



HEALTH

- CHILD POLICY
- CIVIL JUSTICE
- EDUCATION
- ENERGY AND ENVIRONMENT
- HEALTH AND HEALTH CARE
- INTERNATIONAL AFFAIRS
- NATIONAL SECURITY
- POPULATION AND AGING
- PUBLIC SAFETY
- SCIENCE AND TECHNOLOGY
- SUBSTANCE ABUSE
- TERRORISM AND HOMELAND SECURITY
- TRANSPORTATION AND INFRASTRUCTURE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Health](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use.

This product is part of the RAND Corporation reprint series. RAND reprints reproduce previously published journal articles and book chapters with the permission of the publisher. RAND reprints have been formally reviewed in accordance with the publisher's editorial policy.

The RAND-36 measure of health-related quality of life

Ron D Hays and Leo S Morales

The RAND-36 is perhaps the most widely used health-related quality of life (HRQoL) survey instrument in the world today. It is comprised of 36 items that assess eight health concepts: physical functioning, role limitations caused by physical health problems, role limitations caused by emotional problems, social functioning, emotional well-being, energy/fatigue, pain, and general health perceptions. Physical and mental health summary scores are also derived from the eight RAND-36 scales. This paper provides example applications of the RAND-36 cross-sectionally and longitudinally, provides information on what a clinically important difference is for the RAND-36 scales, and provides guidance for summarizing the RAND-36 in a single number. The paper also discusses the availability of the RAND-36 in multiple languages and summarizes changes that are incorporated in the latest version of the survey.

Keywords: generic profile measure; health-related quality of life; RAND-36; SF-36.

Ann Med 2001; 33: 350–357.

The RAND-36 measure of health-related quality of life

Health-related quality of life (HRQoL) refers to how health impacts on an individual's ability to function and his or her perceived well-being in physical, mental and social domains of life. The functioning part of HRQoL includes basic activities, such as self-care (eg, bathing, dressing), as well as work-related activities (whether paid or not) such as housework and career. It also includes the extent to which one is able to interact with family and friends (social functioning).

From the UCLA School of Medicine, Los Angeles, CA; RAND, Santa Monica, CA, USA.

Correspondence: Ron D Hays, PhD, Division of General Internal Medicine and Health Services Research, UCLA Department of Medicine, 911 Broxton Plaza, Room 110, Box 951736, Los Angeles, CA 90095-1736, USA. E-mail: hays@rand.org, Fax: +1 310 7940732.

Functioning is thought to be relatively objective because self-report information can be compared with other sources of data such as observations or performance measures (1). The well-being part of HRQoL is somewhat more subjective than the functioning part, because it relies almost exclusively on the internal, subjective perceptions of the respondent. Included is whether the person feels happy, sad, depressed, or anxious (emotional well-being), whether they are in severe pain or have no pain at all, and whether they are energetic or lethargic. Comprehensive measures of HRQoL include items assessing physical, mental, and social domains of life.

A fundamental distinction between HRQoL measures is whether they are disease-targeted or generic. Disease-targeted HRQoL measures are designed to be applicable and relevant to a particular disease such as diabetes or cancer. In contrast, generic HRQoL measures are designed to be applicable to anyone. They are analogous to intelligence tests in that different people can be compared to one another because they have taken the same test. Generic measures have two basic forms: profile and preference-based. Profile measures are designed to yield scores on multiple aspects of HRQoL. Preference-based measures are designed to produce a single summary score that cuts across the multiple domains of HRQoL. Because this special issue has an economic emphasis, several of the papers are devoted to preference-based measures such as the EQ-5D, Assessment of Quality of Life (AQoL), 15D and Health Utilities Index (HUI). The RAND-36 health survey is an example of a profile measure.

The RAND-36 is perhaps the most widely used HRQoL survey instrument in the world today. It is comprised of 36 items selected from a larger pool of items used in the RAND Medical Outcomes Study (MOS) (2). Twenty of the items are administered using a past 4 weeks' reporting interval. RAND-36 assesses eight health concepts with multi-item scales (35 items): physical functioning (10 items), role limitations caused by physical health problems (4 items), role limitations caused by emotional problems

(3 items), social functioning (2 items), emotional well-being (5 items), energy/fatigue (4 items), pain (2 items), and general health perceptions (5 items). An additional single item assesses change in perceived health during the last 12 months. Physical and mental health summary scores are also derived from the eight RAND-36 scales (see below).

