The Significance of Importance:
An Evaluation of Ferrans and Powers' Quality of Life Index

Peter Hagell[1,2,3], Albert Westergren[1,3,4]

[1] Division of Gerontology and Caring Sciences, Department of Health Sciences, Lund University, Lund, Sweden
[2] Department of Neurology, University Hospital, Lund, Sweden
[3] The Vårdal Institute, the Swedish Institute for Health Science, Lund University, Lund, Sweden
[4] Department of Rehabilitation Medicine, Central Hospital Kristianstad, Northeast Skåne Health Care District, Kristianstad, Sweden

Running head: The Significance of Importance

Corresponding address:
Peter Hagell, RN PhD
Department of Health Sciences
P.O. Box 157
Lund University
SE-221 00 Lund
Sweden
Tel: +46 46 222 1937
E-mail: Peter.Hagell@med.lu.se

ABSTRACT

Ferrans and Powers' Quality of Life Index (QLI) defines and assesses quality of life (QoL) in terms of importance-weighted life satisfaction. This study assessed the value of such weights and explored the relationship between weighted and unweighted (satisfaction only) scores and single-item rated overall life satisfaction (LS) and QoL. Data were collected by a postal survey to 81 Parkinson's disease patients (88% response rate). Correlations between weighted and unweighted QLI scores were ≥0.96, except for one subscale (r_s = 0.85). Item non-response rates ranged between 4.2-45.1% and 1.4-38% for the weighted and unweighted QLI, respectively. Cronbach's alpha exceeded 0.7 for weighted and unweighted versions of two out of the four subscales and the total score. Scaling success rates were similar for weighted and unweighted scores and did not support the current subscale structure. Unexpectedly, weighted total scores correlated stronger with LS than with QoL, and unweighted scores displayed the opposite pattern. This study found no advantages by using importance-weighted satisfaction scores. The correlational pattern with overall LS and QoL challenges the QLI approach to QoL, although these observations may relate to the use of multiplicative item weights. This study has implications also beyond the QLI regarding, e.g., the use of multiplicative weights and the relationship between life satisfaction and QoL.

Key words: Life satisfaction; quality of life; scaling; weights

Abbreviations:

FA – Family subscale of the QLI; HF - Health and functioning subscale of the QLI; LS - Life satisfaction; NHP - Nottingham health profile; PD - Parkinson's disease; PDQ-39 - 39-item Parkinson's disease questionnaire; PS - Psychological-spiritual subscale of the QLI; QLI - Quality of life index; QoL - Quality of life; SE - Social and economic subscale of the QLI

INTRODUCTION

Stemming from work in the social and behavioral sciences, several investigators have defined quality of life (QoL) in terms of life satisfaction [1-9]. Life satisfaction, in turn, has been defined in terms of the perceived discrepancy between aspiration and achievement, i.e., an assessment of the current situation against some standard of comparison, or the degree to which an individual experiences him-/herself as being able to attain his/her goals [e.g., 10, 11]. While a number of instruments have been devised based on these theoretical underpinnings, there appears to be a lack of empirical assessments of the relationship between how ill people rate their QoL and how satisfied they are with their lives. Due to general lack of consensus and individual differences regarding what contributes to QoL, it has been argued that satisfaction with various areas of life should be interpreted in view of their perceived importance [5-8, 12]. One questionnaire that takes these considerations into account is Ferrans and Powers' Quality of Life Index (QLI), which measures QoL in terms of importance-weighted satisfaction with various areas of life [4, 13]. While this approach may appear theoretically appealing, it is not clear to what extent satisfaction ratings are influenced by weighting them by their perceived importance. For example, using the QLI in people with spinal cord injury, May & Warren [14, 15] found close to perfect correlations between importance-weighted and unweighted life satisfaction scores. This indicates that the weighting procedure provides little gain in terms of the resulting scores. However, it is unclear whether these observations extend to other patient groups and if weighted and unweighted scores differ in terms of their psychometric properties.

Here we report results from a study using the QLI in patients with Parkinson's disease (PD). The objectives were two-fold. First, to assess the value of weighting ratings of satisfaction with various areas of life by their perceived importance, regarding (a) the resulting scores, and (b) their psychometric performance. Secondly, to explore the

relationship between weighted and unweighted QLI scores and overall life satisfaction and QoL. Strong correlations were expected between total QLI scores and life satisfaction as well as QoL. It was further expected that in order to support the assumptions underlying the QLI definition of QoL and its scoring system (i.e., that QoL is determined by satisfaction with various areas of life, as weighted by their importance to the individual), importance-weighted QLI scores were expected to correlate stronger with overall QoL than with overall life satisfaction, and unweighted satisfaction QLI scores should display the opposite pattern.

