 2 months a sample of 200 persons was re-examined, of which 159 persons responded. The correlations between test and retest scores for the NHP and the RAND 36-Item Health Survey 1.0 are given in Table 2. As can be observed, for both instruments, scores were highly correlated over time. Next, paired samples *t*-tests were performed to test if the mean scale scores on both measurement points were significantly different. For both scales, none of the differences in scale means were significant. Thus, it seems that the results from both measures did not show much fluctuation over a 2 month period.

Next, we examined the score distributions of the subjects with a zero response (only 'no' answers) on the NHP. For these subjects, Table 3 shows means, standard deviations, the upper and lower limits and the percentage of subjects with an optimal score ('100') for each RAND 36-Item Health Survey 1.0 scale score.

Table 3 shows a remarkable dispersion in RAND 36-Item Health Survey 1.0 scores of this group with

Table 2. Means (M), standard deviations (SD), α values, and test–retest reliability of the NHP and the RAND 36-Item Health Survey 1.0

Dimension	M	SD	α	Test–retest reliability*
NHP				
Physical mobility	0.7	1.3	0.75	0.68
Social isolation	0.6	0.8	0.65	0.79
Emotional reactions	0.9	1.5	0.75	0.80
Energy	0.4	0.8	0.71	0.73
Pain	0.6	1.5	0.88	0.79
Sleep	0.7	1.2	0.74	0.80
RAND 36-Item Health Survey 1.0				
Physical functioning	81.9	23.2	0.92	0.82
Social functioning	86.9	20.5	0.71	0.58
Mental health	76.8	18.4	0.85	0.73
Vitality	67.4	19.9	0.82	0.76
Pain	79.5	25.6	0.88	0.72
Role limitations (physical problem)	79.4	35.5	0.90	0.60
Role limitations (emotional problem)	84.1	32.3	0.86	0.67
General health perception	72.7	22.7	0.81	0.80

* Test–retest correlations after a two month interval ($n = 159$)

Table 3. Means (M), standard deviations (SD), upper and lower limits and percentages of subjects with an optimal score on the RAND 36-Item Health Survey 1.0 scales for 'zero NHP scorers' (n=381)

Dimension	M	SD	Score interval	% with score '100'
Physical functioning	93.0	12.8	[10,100]	54.2
Social functioning	95.7	9.4	[42,100]	77.2
Mental health	85.4	10.7	[40,100]	7.2
Vitality	77.4	12.0	[30,100]	3.8
Pain	92.6	14.0	[22,100]	48.7
Role limitations (physical problem)	94.1	19.7	[0,100]	89.4
Role limitations (emotional problem)	96.0	16.3	[0,100]	93.2
General health perception	83.1	14.0	[20,100]	12.8

* Test–retest correlations after a two month interval (n = 159)

zero-scores on the NHP. Remarkably, although for some of the RAND 36-Item Health Survey 1.0 scales (that is for both scales for role limitations and the scale for social functioning) a large majority of the zero-NHP scorers obtained a maximum score, for most other scales less than half of the subjects obtained a maximum score.

Next, we compared the sensitivity of the NHP and the RAND 36-Item Health Survey 1.0 to the occurrence of chronic illness. As can be seen (Table 4), the five scales of the RAND 36-Item Health Survey 1.0 did not explain significantly more variance in the number of chronic diseases compared with the NHP, respectively, $R^2 = 0.36$ and $R^2 = 0.33$ ($t = 0.97$, ns). For both instruments, the corresponding scales for social functioning (social isolation and social functioning respectively) do not add significant variance to the total amount of explained variance. Thus, it seems that there are no significant differences in the amount of variance explained by the corresponding scales from both instruments.

Second, Table 5 shows the RAND 36-Item Health Survey 1.0 and the NHP scores for subjects who: (1) did not suffer from a chronic disease; (2) suffered from one chronic disease or (3) suffered from more than one chronic disease. As can be seen, significant effects were found of category on the scale scores obtained with both instruments. Thus, the RAND 36-Item Health Survey 1.0 and the NHP both seem good predictors of the number of chronic diseases.