The RAND-36 takes about 7–10 minutes to self-administer. Alternative forms are available for weekly administration (ie, 1-week recall period) and different modes of administration. In addition to self-administration, a trained interviewer can administer the RAND-36 by telephone or in person. McHorney and co-workers (3) found that the telephone mode of administration was more expensive, had a lower response rate, produced fewer missing item responses, and led to more positive reports of health on the RAND-36 than did the mail mode of administration. An internet version of the RAND-36 is also available (4, 5).

The 36 items are distributed by RAND as the RAND 36-Item Health Survey 1.0 (6, 7), by the Psychological Corporation as the RAND-36 Health Status Inventory (8), by the Medical Outcomes Trust as the SF-36 health survey (9, 10), and by the Health Outcomes Institute as the Health Status Questionnaire (see Appendix for RAND-36 items).

RAND-36 scales and two underlying dimensions

The most common scoring approach for the RAND-36 items boils down to transforming every item linearly to a 0–100 possible range (per cent of total possible score) and then averaging all items in the same scale together. However, the Medical Outcomes Trust and the Health Outcomes Institute versions of the RAND-36 have minor deviations in scoring for one or both of the pain and general health scales. Hays and co-workers (6) showed that these differences have minimal effects on scale scores. Nonetheless, the Medical Outcomes Trust scoring increases the correlation between the two pain items because one item is scored conditional on the other. As a result, the internal consistency reliability estimate based on this scoring is biased upward. The Psychological Corporation (RAND-36) version incorporates item response theory (IRT) scoring methods. Correlations between simply-summed and IRT scores for the RAND-36 are substantial (8), but IRT scoring is designed to produce equal interval measures (11).

Different distributors also score physical and mental health summary scores differently. Factor analyses of the RAND-36 health survey in the US provide strong support for a 2-factor model of health, with physical health reflected primarily by measures of physical

Key messages

- The RAND-36 (SF-36) is the most widely used measure of health-related quality of life today.
- The RAND-36 is a profile measure that yields eight scale scores and two summary scores (physical and mental health).
- A preference-based score has recently been developed for the RAND-36.

functioning, pain and role limitations arising from physical health problems, and mental health reflected primarily by measures of emotional well-being and role limitations caused by emotional problems (8, 12). General health perceptions, energy/fatigue and social functioning reflect both dimensions.

The physical (PCS) and mental (MCS) component summary scales derived for the RAND-36 by Ware, Kosinski, Bayliss, and co-workers (13) were forced to be uncorrelated (orthogonal) whereas those derived by Hays, Sherbourne, and Mazel (6) were allowed to correlate (oblique). Product-moment correlations between physical and mental health factors at each of 3 years (baseline, 2-years post-baseline, and 4-years post-baseline) in the MOS ranged from 0.32 to 0.41 (12). Similarly, a correlation of 0.53 between physical and mental health factors was reported in a study of 1053 older individuals (average age 64 years) sampled from an academic general medical clinic (14). In addition, Hays, Prince-Embury and Chen (8) reported a correlation of 0.66 between physical and mental health factors in a sample of 255 females and 245 males stratified by age, race/ethnicity and educational level to reflect the US population.

Orthogonal (uncorrelated) factor rotations yield a less realistic representation of the physical and mental health factors than do oblique rotations (15). In fact, the PCS and MCS scores resulting from orthogonal factor rotation can yield counterintuitive results. For example, a study of 536 primary care patients who initiated antidepressant treatment found that the RAND-36 physical functioning, role limitations caused by physical health, pain and general health perceptions scales improved significantly by 0.28–0.49 SD units, but the PCS did not change significantly (16). Similarly, Nortvedt and co-workers (17) reported in a study of 194 patients with multiple sclerosis large decrements in emotional well-being (0.3 SD), role limitations caused by emotional problems (0.7 SD), energy/fatigue (1.0 SD), and social functioning (1.0 SD) relative to the general population, but the MCS was only 0.20 SD lower. Both of these anomalies occurred because RAND-36 mental health scales

receive negative weightings on the PCS, and RAND-36 physical health scales receive negative weightings on the MCS. Thus, high mental health scale scores tend to lower the PCS and high physical health scores tend to lower the MCS scores.

Example applications of the RAND-36

Generic profile HRQoL measures, such as the RAND-36, have been used to compare the relative burden of disease for patients with different chronic illnesses. For example, the HRQoL of 2864 HIV-infected adults participating in the HIV Cost and Services Utilization Study, a probability sample of adults with HIV receiving health care in the United States, were recently compared with that of patients with other chronic diseases and to the general US population (18). RAND-36 physical functioning scores were about the same for adults with asymptomatic HIV disease as for the US population but were much worse for those with symptomatic HIV disease or for those who met criteria for the AIDS. Patients with AIDS had worse physical functioning than those with some of the other chronic diseases (epilepsy, gastroesophageal reflux disease, clinically localized prostate cancer, clinical depression, diabetes). RAND-36 emotional well-being was comparable among patients with various stages of HIV disease, but was significantly worse than among the general population and patients

with other chronic diseases with the exception of depression (Fig 1).