METHODS

Data were collected by a cross-sectional postal survey for which every fifth patient (n = 81) was selected from a total of 407 patients with clinically diagnosed PD receiving neurological care at Lund University Hospital (Sweden) during one year. To avoid a possible age bias, patients were organized chronologically according to age prior to selection. In addition to the QLI, the survey comprised the 39-item PD Questionnaire (PDQ-39) and the Nottingham Health Profile (NHP). Results regarding the PDQ-39 and NHP have been reported previously [16] and will not be considered here. No reminders were sent out. Response was interpreted as consent to participate. The study was approved by the Research Ethics Committee, Faculty of Medicine, Lund University, Sweden.

Questionnaires

Ferrans and Powers' QLI [4, 13] was developed based on the operational definition of QoL as ""a person's sense of well-being that stems from satisfaction or dissatisfaction with the areas of life that are important to him/her"" [13, p. 29]. In completing the QLI respondents are thus requested to indicate how satisfied they are with 34 areas of life as well as how important they consider each one of them to be. Ratings are made on a 1-6

scale ranging from very dissatisfied/unimportant to very satisfied/important. The instrument yields a total score and four domain scores: health and functioning (HF; 14 items), social and economic (SE; 9 items), psychological-spiritual (PS; 7 items), and family (FA; 4 items). Satisfaction responses are weighted by their paired importance ratings by multiplication. Paired multiplicative satisfaction-importance item scores are then used to calculate total and domain scores (see the official QLI web site for complete scoring algorithm, http://www.uic.edu/orgs/qli/index.htm) that range between 0 and 30 (30 = better life satisfaction/QoL). In this study, the Swedish version [17] of the QLI was used. Previous psychometric information on the Swedish QLI is sparse, but coefficient alpha has ranged between 0.84 and 0.94 [17-19]. An exploratory factor analysis based on responses from 90 mobility disabled elderly people produced a somewhat different factor structure to that by Ferrans & Powers [13], but due to the small sample size the original subscale structure was retained [17, 19].

In order to explore the relationship between weighted and unweighted QLI scores and overall life satisfaction and QoL, patients rated their overall life satisfaction (LS) on a 6-grade single-item Likert scale previously used for validity assessment of the QLI [20]. Patients also rated their overall QoL by placing a mark on a 100-mm horizontal visual analogue scale (VAS) anchored by “worst imaginable quality of life” at the far left (0 mm), and “best imaginable quality of life” at the far right (100 mm) [21].

In addition, data regarding patients’ perceived PD severity (rated as “mild”, “moderate” or “severe”), disease duration and number of daily PD medicine intake occasions were collected.

Analyses

Unweighted QLI scores were calculated by rescoring (from 1-6 to 0-5) and summing raw item responses. To ease interpretation, weighted and unweighted QLI, as well as overall LS scores, were transformed to a possible range between 0 and 100 (observed score/maximum score x 100), where higher scores indicate better life satisfaction.

Psychometric performance was evaluated by means of traditional test theory methodology [22, 23]. Three properties were considered and compared between weighted and unweighted (satisfaction only) QLI scores: data quality, reliability, and scaling success. Data quality relates to the usefulness of an instrument, and is considered high when missing data are few. Data quality was examined in terms of the overall scale and individual item percentages of missing responses, as well as the proportion of respondents with missing item responses, within each scale. Reliability was assessed by means of coefficient alpha, a measure of item interrelatedness and an estimate of reliability that preferably should be around 0.80, and not below 0.70 [24, 25].

Scaling success examines the extent to which items appear to be appropriately grouped into subscales. Support for appropriate grouping of items is obtained if the corrected correlation between an item and its hypothesized subscale score is significantly stronger than that with other subscales. The corrected item-to-total correlation is the correlation between an item and the total score of the other items in its hypothesized subscale [26]. Scaling success was calculated as a percentage indicating how often items within a subscale correlate significantly stronger ($\geq$+2 standard errors, i.e., approximately the upper limit of a 95% confidence interval) with their hypothesized scale (item-convergent validity) than with other scales (item-discriminant validity) [23], which was defined as definite scaling success. Instances when items correlated non-significantly stronger (<+2 standard errors) with their hypothesized scale than with other scales were considered possible scaling success, and scaling failure was defined as instances when items correlated weaker with their hypothesized

scale than with other scales. For total QLI scores, where scaling success is not appropriate, corrected item-to-total correlations were considered instead. According to various recommendations, acceptable thresholds for corrected item-to-total correlations range between 0.20 to 0.40 [22, 25]. In addition, correlations between subscales were considered in relation to coefficient alpha values for the respective domains. Because coefficient alpha is a measure of the average item interrelatedness and all possible split-half correlations [24, 25], it can be viewed as the "correlation between a scale and itself" [23]. To the extent that inter-scale correlations are lower than their respective alpha values, there is thus support that they measure distinct constructs with unique and reliable variance [23].

The relationship between weighted and unweighted QLI scores and overall life satisfaction and QoL was explored by determining the correlations between weighted and unweighted total QLI scores and single-item overall LS and QoL ratings.

Variables were checked regarding assumptions underlying parametric and non-parametric statistics, and analyzed accordingly. All analyses were performed using SPSS 12.0.1 for Windows (SPSS Inc., Chicago, IL).