Next, we considered the sensitivity of the RAND 36-Item Health Survey 1.0 scales in predicting the number of chronic diseases among subjects who scored zero on the NHP. For this group, Table 6 shows the mean numbers of chronic diseases that were found for low and high scorers on each RAND 36-Item Health Survey 1.0 scale and the results of

Table 4. Results of hierarchical linear regression analyses of the NHP and the RAND 36-item Health Survey 1.0 scales[†] on the number of chronic diseases

	R	R^2	F$_{change}$
NHP			
Physical mobility	0.49	0.24	293.56**
Pain	0.52	0.27	41.23**
Social isolation	0.53	0.28	22.39**
Emotional reactions	0.56	0.32	47.67**
Energy	0.57	0.33	15.91**
RAND 36-Item Health Survey 1.0			
Physical functioning	0.51	0.26	324.3**
Pain	0.57	0.32	90.3**
Social functioning	0.58	0.34	27.2**
Mental health	0.60	0.35	29.4**
Vitality	0.60	0.36	5.4*

Significance levels ** $p < 0.001$; * $p < 0.05$.
† For the RAND 36-item Health Survey 1.0 higher scores correspond with better health, for the NHP higher scores correspond with worse health.

univariate tests for the effect of scale score (dichotomized into low vs. high) on the number of chronic diseases. As it shows, for all scales a significant effect of scale score on the number of chronic diseases was found, with the exception of the scales for social functioning and role limitations as a result of physical problems. This finding suggests that the other six scales provide valid information that is not captured by the NHP.

Finally, the construct validity of both scales was considered by relating corresponding RAND 36-Item Health Survey 1.0 and NHP scales to common measures of physical (GARS) and mental health

Table 5. Scores for the NHP and RAND 36-Item Health Survey 1.0 scales[†] that were found for subjects who did not suffer from a chronic disease, who suffered from one chronic disease, and who suffered from more than one chronic disease respectively.

Number of chronic diseases	0	1	>1	F (2,805)
NHP				
Physical mobility	0.21 (0.65)	0.47 (0.93)	1.47 (1.76)	70.54*
Social isolation	0.14 (0.48)	0.26 (0.68)	0.55 (1.01)	24.36*
Emotional reactions	0.44 (0.92)	0.67 (1.27)	1.77 (2.05)	60.82*
Energy	0.11 (0.42)	0.27 (0.62)	0.78 (1.01)	49.64*
Pain	0.08 (0.43)	0.36 (1.09)	1.46 (2.28)	54.55*
Sleep	0.37 (0.80)	0.52 (0.99)	1.29 (1.50)	47.81*
RAND 36-Item Health Survey 1.0				
Physical functioning	91.71 (15.15)	82.94 (20.98)	67.51 (26.71)	89.90*
Social functioning	93.89 (12.93)	88.88 (17.40)	75.33 (25.81)	63.35*
Mental health	82.88 (13.26)	77.64 (17.35)	67.42 (21.57)	66.17*
Vitality	75.83 (14.38)	68.44 (18.53)	54.83 (21.27)	97.47*
Pain	92.37 (15.25)	82.98 (22.02)	64.52 (28.41)	96.66*
Role limitations (physical problem)	93.26 (21.20)	80.64 (33.30)	59.29 (42.97)	64.38*
Role limitations (emotional problem)	92.64 (21.87)	86.14 (30.30)	70.55 (40.62)	36.46*
General health perception	82.89 (14.23)	72.80 (19.73)	53.86 (22.62)	163.79*

Significance levels * $p < 0.001$.
[†]For the RAND 36-Item Health Survey 1.0 higher scores correspond with better health, for the NHP higher scores correspond with worse health.

Table 6. Mean number of chronic illnesses that was found for low vs. high scorers on each RAND 36-Item Health Survey 1.0 scale and the results of univariate ANOVA's for the effect of health status on chronic illness for subjects with a zero score on the NHP (n=344).[†]

RAND 36-Item Health Survey 1.0	Low		High		F (1,343)
	M	SD	M	SD	
Physical functioning	0.73	0.92	0.40	0.73	16.34***
Social functioning	0.67	0.93	0.52	0.81	1.84
Mental health	0.64	0.87	0.48	0.81	3.68*
Vitality	0.69	0.99	0.43	0.67	7.74*
Pain	0.80	1.01	0.43	0.71	20.03***
Role limitations (physical problem)	0.80	0.82	0.52	0.83	2.42
Role limitations (emotional problem)	0.85	1.01	0.53	0.82	4.47*
General health perception	0.80	1.03	0.38	0.62	23.47***

Significance levels *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.
[†]As a result of missing values the number of subjects included in the multivariate analysis was lower than the total number of zero-scorers (n =381) that was reported in Table 3.