The RAND-36 has also been used to assess HRQoL between patients with different diseases over time. For example, patients with hypertension, diabetes and depression (major depression, subthreshold depression) in the MOS were compared at baseline and 2-years post-baseline (19). As expected, patients with depression at baseline had substantially worse emotional well-being than patients with the chronic medical illnesses (hypertension, diabetes) at baseline of the study. Would these differences persist over time? Figure 2 shows the course of emotional well-being from baseline to 2 years later. This picture indicates a relatively flat line for the chronic medical illnesses and positive gains in emotional well-being over time for the patients who were depressed at baseline, especially those with major depression. Nonetheless, patients depressed at baseline continued to demonstrate relatively poor emotional well-being 2 years later.

Another longitudinal application of the RAND-36 was illustrated in a study of the associations of physical activity and exercise with HRQoL in the MOS sample (20). For illustrative purposes (Fig 3), MOS patients with chronic medical illness (hypertension, diabetes) or depression were classified at baseline into those spending little or no time exercising (low) vs those spending a lot of time exercising (high). Those who exercised more had better physical func-

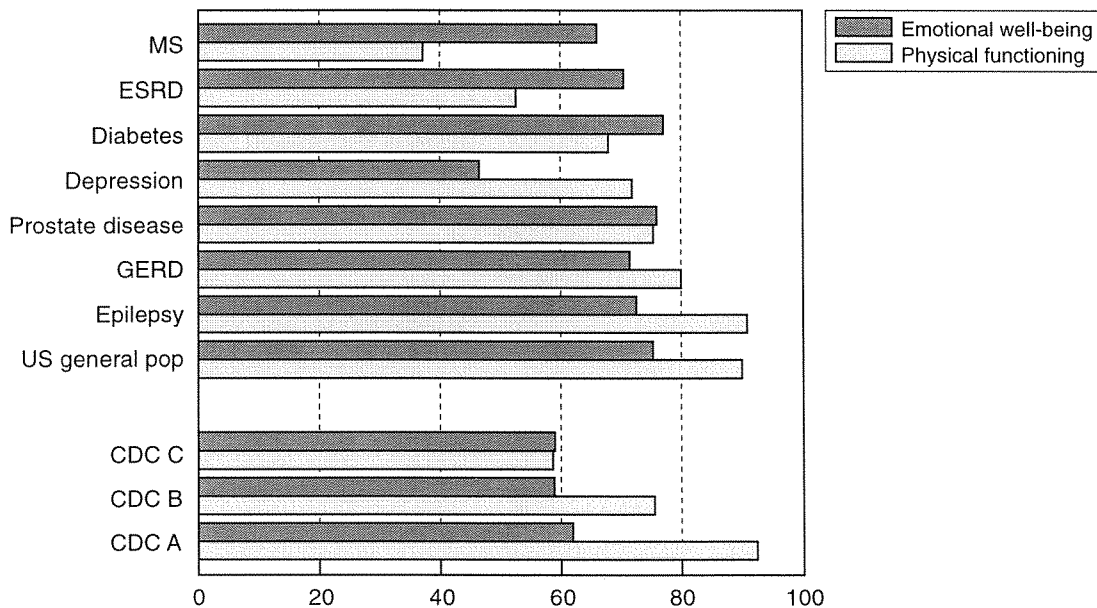


Figure 1. Comparison of RAND-36 emotional well-being and physical functioning scales for 2864 people with HIV infection, patients with chronic disease, and the general US population. MS, multiple sclerosis ($n = 178$); ESRD, end stage renal disease ($n = 165$); diabetes, type II diabetes ($n = 541$); depression, clinical depression ($n = 502$); prostate disease, prostate cancer ($n = 98$); GERD, gastroesophageal reflux disease ($n = 516$); US General Pop., US general population ($n = 2474$); CDC C, AIDS ($n = 1126$); CDC B, symptomatic HIV ($n = 1495$); CDC A, asymptomatic HIV ($n = 243$).