RESULTS

Seventy-one (88%) of the questionnaire packages were returned. Sixty-two percent of responders were males. Responders' mean (SD) age was 69.1 (10.2) years, the mean (SD) duration of their PD was 10.4 (9.4) years, and the mean (SD) number of PD medicine intakes per day was 5.6 (2.7). Non-responders were older (76.8 ± 6 years old) than respondents ($p = 0.027$; unpaired t-test). A majority of patients (n=35) perceived their PD as moderate, and 18 and 17 patients perceived their PD as mild and severe, respectively (one patient did not report her perceived disease severity).

[Table 1 about here]

The mean (SD) nontransformed (0-30) total QLI score was 20.2 (4.3). Mean (SD) nontransformed domain scores were 16.9 (5.9), 22.9 (4.8), 21.0 (4.7) and 25.8 (3.9) for the HF, SE, PS and FA subscales, respectively. Weighted and unweighted transformed (0-100) QLI scores were very similar, with mean and median scores differing by 0.1-1.6 and 0.5-3.7 points, respectively (Table 1). Spearman correlations (r_s) between weighted and unweighted scores were ≥0.96 in all instances but for the SE subscore, where it was 0.85. Figure 1 shows the relationship between weighted and unweighted total QLI scores. As also illustrated in Figure 1, the distribution of unweighted satisfaction and importance scores displayed different patterns. Whereas transformed satisfaction scores ranged between 44.7 and 87.6, importance scores ranged between 70.6 and 98.8. The mean (SD; min-max) raw item score (possible range, 1-6) was 4.38 (1.34; 3.03-5.62) for satisfaction items and 5.23 (0.87; 4.06-5.91) for importance items.

[Figure 1 about here]

There were few substantial differences between weighted and unweighted QLI scores regarding their psychometric performance (Table 2). The overall proportion of missing item responses within the various scales ranged between 5.4% and 19.1% and the proportion for individual items ranged between 1.4% and 45.1%, with lower values for unweighted scores. The proportion of respondents with missing items within the respective scales was >20% for all but the PS and the unweighted FA domains. About one third of respondents (35.2%) returned questionnaires without any missing item responses. No item was answered by all respondents. Reliability was similar for the two QLI score versions (Table 2). Desirable

levels of reliability were observed for weighted and unweighted HF, PS, and total QLI scores. Evaluations of scaling success also yielded similar results for weighted and unweighted scores. Definite scaling success rates did not exceed 50% in more than half of the instances (weighted and unweighted HF, weighted PS, and unweighted FA subscales) and never reached 70% (Table 2). All scale versions but the unweighted FA subscale exhibited some degree of scaling failure. The mean (min, max) corrected item-to-total correlations for weighted and unweighted total QLI scores were 0.37 (-0.18, 0.72) and 0.39 (-0.24, 0.74), respectively. For the weighted total QLI score, 24 (71%) and 19 (56%) items met the suggested item-to-total criteria of 0.20 and 0.40, respectively. For the unweighted score, the corresponding figures were 26 (76%) and 17 (50%), respectively. Coefficient alpha exceeded inter-scale correlations by ≥0.11 (r) and ≥0.14 (r_s) for all domain scores but the weighted SE subscale (alpha, 0.69), which correlated strongly (r_s and r, 0.68) with the weighted PS subscale.

[Table 2 about here]

Median (interquartile range) single-item overall QoL and transformed (0-100) LS scores were 56 (34-78) and 80 (60-80), respectively. Corresponding mean (SD) scores were 53.8 (25.5) and 73.2 (23.8), respectively. Overall QoL ratings correlated non-significantly stronger with unweighted satisfaction scores than with importance-weighted total QLI scores (r_s = 0.692 and 0.576, respectively). Correlations between total unweighted and weighted QLI scores and LS displayed the opposite statistically non-significant pattern (r_s = 0.529 and 0.657, respectively). The correlation (r_s) between single-item overall QoL and LS ratings was 0.477.

DISCUSSION

This study provides an assessment of the relative gain of weighting satisfaction ratings by their paired importance ratings in terms of the resulting scores and their psychometric performance. In addition, the QLI definition of QoL in relation to overall ratings of life satisfaction and QoL among people with PD was addressed. To the best of our knowledge, this is the first evaluation of these aspects of the QLI.

The main study limitation was the relatively small sample size, precluding the use of, e.g., factor analysis. However, it is unlikely that this substantially influenced the main findings. While it is recognized that traditional psychometric test theory estimates are sample dependent, experience indicate that stable interpretations of psychometric indices such as those used here can be obtained with sample sizes of 20 to 40 subjects [27].