(GHQ, STAI and CES-D). As Table 7 shows, scales from both instruments show high correlations with the corresponding measures of physical and mental health, that is, the physical scales are highly correlated with the GARS subscales, whereas the scales for psychological health are highly correlated with the scales for anxiety, depression and the GHQ. Note that the correlations between the RAND 36-Item Health Survey 1.0 scale for mental health and the three other measures for psychological well-being are significantly higher than the correlations between the emotional reaction scale from the NHP and these three measures. In a similar vein, correlations between the vitality scale and these indicators of

Table 7. Correlations of the NHP and the RAND 36-Item Health Survey 1.0 scales with scales from the GARS, GHQ, CES-D, and STAI[†‡]

	GARS (n = 323)	ADL	HDL	GHQ (n = 633)	CES-D	STAI (n = 272)
NHP						
Physical mobility	0.68**	0.62**	0.65**	0.17**	0.26**	0.26*
Pain	0.48**	0.40**	0.48**	0.19**	0.33**	0.23*
Social isolation	0.07	0.08	0.06	0.54**	0.62**	0.48**
Emotional reactions	0.06	0.02	0.09	0.57**	0.68**	0.62**
Energy	0.40**	0.37**	0.38**	0.44**	0.39**	0.34**
RAND 36-Item Health Survey 1.0						
Physical functioning	-0.65**	-0.59**	-0.65**	-0.28**	-0.28**	-0.29**
Pain	-0.41**	-0.37**	-0.42**	-0.39**	-0.37**	-0.37**
Social functioning	-0.41**	-0.39**	-0.40**	-0.64**	-0.65**	-0.59**
Mental health	-0.08	-0.10	-0.06	-0.76**	-0.81**	-0.72**
Vitality	-0.34**	-0.31**	-0.35**	-0.61**	-0.68**	-0.59**

Significance levels ** $p < 0.001$, * $p < 0.01$.
†For the RAND 36-Item Health Survey 1.0 scale higher scores correspond with better health, for the other scales higher scores correspond with worse health. ‡GARS=Groningen Activity Restriction Scale; GHQ=General Health Questionnaire; CES-D=Center for Epidemiological Studies Depression Scale; STAI =State and Trait Anxiety Questionnaire.

psychological health are also significantly higher than the correlations between the energy scale and these indicators. The energy scale from the NHP shows comparable correlations with the physical scales (GARS-subscales) and the psychological scales (STAI, CES-D and GHQ). Further, the GARS subscales are more highly correlated with the RAND 36-Item Health Survey 1.0 social functioning scale (r's of -0.39 to -0.41) than with the NHP social isolation scale (r's of 0.06 to 0.08).

Discussion

This study shows that, compared with the NHP, the RAND 36-Item Health Survey 1.0 is a more reliable measure of health status in the sample surveyed. The internal consistency reliability of all RAND 36-Item Health Survey 1.0 scales was high, and on the average higher than for the NHP scales. High test–retest correlations (2 month interval) were found for both instruments. Test–retest correlations (mean correlation 0.77 and 0.71 for the NHP and the RAND 36-Item Health Survey 1.0, respectively) in this study were similar to those reported by Brazier and his colleages[11] (mean correlation 0.71).

Previous studies using the NHP, have reported a modal response of 0 (only 'no'-answers), making the instrument an insensitive instrument for use in popu-

lation surveys, whilst not discriminating between subjects suffering from mild health problems. In this study, we found considerable dispersion in RAND 36-Item Health Survey 1.0 scores within this zero response group. Hence, the RAND 36-Item Health Survey 1.0 may be better suited for use in population surveys, as compared with the NHP.

