UNIVERSITÀ DEGLI STUDI FIRENZE

FLORE
Repository istituzionale dell'Università degli Studi di Firenze

Item re-scaling of an Italian version of the Sickness Impact Profile: Effect of age and profession of the observers

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

Original Citation:

Item re-scaling of an Italian version of the Sickness Impact Profile: Effect of age and profession of the observers / Niccoló Marchionni;Luigi Ferrucci;Samuele Baldasseroni;Stefano Fumagalli;Jack M. Guralnik;Maura Bonazinga;Francesca Cecchi;Giulio Masotti. - In: JOURNAL OF CLINICAL EPIDEMIOLOGY. - ISSN 0895-4356. - STAMPA. - 50:(1997), pp. 195-201. [10.1016/S0895-4356(96)00318-6]

Availability:

This version is available at: 2158/771343 since:

Published version:

DOI: 10.1016/S0895-4356(96)00318-6

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf)

Publisher copyright claim:

(Article begins on next page)

12 October 2024

J Clin Epidemiol Vol. 50 No. 2, pp. 195–201, 1997
Copyright © 1997 Elsevier Science Inc.

ELSEVIER

0895-4356/97/$17.00
PII S0895-4356(96)00318-6

Item Re-Scaling of an Italian Version of the Sickness Impact Profile: Effect of Age and Profession of the Observers

Niccoló Marchionni,[1,*] *Luigi Ferrucci,*[2] *Samuele Baldasseroni,*[1] *Stefano Fumagalli,*[1] *Jack M. Guralnik,*[3] *Maura Bonazinga,*[1] *Francesca Cecchi,*[2] *and Giulio Masotti*[1]

[1]The University Department of Gerontology and Geriatric Medicine, Florence, Italy 50141, the [2]INRCA Department of Geriatric Medicine, Florence, Italy 50139, and the [3]National Institute on Aging, Bethesda, Maryland 20892, United States

ABSTRACT. An Italian version of the Sickness Impact Profile (SIP) obtained by professional and nonprofessional translators was checked for cross-cultural equivalence using a back-translation method followed by two scaling studies. The first scaling study involved 30 health professionals who ranked the items within each category for severity of dysfunction. By comparing Italian and US average ranks, 14 highly discordant items were identified. A revised translation was evaluated in a new study involving 120 observers stratified by age (<65 versus ≥65 years) and profession (health versus non health professionals) into 4 groups of the same size. The Italian and American item rank orders were almost equivalent, independently of the age and profession of the observers (93% of the ranks showing differences <2), suggesting that this Italian version of SIP is cross-culturally unbiased. However, older age was associated with higher variability in the rank orders, and some caution is required for use in the geriatric population. *Copyright © 1997 Elsevier Science Inc.* J CLIN EPIDEMIOL 50;2:195–201, 1997.

KEY WORDS. SIP, quality of life, psychometric methods, geriatric medicine, multidimensional assessment, disability

INTRODUCTION

In health care research, Quality of life (QoL) has become a critically important outcome when assessing the effects of medical interventions. Several studies have suggested that the efficacy of health care provision assessed by conventional criteria such as mortality, morbidity, or functional status, do not always translate into positive changes in quality of life [1,2]. This issue is particularly important in older people, who are generally more likely to suffer adverse reactions to different medical interventions, including pharmacological treatments [3], hospitalization [4], and surgical procedures [5]. However, the choice of the most appropriate instrument for the assessment of QoL, which is not an easy task in general, becomes even more difficult in selected subgroups of the population, such as older persons. A number of papers have previously addressed the conceptualization of QoL and the identification of the elements most pertinent to its definition [6–8]. Nevertheless, how to measure QoL is still an issue under debate [6,7]. Many of the instruments proposed in the literature have been designed for patients with particular clinical characteristics [6,9–11]. These disease-specific instruments cannot be used to compare populations with different illnesses or problems and, therefore, they are of particularly limited value in older patients who are often affected by multiple pathologic conditions. In the geriatric setting, a less specific questionnaire that can reliably measure QoL across a wide range of clinical circumstances would be highly desirable, provided that its psychometric characteristics are assessed in an older population affected by a large variety of chronic diseases, and over a wide range of functional conditions [12]. The Sickness Impact Profile (SIP) [13,14], may be the instrument of choice for this purpose. In fact, the SIP is a questionnaire intended to measure health-related QoL by assessing changes in patients self-perceived health and behavior that can be attributed to chronic diseases. It provides a measurement of QoL through a range of different health-related domains so wide as to become a good candidate for extended utilization also in the older population.