Another potential limitation is the concurrent use of multiple questionnaires in a postal survey, which may have influenced item response rates, responses and psychometric performance. However, use of multiple questionnaires is common practice in studies addressing QoL and related issues, and questionnaire batteries similar to that used here have been employed also in previous studies involving the QLI [e.g., 28-32]. Evaluating questionnaire performance under circumstances such as those employed in this study may provide a means of assessing how acceptable instruments are to respondents, since response rates can indicate respondents' understanding and acceptance of questionnaires [22]. The high proportion of missing item responses observed here may therefore indicate that the QLI is perceived as difficult, lacks in saliency, or is inappropriate for self-administration. The somewhat higher item response rates for the unweighted, as compared to the weighted, QLI are expected since the former is not affected by missing importance item responses. The QLI has previously been used in a number of postal surveys, involving people with, e.g., epilepsy, cardiac disease and cancer, including samples of similar ages as those participating in this

study, without any reported problems with low questionnaire response rates [18, 29, 30, 33]. However, item non-response does not appear to be a unique phenomenon for this study. For example, in postal surveys among middle-aged breast cancer and hemodialysis patients, respectively, Ferrans and Powers [13] and Ferrans [20] had to drop 20-30% of their samples from the final analysis of the QLI due to ≥15% missing data. Taken together, there appears to be strong reason to question the appropriateness of the QLI for self-administration.

We found striking similarities between importance-weighted and unweighted (satisfaction only) QLI scores, and the correlations between the two were close to perfect in four instances and very strong in one. One explanation probably relates to the restricted distribution of respondents' importance ratings, i.e., patients tended to assign high levels of importance to most assessed areas of life. These observations are in close accordance with those by May and Warren [14, 15] when using the QLI in people with spinal cord injury. Similarly, in their validation study of the Quality of Life Inventory (another questionnaire where satisfaction with various domains of life is weighted by their paired importance ratings) Frisch et al. [5] reported that respondents typically indicated that they perceived all assessed areas of life as important. In addition to the close similarities between weighted and unweighted QLI scores, neither score version performed psychometrically substantially or consistently better than the other. Experiences from the use of other weighting schemes intended to reflect item severity when deriving scores on health-related outcome measures have also typically failed to demonstrate any advantages as compared to unweighted scores [34-36].

Another important aspect to consider in the use of patient-reported outcomes is the burden they pose upon those requested to complete them, particularly since research protocols often involve multiple questionnaires [37, 38]. Given the similarities in scores and the closely resembling psychometric performance of weighted and unweighted QLI scores,

importance ratings appear to add little but undue respondent burden, which further argues against their use. Two exceptions are, however, obvious. First, when the relative importance of various areas of life is an integral part of study objectives. Second, to the extent that the QLI or similar questionnaires are used in clinical settings, importance ratings may serve as an aid in setting priorities for individual patient interventions. However, regardless of the rationale for assessing importance there are good reasons not to use these ratings as multiplicative satisfaction weights, but to treat them as independent scores instead (see below).

Psychometrically, the QLI performed poorly in this study. Assessments of scaling assumptions did not support the current grouping of items into subscales, or the use of a total QLI score. On all subscales but the unweighted FA version, there were items that correlated stronger with other domains than with their hypothesized subscales, and only in about half of instances was a definite scaling success found. In particular, the validity of the SE subscale can be questioned. As for the total QLI score, up to half of items failed to meet suggested criteria for corrected item-total correlations, and some items displayed negative coefficients, indicating that they do not behave in harmony with the rest. The subscale grouping of QLI items is based on an exploratory factor analysis of responses from 349 American hemodialysis patients [13]. Studies re-assessing the factorial validity of the QLI are few [15, 17, 39] and have been hampered by relatively small sample sizes, rendering them largely inconclusive [40]. However, given that the QLI was conceptualized, developed and grouped into its current subscale structure in the USA, cultural differences cannot be excluded as potential reasons for the observed scaling discrepancies. Hypothetically, such discrepancies could also relate to differences in patient characteristics (e.g., hemodialysis vs PD).

Contrary to expectations based on the rationale behind the QoL definition and scoring algorithm of the QLI, correlations between importance-weighted QLI scores and

overall QoL ratings were weaker than those with overall LS, whereas the opposite was observed for unweighted satisfaction scores. Earlier assessments of the validity of the weighted QLI have typically found correlations around 0.8 with single-item overall life satisfaction ratings [13, 20, 41], which is stronger than the 0.657 coefficient observed here. However, values closer to 0.6 have also been reported [4]. While the strength of the correlation between overall LS and the weighted QLI observed in this study is weaker than that reported in most previous studies, the more intriguing finding is the unexpected (although statistically non-significant) pattern of correlations. These observations could be interpreted in several ways. First, the weaker correlation between overall QoL and the weighted, as compared to the unweighted QLI speaks against the weighting procedure as a means of reflecting subjective QoL. Second, the stronger relationship between weighted QLI scores and overall LS, as compared to overall QoL, indicates that the importance-weighted QLI primarily has validity as a measure of life satisfaction rather than QoL. Third, the stronger relationship between overall QoL and unweighted QLI scores, as compared to weighted QLI scores, indicates that the better (albeit yet suboptimal) QoL proxy derived from the QLI is the unweighted satisfaction score.