The sensitivity of the RAND 36-Item Health Survey 1.0 to the occurrence of chronic illness does not differ significantly when compared with the NHP. Hierarchical regression analyses with the inclusion of only the five corresponding scales showed that the RAND 36-Item Health Survey 1.0 scales do not explain significantly more variance in the number of chronic diseases than do the NHP scales. Moreover, forming three groups according to the number of chronic diseases, significant effects of group on the common scale scores from both instruments were found. Thus, both scales seem capable of predicting chronic illness. Interestingly, we did find support that the RAND-scales provide valid information that is not captured by the NHP. Among individuals with a zero score on the NHP, high and low scorers on each RAND 36-Item Health Survey 1.0 scale differed significantly in the number of chronic diseases. Apparently, even among subjects with the best health status as measured by the NHP, the RAND 36-Item Health Survey 1.0 is able to account for variation in the number of chronic diseases.

In this study, the presence of chronic diseases was measured by the 'List of Chronic Diseases'. The validity of taking 'number of chronic complaints' as a quantifying indicator of chronic illness might be criticized. The number of diseases gives no indication of the seriousness of the complaints. A combination of migraine and dizziness is not necessarily twice as worse, compared with a form of cancer. Besides, an illness in a progressive stage can not be compared with a disease in a beginning stage. Future studies should take the seriousness of the specific case of illness into account.

Examining the relation between the RAND 36-Item Health Survey 1.0 scales and the NHP on the one hand and scales from other instruments on the other hand, it was found that for both instruments correlations between subscales and corresponding measures of physical and mental health were higher than the correlations between the subscales and non-corresponding measures. Both instruments did not differ very much in this respect. Noteworthy is that vitality as measured by the RAND 36-Item Health Survey 1.0 seems to reflect mental health whereas energy as measured by the NHP reflects both physical and mental health. In addition, the social functioning scale from the RAND 36-Item Health Survey 1.0 was strongly correlated with measures of physical health, whereas the corresponding NHP scale for social isolation only correlated significantly with mental health. It seems that the 'common scales' for vitality and social functioning measure different things. In an earlier study,[27] factor analysis revealed that the RAND 36-Item Health Survey 1.0 vitality scale and social functioning scale both loaded substantially on two factors one reflecting positive psychological well-being, and the other reflecting negative psychological well-being. Further research is necessary on the construct validity of these scales.

Summarizing, compared with the NHP, the RAND 36-Item Health Survey 1.0 seems a highly reliable instrument capable of discriminating between subjects who only suffer from mild health problems. The RAND 36-Item Health Survey 1.0 is even able to discriminate among subjects with a score of zero on the NHP. How both instruments compare with respect to their sensitivity to chronic illness should be further investigated using more valid criteria, like diagnose by doctors or classifications of the severity of diseases. However, on the basis of the current study it seems reasonable to conclude that the RAND 36-Item Health Survey 1.0 seems to be a more sensitive instrument than the NHP for the use in population samples (see also VanderZee, Heyink and Sanderman, 1994).[28]

References

1. Hunt S, McKenna SP, McEwen J. *The Nottingham Health Profile User's Manual*. Manchester: Galen Research & Consultancy, 1989.
2. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Economics* 1993; **2**: 217–227.
3. RAND Health Sciences Program. *RAND 36-Item Health Survey 1.0*. Santa Monica, CA: RAND, 1992.
4. VanderZee KI, Sanderman R. *Het meten van de algemene gezondheidstoestand met de RAND-36. Een handleiding.* [Measuring general health status with the RAND-36. Users Manual]. Groningen: Northern Center of Health Care Research, 1994
5. Hunt S, McKenna SP, McEwen J, Williams J, Papp, E. The Nottingham Health Profile: Subjective health status and medical consultations. *Soc Sci Med* 1981; **15a**: 221–229.
6. Jenkinson C, Fitzpatrick R, Argyle M. The Nottingham Health Profile: An analysis of its sensitivity in differentiating illness groups. *Soc Sci Med* 1988; **27(12)**: 1411–1414.
7. Kind P, Carr-Hill R. The Nottingham Health Profile: A useful tool for epidemiologists? *Soc Sci Med* 1987; **25**: 905–910.
8. Stewart AL, Sherbourne CD, Hays RD, *et al*. Summary and discussion of MOS measures. In Stewart AL, Ware JE, Jr, eds. *Measuring functioning and well-being: The medical outcomes study approach*. Durham: Duke University Press, 1993.
9. Brook RJ, Ware JE, Davies-Avery A, Stewart AL, Donald CA, Rogers WH, Williams KN, Johnston SA. Overview of adult health status measures fielded in RAND's Health Insurance Study. *Med Care* 1979; **15**; 724–735.
10. WHO. Constitution of the World Health Organization. *Basic Documents*. Geneva: WHO, 1948.
11. Brazier JE, Harper R, Jones NMB, Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health questionnaire: Outcome measure for primary care. *Brit Med J* 1992; **305**: 160–164.
12. Jenkinson C, Coulter A, Wright L. Short form (SF-36) health survey questionnaire: normative data for adults in working age. *Brit Med J* 1993; **306**:1437–1440.
13. Garatt 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? *Brit Med J* 1993; **306**:1440–1443.
14. Ware JE, Sherbourne CD. The SF-36 Short-form Health Status Survey: I. Conceptual framework and item selection. *Med Care* 1992; **30(6)**:473–483.
15. McHorney CA, Ware JE, Rogers W, *et al*. The validity and relative precision of MOS short- and long-form health status scales and Darthmouth COOP charts; Results from the Medical Outcomes Study. *Med Care* 1992; **30(5)**:253–265.
16. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short Form Health Status Survey (SF-36). II. Psychometric and clinical tests of validity in measuring physical and