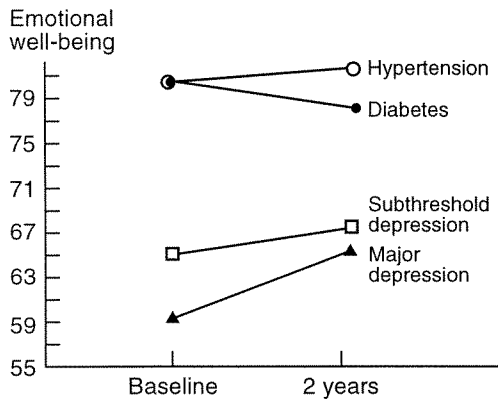


Figure 2. Course of emotional well-being over 2 years for patients in the Medical Outcomes Study.

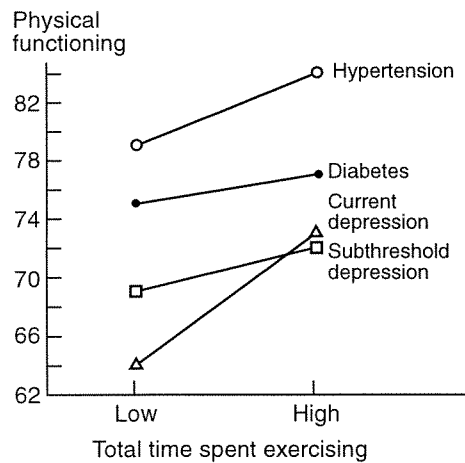


Figure 3. Association of baseline exercise with physical functioning 2 years after baseline in the Medical Outcomes Study.

tioning 2 years post-baseline than those who exercised less, suggesting a possible beneficial effect of exercise.

Minimally clinically important difference

It is generally believed that small differences in HRQoL may be statistically significant yet unimportant. The concept of minimally clinically important difference (MCID) has been proposed to refer to the smallest difference in a score that is considered to be worthwhile or important. Samsa and co-workers (21) reviewed the existing literature and concluded that the MCID for the RAND-36 is ‘typically in the range of 3 to 5 points’ (p. 149).

The literature on the RAND-36 shows that very small differences on the survey could be interpreted as clinically important. The 3–5 point difference in RAND-36 scale scores noted by Samsa and co-workers (21) as the MCID translates into a 0.09–0.28

effect size range. Consistent with this, persons with mild asthma scored 0.86 points (0.09 SD) below the average US adult on the PCS (10). Low back therapy was associated with a 1.3-point (0.13 SD) improvement on the MCS (22). Intermittent treatment for duodenal ulcer has been shown to be associated with an improvement of 3.2 points (0.32 SD) on the PCS (23). All of these changes are arguably clinically important.

The quest to identify a clinically meaningful difference in HRQoL research is part of a more general goal of providing familiar anchors to unfamiliar units to aid interpretation. This is a worthwhile endeavour that can help researchers and clinicians better understand the measures they are using. However, the identification of the MCID is often the ‘most contentious issue discussed by grant review panels.... In most cases, the opinions of both the researchers and reviewers are arbitrary and unsupported by evidence’ (24).

Absolute MCID thresholds are suspect, because they ignore the cost or resources required to produce a change in HRQoL. Just as cost-effectiveness and cost-utility are more comprehensive approaches than examination of outcomes alone, it is also true that the worth of a HRQoL score change depends on the cost to produce it. A small positive change has greater worth by definition if achieving that improvement is possible at lower cost (worth = ratio of incremental improvement to incremental cost). In the extreme case, a therapy that provides an improvement in HRQoL that does not exceed the MCID might still be worthwhile if it were cheap enough or free (25). Hence, we recommend appropriate caution in interpreting 3–5 points on the RAND-36 scales as the MCID.

Summary scores for the RAND-36

If a study shows improvement in some RAND-36 measures and decrements in others, it can be difficult to draw an overall conclusion. For example, Figure 4 illustrates a situation where a new treatment (x) looks better than standard care (o) on physical functioning, but a little worse on pain and emotional well-being, and there is no difference on social functioning. Is the new treatment better than standard care? To make concluding statements, it is necessary to have some method of summarizing the multiple scale scores.