Such interpretations should, however, be made with caution, since (in addition to other concerns addressed above), the explanation for the observed and unexpected correlational pattern may lie in the scoring algorithm of the weighted QLI. As noted above, this algorithm uses importance ratings as multiplicative weights of satisfaction scores. This, in itself, can be a source of dubious correlations with a third variable such as overall LS and QoL, since multiplicative operations can cause dramatic fluctuations in correlation coefficients with a third variable [42]. The explanation lies in multiplicative effects on the mean and variance of the weight, as well as the covariance between the weight and the third variable [43]. It is noteworthy that the use of multiplicative weights has survived the

development of the QLI as well as other similar instruments [e.g., 5, 6, 44], despite being cautioned against for decades [25, 42, 43].

Regardless of the methodological doubts related to the QLI weighting algorithm, the moderate correlation between overall LS and QoL does not support the view that the two constructs are synonymous. Indeed, theoretical arguments have suggested that life satisfaction should not be considered equivalent with, but a critical attribute or an outcome of QoL, which instead has been thought of as a broader and more encompassing term [45, 46]. Empirical support for this view can be found in general population surveys. In a survey conducted in British Columbia (Canada), Michalos et al. [47] found that the amount of variance in overall QoL explained by satisfaction with various life domains and health status indices was 63%, of which domain satisfaction explained 82% (i.e., about 51% of the total variance). In a cross-national UK general population survey, Bowling and Windsor [48] found that health and demographic variables explained 5% of the variance in subjective QoL, and addition of satisfaction with prioritized areas of life increased the explanatory power to 16%. While these two surveys differed on various accounts, both support the notion that life satisfaction appears to be an important contributor to, rather than synonymous with, subjective QoL. Several outcome instruments within the health sciences conceptualize QoL in terms of life satisfaction [e.g., 3, 5, 6, 9, 44]. We are, however, not aware of any previous studies of the relationship between these (or the QLI) and subjective overall QoL in chronically ill people. Together with previous observations in the general population, our findings illustrate that life satisfaction not necessarily is synonymous with QoL, and that investigators should be cautious when using these and related terms in order to avoid conceptual confusion.

In conclusion, this study does not support the practice of weighting ratings of satisfaction with various areas of life by their perceived importance, neither in terms of the resulting scores nor regarding their psychometric performance. Furthermore, use of

multiplicative weighting procedures render correlations with other variables spurious and cannot be recommended for methodological as well as practical reasons related to respondent burden. However, less than optimal data quality and psychometric performance of the unweighted QLI suggest that there is reason to consider other, simpler, life satisfaction questionnaires covering similar domains [e.g., 3, 49] instead. If deemed necessary to consider, ratings of perceived importance of various areas of life are probably best used separately from satisfaction ratings. These observations have implications not only for the QLI but also for other importance weighted life satisfaction questionnaires [e.g., 5, 6, 44].

ACKNOWLEDGEMENTS

The authors wish to thank the patients who participated in the study. This study was supported by the Swedish Research Council, the Skane County Council's Research and Development Foundation, and the Vårdal Institute, the Swedish Institute for Health Science, Lund University, Sweden.

REFERENCES

1. Anderson KL, Burckhardt CS. Conceptualization and measurement of quality of life as an outcome variable for health care intervention and research. J Adv Nurs 1999; 29(2): 298-306.

2. Beckie TM, Hayduk LA. Using perceived health to test the construct-related validity of global quality of life. Soc Indic Res 2004; 65: 276-298.

3. Burckhardt CS, Woods SL, Schultz AA, Ziebarth DM. Quality of life of adults with chronic illness: a psychometric study. Res Nurs Health 1989; 12(6): 347-354.

4. Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. ANS Adv Nurs Sci 1985; 8(1): 15-24.

5. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. Psychol Assess 1992; 4(1): 92-101.

6. Henrich G, Herschbach P. Questions on life satisfaction (FLZ^M) – a short questionnaire for assessing subjective quality of life. Eur J Psychol Assess 2000; 16: 150–159.

7. Holmes S. Assessing the quality of life--reality or impossible dream? A discussion paper. Int J Nurs Stud 2005; 42(4): 493-501.

8. Oleson M. Subjectively perceived quality of life. Image J Nurs Sch 1990; 22(3): 187-190.

9. Pavot W, Diener E. Review of the Satisfaction With life Scale. Psychol Assess 1993; 5(2): 164-172.

10. Campbell A, Converse PE, Rodgers WL. *The Quality of American Life: Perceptions, Evaluations and Satisfactions*. New York: Russell Sage Foundation, 1976.

11. Melin R. On life satisfaction and vocational rehabilitation outcome in Sweden. [Thesis] Uppsala: Uppsala University, 2003.

12. Birnbacher D. Quality of life - evaluation or description? Ethical Theory Moral Pract 1999; 2(1): 25-36.

13. Ferrans CE, Powers MJ. Psychometric assessment of the Quality of Life Index. Res Nurs Health 1992; 15(1): 29-38.

14. May LA, Warren S. Measuring quality of life of persons with spinal cord injury: substantive and structural validation. Qual Life Res 2001; 10(6): 503-515.