mental health constructs. *Med Care* 1993; **31**:247–263.

17. McHorney CA, Ware JE, Lu JFR, Sherbourne CD. The MOS 36-Item Short Form Health Survey (SF-36). III. Tests of data quality, scaling assumptions and reliability across diverse patient groups. *Med Care* 1994; **32(1)**:40–66.

18. Aaronson NK, Acquadro C, Alonso J, Apolone G, Bucquet D, Bullinger M, *et al*. International Quality Of Life assessment (IQOLA) project: special communication. *Qual Life Res* 1992; **1**:349–351.

19. VandenBos GAM. Zorgen van en voor chronische zieken. (Cares of and caring for those who suffer from chronic illness). Dissertation, University of Amsterdam. Utrecht/ Antwerpen: Bohn, Scheltema and Holkema, 1989.

20. Katz S, Ford AB, Moskowitz RW, Jackson BA Jaffe MW. Studies of illness in the aged: the index of ADL: A standardized measure of biological and psychosocial function. *J Amer Med Assoc* 1963; **185**: 914–919.

21. Kempen GIJM, Doeglas DM, Suurmeijer ThPBM. *Het meten van problemen met zelfredzaamheid op verzorgend en huishoudelijk gebied met de Groningen Activiteiten Restrictieschaal (GARS). Een handleiding.* [Assessment of problems with activities of daily living and instrumental activities of daily living with the Groninger Activity Restriction Scale. Manual]. Groningen: University of Groningen, 1993.

22. Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. *App Psycho Measure* 1977; **1**: 385–401.

23. Ploeg HM, Defares PB, Spielberger CD. *Handleiding bij de Zelf-Beoordelings Vragenlijst (ZBV): een Nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory (STAI-DY).* [Manual for the Dutch version of the Spielberger State-Trait Anxiety Inventory]. Lisse: Swetz & Zeitlinger, 1981.

24. Goldberg D, Williams P. *A User's Guide to the General Health Questionnaire*. Windsor, Berkshire: NFER-NELSON, 1988.

25. Sanderman R, Stewart R. The assessment of psychological distress: Psychometric properties of the General Health Questionnaire (GHQ). *Int J Health Sci* 1991; **1(3)**: 195–202.

26. Feldt LS, Woodruff DJ, Salih FA. Statistical inference for coefficient alpha. *App Psycho Measure* 1987; **11(1)**: 93–103.

27. VanderZee KI, Sanderman R, Heyink J. De psychometrische kwaliteiten van de MOS 36-item Short Form Health Survey (SF-36) in een Nederlandse populatie [The psychometric qualities of the MOS 36-item Short Form Health Survey (SF-36) in a Dutch population sample]. *Tijdschrift voor Sociale Gezondheidszorg* 1993; **71**: 183–191.

28. VanderZee KI, Heyink J, Sanderman R. A comparison of two multidimensional measures of health status: The RAND-36 and the NHP. Poster presented at the First Meeting of the International Society of Quality of Life Research, Brussels, 1994.