A variety of approaches have been used to help summarize overall HRQoL impact represented in profile measures. As noted above, physical and mental health summary scores have been derived for the RAND-36. Reliability estimates for summary measures have exceeded 0.90 (8). Extensive support for the construct validity of these summary scores has also

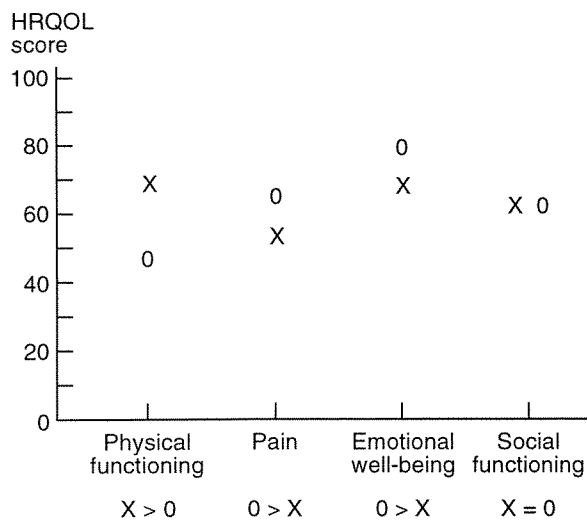


Figure 4. Is new treatment (X) better than standard care (O)?

been provided. For example, the RAND-36 PCS has been demonstrated to be strongly predictive of mortality rates 5 years later (10). Because the PCS and MCS scores were derived from a model that is inconsistent with observed data (ie, uncorrelated factors) and the potential for anomalous results discussed above, it is highly recommended that users employ the summary scores derived by the Psychological Corporation or derive scores by using an oblique factor analysis model.

Regression-weighting methods have been employed to derive a single summary score. Weights are derived by regressing a criterion measure (eg, current health perceptions, visual analogue rating, probability of being in excellent, very good or good health in the future, probability of being dead in the future) on HRQoL scale scores. These weights reflect the relative importance of the scales in predicting the criterion (26).

Attrition of participants in a study over time because of mortality creates a noteworthy problem in applying profile measures. If those who die are dropped from the analysis, results can be biased. Factor analytic and regression-weighting approaches are useful summary methods, but preference-based methods are more desirable because they explicitly combine morbidity and mortality (27). Preference-based measures are designed to assess the value or desirability of health states. 'It is important to know what values people attach to different health outcomes in order to provide, as efficiently as possible, more of the outcomes that are desired and fewer of those that are not' (28). Thus, it has been recommended that the RAND-36 be supplemented with a preference measure whenever possible (27).

If a preference measure has not been administered,

it is possible to estimate one. For example, a study of 363 community-dwelling older persons found that comprehensive geriatric assessment lead to improvements in RAND-36 physical functioning scores by 4.69 relative to standard care (29). The estimated 64-week intervention costs were US\$ 457 per person. Thus, the cost-effectiveness of the intervention was estimated to be US\$97 per SF-36 physical functioning scale point. Is this an intervention that is worth paying for or not? To help answer this question, Keeler and co-workers (30) estimated the change in quality-adjusted life years (QALYs) equivalent to the 4.69 change in SF-36 physical functioning score. They did this by translating a 0.69 correlation between the SF-36 physical functioning score and the Quality of Well-Being (QWB) scale into a simple regression coefficient (0.003) and multiplying this coefficient by 4.69 to obtain 0.014 as the equivalent change in the QWB. Next, they estimated an increase of 0.07 QALYs over 5 years and a cost per QALY between US\$ 10 600 and 26 500. This is useful information because previously published data indicate that many common medical interventions cost from US\$ 10 000 to 40 000 per QALY, with interventions producing QALYs for less than US\$ 20 000 regarded as worthwhile (31). If Keeler and co-workers had had a preference-based score, they would not have been forced to rely on this crude approach to estimating QALYs.

The use of preference measures, such as the standard gamble and time trade-off, as criteria for regression weighting has also been considered, but profile measures tend to account for a relatively small amount of variance (18–43%) in these criteria (32). Fryback and co-workers (33) derived regression equations to predict the QWB Scale from the RAND-36. A 6-variable regression model accounted for 57% of the variance in QWB scores.

One investigative team derived a 6-dimensional health classification scheme by using a subset of items and categories from the RAND-36 physical functioning, social functioning, pain, emotional well-being, energy/fatigue, and role functioning scales (34, 35). Multiattribute utility theory was used to derive preference weights for 9000 possible health states. Visual analogue and standard gamble estimating equations were developed to predict preference scores for each possible health state. The 166 participants in the valuation study consisted of health professionals, health service managers and administrators, staff at the University of Sheffield Medical School, undergraduates, and patients at hospital outpatient clinics. A larger study of 611 people from the UK general population has recently been completed by this group of investigators.

Another investigative team identified six health states from the SF-12 in depressed patients in the

MOS (36). The six scales were derived by using cluster analysis in combination with clinical criteria. For each cluster, health state descriptions were written based on SF-12 item responses given by at least half the patients in the cluster. Preference ratings were then derived for the six health states (37).