15. May LA, Warren S. Measuring quality of life of persons with spinal cord injury: external and structural validity. Spinal Cord 2002; 40(7): 341-350.

16. Hagell P, Whalley D, McKenna SP, Lindvall O. Health status measurement in Parkinson's disease: validity of the PDQ-39 and Nottingham Health Profile. Mov Disord 2003; 18: 773-783.

17. Gustafsson G. Quality of life and functional capacity among elderly with locomotor disability. [Thesis] Linköping: Linköping University, 1996.

18. Räty L, Hamrin E, Söderfeldt B. Quality of life in newly-debuted epilepsy. An empirical study. Acta Neurol Scand 1999; 100(4): 221-226.

19. Johansson IS, Hamrin EKF, Larsson G. Psychometric testing of the NEECHAM confusion scale among patients with hip fracture. Res Nurs Health 2002; 25(3): 203-211.

20. Ferrans CE. Development of a quality of life index for patients with cancer. Oncol Nurs Forum 1990; 17(3 Suppl): 15-19.

21. de Boer AG, van Lanschot JJ, Stalmeier PF, van Sandick JW, Hulscher JB, de Haes JC, Sprangers MA. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? Qual Life Res 2004; 13(2): 311-320.

22. 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: 40-66.

23. Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. J Clin Epidemiol 1998; 51(11): 945-952.

24. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol 1993; 78: 98-104.

25. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use. 2nd ed*. Oxford: Oxford University Press, 1995.

26. Howard KI, Forehand GA. A method for correcting item-total correlations for the effect of relevant item inclusion. Educ Psychol Measmt 1962; 22(4): 731–735.

27. Hobart JC, Cano SJ, Thompson AJ. What sample sizes for reliability and validity. Qual Life Res 2002; 11: 636.

28. Dougherty CM, Dewhurst T, Nichol WP, Spertus J. Comparison of three quality of life instruments in stable angina pectoris: Seattle Angina Questionnaire, Short Form Health Survey (SF-36), and Quality of Life Index-Cardiac Version III. J Clin Epidemiol 1998; 51(7): 569-575.

29. Faris JA, Stotts NA. The effect of percutaneous transluminal coronary angioplasty on quality of life. Prog Cardiovasc Nurs 1990; 5(4): 132-140.

30. Smith HJ, Taylor R, Mitchell A. A comparison of four quality of life instruments in cardiac patients: SF-36, QLI, QLMI, and SEIQoL. Heart 2000; 84(4): 390-394.

31. Stuifbergen AK. Cognitive impairment and perceptions of quality of life among individuals with multiple sclerosis. Rehabil Nurs Res 1995; 4(1): 11-18.

32. Weitzner MA, Meyers CA, Gelke CK, Byrne KS, Cella DF, Levin VA. The Functional Assessment of Cancer Therapy (FACT) scale. Development of a brain subscale and revalidation of the general version (FACT-G) in patients with primary brain tumors. Cancer 1995; 75(5): 1151-1161.

33. Yarbro CH, Ferrans CE. Quality of life of patients with prostate cancer treated with surgery or radiation therapy. Oncol Nurs Forum 1998; 25(4): 685-693.

34. Demers L, Ska B, Giroux F, Weiss-Lambrou R. Stability and reproducibility of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). J Rehabil Outcomes Meas 1999; 3(4): 42-52.

35. Jenkinson C, Mant J, Carter J, Wade D, Winner S. The London handicap scale: a re-evaluation of its validity using standard scoring and simple summation. J Neurol Neurosurg Psychiatry 2000; 68(3): 365-367.

36. Prieto L, Alonso J, Viladrich MC, Anto JM. Scaling the Spanish version of the Nottingham Health Profile: evidence of limited value of item weights. J Clin Epidemiol 1996; 49(1): 31-38.

37. Doward LC, Meads DM, Thorsen H. Requirements for quality of life instruments in clinical research. Value Health 2004; 7(Suppl. 1): S13-S16.

38. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002; 11: 193-205.

39. Rustøen T, Wiklund I, Hanestad BR, Burckhardt CS. Validity and reliability of the Norwegian version of the Ferrans and Powers Quality of Life Index. Scand J Caring Sci 1999; 13(2): 96-101.

40. de Vet HCW, Adér HJ, Terwee CB, Pouwer F. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. Qual Life Res 2005; 14(6): 1203-1218.

41. Anderson J, Ferrans C. The quality of life of persons with chronic fatigue syndrome. J Nerv Ment Dis 1997; 186(6): 359-367.

42. Schmidt FL. Implications of a measurement problem for expectancy theory research. Organizational Behavior and Human Performance 1973; 10: 243-251.