The SIP has been extensively validated in the past decade in a variety of clinical settings [15] and in patients with a wide age range affected by many different chronic conditions. Translations into many different languages are

*Address for correspondence: Niccoló Marchionni, MD, Department of Gerontology and Geriatric Medicine, University of Florence. Via delle Oblate, 4.50141, Florence. Italy.
Accepted for publication on 7 August 1996.

already available [16–18]. Thus, scientists of different countries may use this instrument to collate their data and to perform cross-national comparisons [19,20]. Unfortunately, local translations have often been produced by a single well-meaning translator, and in very few cases have translated versions undergone a process of formal psychometric evaluation. The validity of such straightforward translations is questionable. In particular, while several translations of the SIP have appeared in Italy in the last few years, no Italian version has been acknowledged until now by the original authors, nor have any been tested for cross-cultural correspondence, reliability, or validity.

The translation of an instrument into a language different from the original requires a new, complete process of validation of the instrument in its translated form. This validation is an essential prerequisite for the use of any new Italian version of the SIP. The SIP global and category-specific scores are calculated by adding the weighted scores of individual items checked by the subject and then dividing by the maximum possible dysfunction score. In the original version of the SIP, the weights assigned to each item were determined according to the average judgment expressed by many different observers, using a scaling technique that assumes equal appearing intervals [21]. The verification of a similar hierarchical order in the Italian version, and its stability over a wide range of age and subjective perspectives on health-related changes in life-standard (for example the perspective of a health professional versus the perspective of a health service user) is an essential precondition for the utilization of an Italian version of the SIP both in research projects and in clinical practice.

This article describes the procedure that has been followed to develop an Italian version of the SIP, to test the cross-cultural similarity with the hierarchical order of the weights assigned to the items in the original version, and to determine whether such order may be affected by the age or the profession (health versus non-health professionals) of the examiner.

METHODS

The procedure consisted of three distinct phases: the translation of the instrument, and two sequential scaling studies.

Translation of the SIP

The SIP consists of a questionnaire of 136 items. The subject is asked to identify those most appropriately describing his/her own condition. The items are partitioned into 12 categories that describe several components of common daily life: sleep and rest (SR); emotional behavior (EB); body care and movement (BCM); home maintenance (HM); mobility (M); social interactions (SI); ambulation (A); alertness behavior (AB); communication (C); work (W); recreation and pastimes (RP); eating (E).

The SIP questionnaire was translated into Italian by a member of the research team who had not previously had experience with the instrument. Whenever a literal translation of the item would not adequately represent the condition described in the original item, the sentence was rephrased to maintain conceptual cross-cultural equivalence. The Italian version was translated back into English by another member of the research team and by a professional English translator who had never seen the original instrument. During the back-translation procedure, disagreement between the two translators was always resolved by discussion, and a consensus was always reached to obtain a common English back-translated version. The two English versions (the original and the back-translated one) were examined and approved for identity of meaning of the 136 items (content validity) by an editorial committee that was composed of a member of the research group, two bilingual medical professionals and two bilingual non-medical professionals. Finally, the equivalence between the back-translated English version and the original version was verified by one of the authors of the original SIP (M. Bergner,[1] Johns Hopkins Medical School, Baltimore, MD, USA) [13,14]. Having successfully translated the SIP, the research team proceeded to a first study of hierarchical item scaling, following a protocol similar to that previously used in the validation process of the French version [16].

First Item Scaling Study

The first study involved a convenience sample of 30 health professionals (14 men and 16 women, mean age 46 ± 13 years) with no clinical evidence of cognitive impairment, depression or sensory deficit: there were 12 physicians, 9 nurses, and 9 physical therapists. All of them received a summary description of the SIP, of the general project of cross-cultural adaptation, and of their specific role in this process.

The items of the SIP were printed on individual colored cards, the color being different for each of the 12 SIP categories. Study participants were individually asked to order the colored cards within each category according to their perception of severity of dysfunction, without considering any particular pathologic condition or prognostic element. A member of the research group was present to offer information or clarification if needed, but in no case did he or she give any suggestion on the ordering procedure.

An average rank across raters was then calculated for each item and these average values were used to obtain a hierarchical rank order within each category, which was then compared with the rank order of the original SIP.

The data of this first item scaling study were extensively reviewed by the research group and a professional translator. Particular attention was given to the items whose rank differed more than two positions from the original U.S. version. Three possible explanations for differences in ranking

[1]Deceased, December, 1992.