Availability of RAND-36 in multiple languages

Although the RAND-36 (aka SF-36) has been used in multiple languages, there have been few systematic efforts to produce equivalent translations for use across cultural and linguistic groups. The most important of these efforts is the International Quality of Life Assessment (IQOLA) Project. The IQOLA project is translating and validating the SF-36 for use in 45 countries, including Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, the UK (English version), and the USA (English and Spanish versions). The purpose of the IQOLA Project is to produce translations of the SF-36 for use in multinational clinical trials and other international studies (38).

The translation protocol used by the IQOLA Project has four major stages: translation following a standard protocol including multiple forward and backward translations; qualitative and quantitative methods to evaluate the quality of a translation and its conceptual equivalence with the original survey; psychometric studies to test the scaling and scoring assumptions; and analysis of data from clinical trials and other studies to assess validity and comparability of the survey data across countries (39). In addition, normative data are being collected in general population surveys in 11 countries for purposes of norm-based interpretation.

Published IQOLA Project translations and English-language adaptations are distributed by the New England Medical Center (NEMC) Health Assessment Lab. (Boston, MA) Currently, published forms are available for Australia/New Zealand (English), Belgium (Dutch and French), Canada (English and French), Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, and the UK (English) (40).

References

1. Reuben DB, Valle LA, Hays RD, Siu, AL. Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. *J Am Geriatr Soc* 1995; 43: 17–23.
2. Stewart AL, Sherbourne CD, Hays RD, Wells KB, Nelson

SF-36 Version 2

The SF-36 Version 2 differs from the original survey instrument (version 1) in several respects including: wording changes in instructions and survey items; changes in the layout for questions and response options (41); 5-level response choices in place of dichotomous response choices for seven items in the two role-functioning scales; and 5-level in place of 6-level response choices to nine items in the emotional well-being and energy/fatigue scales. These changes are purported to increase the precision of the scores generated by the SF-36 and to improve respondents' comprehension of the instrument. However, unpublished analyses of existing datasets suggest a possible lack of equivalence between the old and new versions of the energy/fatigue items. The original 'full of pep' item was replaced with a 'full of life' item in the second version of SF-36. Product-moment correlations of 0.61 and 0.64 were found between these two items in samples of people in a smoking cessation programme and people with chronic liver disease, respectively (42, 43).

A detailed description of the differences between versions 1 and 2 is available on the internet (44). Although a conversion formula between the two versions of the SF-36 is available (44), studies to further evaluate the equivalence of two versions are currently ongoing.

Conclusion

The RAND-36 is a comprehensive short-form generic profile HRQoL measure. It yields eight scale and two summary scores. When a single summary score is required (eg, cost-utility studies) it is most desirable to use the RAND-36 in combination with a standard preference-based measure (45). When a preference-based measure has not been administered, it is possible to estimate such a score from the RAND-36, but the validity of the existing estimates is subject to question.

The support of the Yrjö Jahnsson Foundation is greatly appreciated. ©1986, 1992, 2001 by RAND.

- EC, Kamberg C J, et al. Summary and discussion of MOS measures. In: Stewart A L, Ware J E, eds. *Measuring functioning and well-being: The Medical Outcomes Study approach*. Durham, NC: Duke University Press; 1992: 345–71.
3. McHorney CA, Kosinski M, Ware JE. Comparisons of the

- costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: results from a national survey. *Med Care* 1994; 32: 551-67.
4. Bell DS, Kahn CD. Health status assessment via the World Wide Web. *Proc AMIA Annu Fall Symp* 1996; 338-42.
 5. Your Health [Online]. The medical informatics and decision science consortium. Available at: <http://www.mcw.edu/midas/health/> (2000, Sept 4).
 6. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-item Health Survey 1.0. *Health Econ* 1993; 2: 217-27.
 7. RAND-36 Item Health Survey 1.0 [Online]. RAND Health Program. Available at: <http://www.rand.org/organization/health/healthpubnav.html> (2000, Sept 5).
 8. Hays RD, Prince-Embury S, Chen H. *RAND-36 health status inventory*. San Antonio, TX: The Psychological Corporation; 1998.
 9. Ware JE, Sherbourne C. The MOS 36-item short form health survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-83.
 10. Ware JE, Kosinski M, Keller SD. *SF-36 Physical and mental health summary scales: a user's manual*. Boston, MA: The Health Institute: New England Medical Center; 1994.
 11. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000; 38: II-28-II-42.
 12. Hays RD, Marshall GN, Wang EYI, Sherbourne CD. Four-year cross-lagged associations between physical and mental health in the Medical Outcomes Study. *J Consult Clin Psychol* 1994; 62: 441-9.
 13. Ware JE, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparisons of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care* 1995; 33: AS264-79.
 14. Dexter PR, Stump TE, Tierney WM, Wolinsky FD. The psychometric properties of the SF-36 health survey among older adults in a clinical setting. *J Clin Geropsychol* 1996; 2: 225-37.
 15. Rummel RJ. *Applied factor analysis*. Evanston, IL: Northwestern University Press; 1970.
 16. Simon GE, Revicki DA, Grothaus L, Vonkorff M. SF-36 summary scores: are physical and mental health truly distinct. *Med Care* 1998; 36: 567-572.
 17. Nortvedt M, Riise T, Myhr K, Nyland H. Performance of the SF-36, SF-12 and RAND-36 summary scales in a multiple sclerosis population. *Med Care* 2000; 38: 1022-8.
 18. Hays RD, Cunningham WE, Sherbourne CD, Wilson IB, Wu AW, Cleary PD, et al. Health-related quality of life in patients with human immunodeficiency virus infection in the United States: results from the HIV Cost and Services Utilization Study. *Am J Med* 2000; 108: 714-22.
 19. Hays RD, Wells KB, Sherbourne CB, Rogers WH, Spritzer K. Functioning and well-being outcomes of patients with depression compared to chronic general medical illness. *Arch Gen Psychiatry* 1995; 52: 11-19.
 20. Stewart AL, Hays RD, Wells KB, Rogers WH, Spritzer KL, Greenfield S. Long-term functioning and well-being outcomes associated with physical activity and exercise in patients with chronic conditions in the Medical Outcomes Study. *J Clin Epidemiol* 1994; 47: 719-30.
 21. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 1999; 15: 141-55.
 22. Lansky D, Butler JB, Waller FT. Using health status measures in the hospital setting: from acute care to 'outcomes management.' *Med Care* 1992; 30: MS57-73.
 23. Rampal P, Martin C, Marquis P, Ware JE, Bonfils S. A quality of life study in five hundred and eighty-one duodenal ulcer patients: maintenance versus intermittent treatment with nizatidine. *Scand J Gastroenterol* 1994; 206: 44-51.
 24. van Walraven C, Mahon JL, Moher D, Bohm C, Laupacis A. Surveying physicians to determine the minimal important difference: implications for sample-size calculation. *J Clin Epidemiol* 1999; 52: 717-23.
 25. Hays RD, Woolley JM. How meaningful is the concept of clinically meaningful difference in health-related quality of life research? *Pharmacoeconomics* 2000; 18: 419-23.
 26. Bozzette SA, Hays RD, Berry S, Kanouse D. A perceived health index for use in persons with advanced HIV disease: derivation, reliability, and validity. *Med Care* 1994; 32: 716-31.
 27. Hays RD, Alonso J, Coons SJ. Possibilities for summarizing health-related quality of life when using a profile instrument. In: Staquet M, Hays R, Fayers P, eds. *Quality of life assessment in clinical trials: methods and practice*. Oxford: Oxford University Press; 1998: 143-53.
 28. Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet ND, et al. Identifying and valuing outcomes. In: Gold M R, Siegel J E, Russell L B, et al, eds. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996: 82-134.
 29. Reuben DB, Frank JC, Hirsch SH, McGuigan KA, Maly RC. A randomized clinical trial of outpatient comprehensive geriatric assessment coupled with an intervention to increase adherence to recommendations. *J Am Geriatr Soc* 1999; 47: 269-76.
 30. Keeler EB, Robalino DA, Frank JC, Hirsch SH, Maly RC, Reuben DB. Cost-effectiveness of outpatient geriatric assessment with an intervention to increase adherence. *Med Care* 1999; 37: 1199-206.
 31. Tengs TO, Adams ME, Pliskin JS, Safran DG, Siegel JE, Weinstein MC, et al. Five hundred life-saving interventions and their cost-effectiveness. *Risk Analysis* 1995; 15: 369-90.
 32. Revicki DA, Kaplan RM. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Qual Life Res* 1993; 2: 477-87.
 33. Fryback DG, Lawrence WF, Martin PA, Klein R, Klein BEK. Predicting Quality of Well-Being scores from the SF-36: results from the Beaver Dam Health Outcomes Study. *Med Decis Making* 1997; 17: 1-9.
 34. Brazier J. The Short-Form 36 (SF-36) Health Survey and its use in pharmacoeconomic evaluation. *Pharmacoeconomics* 1995; 7: 403-15.
 35. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 health survey. *J Clin Epidemiol* 1998; 51: 1115-28.
 36. Sugar CA, Sturm R, Lee TT, Sherbourne CD, Olshen RA, Wells KB, et al. Empirically defined health states for depression from the SF-12. *Health Services Res* 1998; 33: 911-28.
 37. Lenert LA, Sherbourne CD, Sugar C, Wells KB. Estimation of utilities for the effects of depression from the SF-12. *Med Care* 2000; 38: 763-70.
 38. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998; 51: 903-12.
 39. Bullinger M, Alonso J, Apolone G, Leplege A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol* 1998; 51: 913-23.
 40. SF-36 Health Survey [Online]. Health Assessment Lab. Available at: <http://www.sf36.com> (2000, Sept 4).
 41. Mullin PA, Lohr KN, Bresnahan BW, McNulty P. Applying cognitive design principles to formatting HRQOL instruments. *Qual Life Res* 2000; 9: 13-27.