43. Evans MG. The problem of analyzing multiplicative composites: Interactions revisited. Am Psychol 1991; 46: 6-15.

44. Kuehler A, Henrich G, Schroeder U, Conrad B, Herschbach P, Ceballos-Baumann A. A novel quality of life instrument for deep brain stimulation in movement disorders. J Neurol Neurosurg Psychiatry 2003; 74(8): 1023-1030.

45. Haas BK. Clarification and integration of similar quality of life concepts. Image J Nurs Sch 1999; 31(3): 215-220.

46. Meeberg GA. Quality of life: a concept analysis. J Adv Nurs 1993; 18: 32-38.

47. Michalos AC, Zumbo BD, Hubley A. Health and the quality of life. Soc Indic Res 2000; 51: 245-286.

48. Bowling A, Windsor J. Towards the good life: a population survey of dimensions of quality of life. J Happiness Stud 2001; 2: 55-81.

49. Fugl-Meyer AR, Melin R, Fugl-Meyer KS. Life satisfaction in 18- to 64-year-old Swedes: in relation to gender, age, partner and immigrant status. J Rehabil Med 2002; 34(5): 239-246.

Legends to Figures

Fig. 1:

Scatterplots illustrating the relationships between total importance-weighted scores (y-axis) and total unweighted (A) satisfaction ($r_s = 0.962$) and (B) importance ($r_s = 0.434$) scores on Ferrans and Powers' Quality of Life Index. All scores were transformed to a possible range between 0 and 100 (observed score/maximum score x 100), where higher scores indicate better life satisfaction/higher importance.

Table 1. Descriptive statistics for weighted (satisfaction by importance) and unweighted (satisfaction only) scores on Ferrans and Powers' Quality of Life Index [a]

	QLI total		HF		SE		PS		FA	
	W	UW	W	UW	W	UW	W	UW	W	UW
Score mean (SD)	68.4 (9.1)	69.6 (10.9)	57.8 (18.9)	57.7 (21.0)	74.6 (10.9)	76.1 (12.0)	70.3 (14.6)	71.9 (15.8)	85.4 (12.9)	85.9 (12.9)
Score median (IQR)	68.2 (58.9, 75.6)	69.4 (65.3, 75.9)	56.2 (42.6, 69.1)	57.8 (45.7, 72.1)	76.1 (66.8, 89.8)	75.6 (66.7, 84.4)	70.6 (63.2, 78.8)	74.3 (65.0, 80.7)	89.4 (79.9, 95.0)	90.0 (80.0, 95.0)

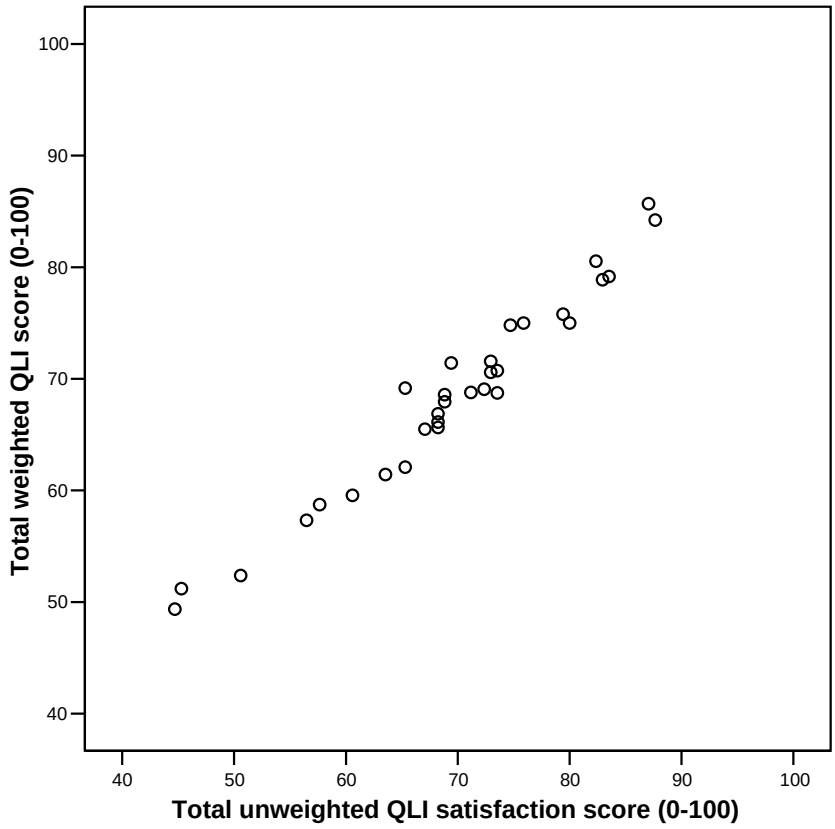
[a] All QLI scores are transformed to a possible range between 0-100.
QLI, Ferrans & Powers' Quality of Life Index; HF, Health and Functioning subscale; PS, Psychological/Spiritual subscale; SE, Social and Economic subscale; FA, Family subscale; W, weighted (satisfaction by importance) score; UW, unweighted (satisfaction only) score; SD, standard deviation; IQR, inter-quartile range.