42. Olufade AO, Shaw JW, Foster SA, Leischow SJ, Andrejasich CM, Hays RD, et al. Development of the smoking cessation quality of life (SCQoL) questionnaire. *Clin Therapeutics* 1999; 21: 2113-30.

43. Gralnek IM, Hays RD, Kim S, Rosen HR, Keeffe EB, Artinian L, et al. Development and evaluation of the Liver Disease Quality of Life instrument in patients with advanced, chronic liver disease - LDQOL 1.0. *Am J Gastroenterol*

2000; 95: 3552-65.

44. Download and Demo SF-36 (Online). Quality Metric Inc. Available at: <http://www.qmetric.com/products/assessments/sf36/sf36v2.php3/> (2000, Sept 7).

45. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. *Pharmacoeconomics* 2000; 17: 13-35.

Appendix: RAND-36 items

Your Health
 This survey includes a wide variety of questions about your health and your life. We are interested in how you feel about each of these issues.

1. In general, would you say your health is: [Mark an in the one box that best describes your answer.]

Excellent Very good Good Fair Poor
 1 2 3 4 5

2. Compared to one year ago, how would you rate your health in general now?

Much better now than one year ago Somewhat better now than one year ago About the same as one year ago Somewhat worse now than one year ago Much worse now than one year ago
 1 2 3 4 5

3. The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? [Mark an in a box on each line.]

Yes, limited a lot Yes, limited a little No, not limited at all

a Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports 1 2 3

b Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf 1 2 3

c Lifting or carrying groceries 1 2 3

d Climbing several flights of stairs 1 2 3

e Climbing one flight of stairs 1 2 3

f Bending, kneeling, or stooping 1 2 3

g Walking more than a mile 1 2 3

h Walking several blocks 1 2 3

i Walking one block 1 2 3

j Bathing or dressing yourself 1 2 3

4. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

a Cut down the amount of time you spent on work or other activities Yes 1 No 2

b Accomplished less than you would like Yes 1 No 2

c Were limited in the kind of work or other activities Yes 1 No 2

d Had difficulty performing the work or other activities (for example, it took extra effort) Yes 1 No 2

5. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

a Cut down the amount of time you spent on work or other activities Yes 1 No 2

b Accomplished less than you would like Yes 1 No 2

c Didn't do work or other activities as carefully as usual Yes 1 No 2

6. During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

Not at all Slightly Moderately Quite a bit Extremely
 1 2 3 4 5

7. How much bodily pain have you had during the past 4 weeks?

None Very mild Mild Moderate Severe Very severe
 1 2 3 4 5 6

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all A little bit Moderately Quite a bit Extremely
 1 2 3 4 5

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks...

All of the time Most of the time A good bit of the time Some of the time A little of the time None of the time

a Did you feel full of pep? 1 2 3 4 5 6

b Have you been a very nervous person? 1 2 3 4 5 6

c Have you felt so down in the dumps that nothing could cheer you up? 1 2 3 4 5 6

d Have you felt calm and peaceful? 1 2 3 4 5 6

e Did you have a lot of energy? 1 2 3 4 5 6

f Have you felt downhearted and blue? 1 2 3 4 5 6

g Did you feel worn out? 1 2 3 4 5 6

h Have you been a happy person? 1 2 3 4 5 6

i Did you feel tired? 1 2 3 4 5 6

10. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

All of the time Most of the time Some of the time A little of the time None of the time
 1 2 3 4 5

11. Please choose the answer that best describes how true or false each of the following statements is for you.

Definitely true Mostly true Don't know Mostly false Definitely false

a I seem to get sick a little easier than other people 1 2 3 4 5

b I am as healthy as anybody I know 1 2 3 4 5

c I expect my health to get worse 1 2 3 4 5

d My health is excellent 1 2 3 4 5

Thank you for completing these questions!