Table 2. Psychometric properties for weighted (satisfaction by importance) and unweighted (satisfaction only) scores on Ferrans and Powers' Quality of Life Index

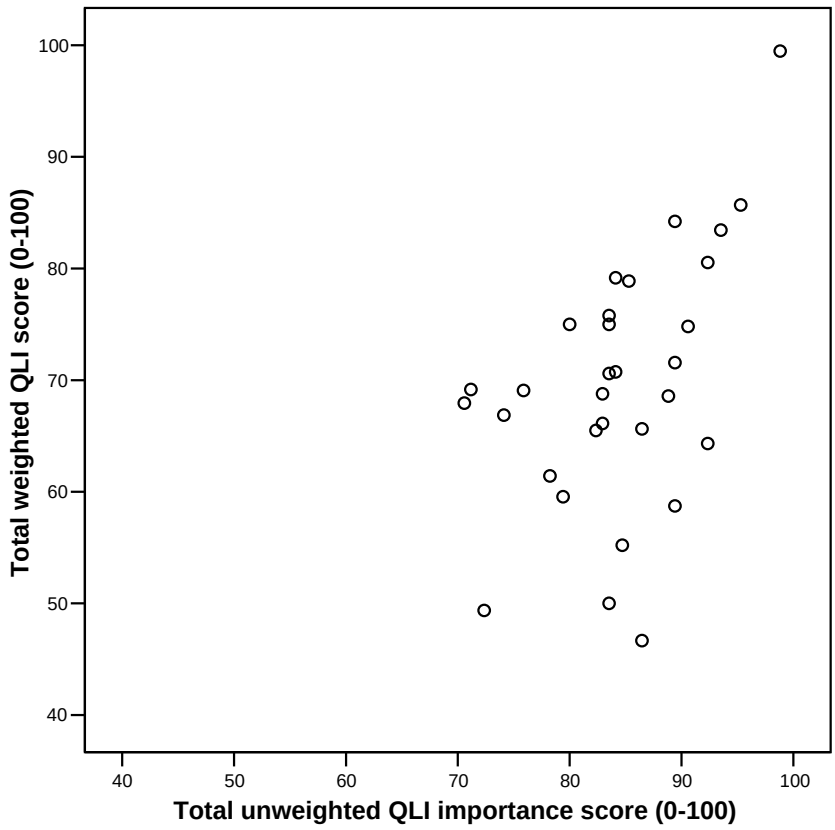
	QLI total		HF		SE		PS		FA	
	W	UW	W	UW	W	UW	W	UW	W	UW
Overall (individual) item non-response (%)	11.9 [a] (4.2-45.1) [b]	8.2 (1.4-38.0)	8.6 (4.2-18.3)	5.4 (1.4-16.9)	19.1 (5.6-45.1)	14.2 (4.2-38.0)	9.4 (7.0-12.7)	6.0 (4.2-9.9)	11.3 (7.0-18.3)	8.4 (5.6-14.1)
Respondents with missing item responses (%)	64.8	56.3	33.8	29.6	57.7	50.7	16.9	12.7	21.1	19.7
Reliability (min, max if item deleted) [c]	0.86 (0.84, 0.87)	0.88 (0.87, 0.89)	0.92 (0.90, 0.92)	0.92 (0.90, 0.92)	0.69 (0.63, 0.74)	0.62 (0.56, 0.62)	0.87 (0.84, 0.87)	0.83 (0.77, 0.83)	0.57 (0.42, 0.55)	0.56 (0.41, 0.56)
Definite scaling success (%) [d]	-	-	64.3	64.3	11.1	7.4	52.4	42.8	50	66.7
Possible scaling success (%) [e]	-	-	30.9	30.9	48.1	44.4	42.8	52.4	41.7	33.3
Scaling failure (%) [f]	-	-	4.8	4.8	44.4	48.2	4.8	4.8	8.3	0

[a] Proportion of item non-response for whole scales.
[b] Range of proportions of item non-response for individual items within scales.
[c] Cronbach's coefficient alpha; should exceed 0.70.
[d] Percentage of instances when items correlated significantly stronger (≥2 standard errors) with their hypothesized scale than with other scales.
[e] Percentage of instances when items correlated non-significantly stronger (<2 standard errors) with their hypothesized scale than with other scales.
[f] Percentage of instances when items correlated stronger with other scales than with their hypothesized scale.
QLI, Ferrans & Powers Quality of Life Index; HF, Health and Functioning subscale; PS, Psychological/Spiritual subscale; SE, Social and Economic subscale; FA, Family subscale; W, weighted (satisfaction by importance) score; UW, unweighted (satisfaction only) score.

a


Total weighted QLI score (0-100)
100
90
80
70
60
50
40
40
50
60
70
80
90
100
Total unweighted QLI satisfaction score (0-100)

b


Total weighted QLI score (0-100)
100
90
80
70
60
50
40
40
50
60
70
80
90
100
Total unweighted QLI importance score (0-100)