TABLE 1. Examples of changes made to the wording of some items of the Italian SIP version between the first and the second item scaling study

Item	Original U.S. version	First Italian version	Second Italian version
BCM 8[a]	I am in a restricted position all the time.	Sono sempre in una posizione costretta.	Posso stare solo in una posizione obbligata.
16[a]	I use bedpan with assistance.	Mi devono portare la padella a letto.	Ho bisogno di aiuto per usare la padella.
HM 5[b]	I am not doing *any* of the shopping that I would usually do.	Non faccio nessuna delle compere che avrei fatto di solito.	Non vado più a fare *nessuna* delle compere che avrei fatto di solito.
SI 4[c]	I often act irritable toward those around me, for example, snap at people, give sharp answers, criticize easily.	Spesso mi comporto in modo irritante nei confronti di coloro che mi stanno vicino, per esempio alzo la voce, do risposte brusche, li critico facilmente.	Spesso mi comporto in modo irritabile nei confronti di coloro che mi stanno vicino, per esempio rispondo male, in modo brusco, oppure critico facilmente.
8[a]	I am avoiding social visits from others.	Evito di ricevere visite di cortesia.	Evito di ricevere visite di amici e conoscenti.
10[d]	I often express concern over what might be happening to my health.	Esprimo spesso preoccupazioni su quello che potrebbe accadere alla mia salute.	Esprimo spesso preoccupazioni sulla mia salute.
12[c]	I make many demands, for example, insist that people do things for me, tell them how to do things.	Faccio molte richieste, per esempio, insisto che gli altri facciano delle cose per me, dicendo loro come farle.	Ho molte pretese, per esempio, insisto che gli altri facciano delle cose per me, dicendo loro come farle.

[a]Literal translation assumes a different meaning in Italian.
[b]The severity of the condition needs to be stressed by underlining.
[c]Error in the Italian translation detected.
[d]Unusual expression in Italian language, potentially misleading.

were identified: (1) some expressions (e.g., those concerning peculiar and difficult movements, use of mechanical aids, the meaning of "social visits") that were appropriate in the U.S. version were obscure or open to misinterpretation in the Italian version; (2) the translated version failed to maintain the emphasis on certain words or ideas found in the original (e.g., "I do not do *any* of the previous activities") that had not been maintained in the translated version; (3) real cross-cultural differences were considered possible in the evaluation of some activities (e.g., sexual activity). Revisions were made to items where translation problems were felt to account for differences in ranking. Selected examples of changes made to the wording of some items between the first and the second scaling study are given in Table 1. The revised version underwent a second scaling study.

Second Item Scaling Study

The second study involved 120 observers (58 men and 62 women) who, according to their age (<65 versus ≥65 years) and profession (health versus non-health professionals) were stratified into four groups of 30 individuals each: (1) young non-health professionals (mean age 29 ± 8 years); (2) young health professionals (mean age 35 ± 9 years); (3) older non-health professionals (mean age 74 ± 8 years); (4) older health professionals (mean age 74 ± 7 years). Clinical evidence of cognitive impairment, depression or sensory impairment were considered as exclusion criteria. The method of scaling the 136 items of the Italian SIP instrument followed the same guidelines described in the previous section.

Statistical Analysis

Data were analyzed using the SAS statistical package [22]. Both in the first and in the second scaling study, the strength of the correlation between the average ranks of the U.S. and the Italian version was estimated by computing Spearman rank order correlation coefficients. The size of these coefficients indicate the level of concordance between the average item rank order expressed by the Italian group of reviewers and the item rank order reported for the original SIP version.

The level of concordance was also estimated for each subject, separately for each category, by computing a correlation coefficient between his/her specific rank order, and the rank order of the original SIP. The distribution of these correlation coefficients provides information on variability in ranking that may be lost if only the average rank order is considered in the analysis. The effect of age and professional status (health versus non-health professionals) on the variability of the observers' judgments was tested by full factorial analysis of variance models performed separately for each of the categories.

RESULTS

First and Second Item Scaling Study

The average ranks across raters of the items of the Italian SIP that were calculated for the first item scaling study are compared with the original US version in the left panel of Fig. 1. In spite of a high correlation in the series as a whole ($r = 0.89, p < 0.001$), the correlation coefficients calculated

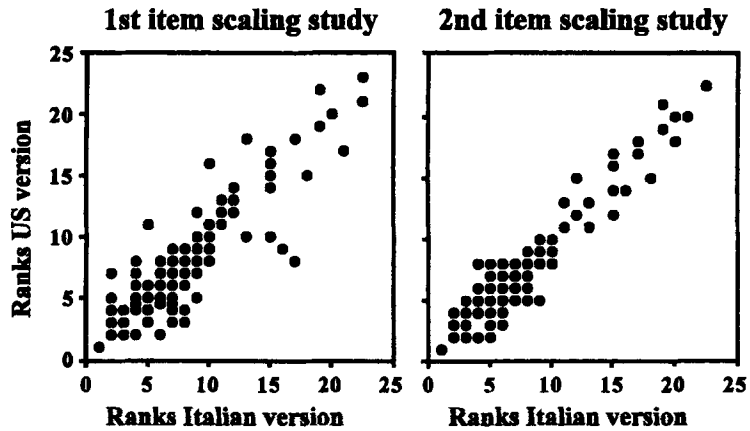

FIGURE 1. Comparison of the average ranks of severity attributed to the 136 items of the Italian SIP with the ranks of the original U.S. version during the first (*left panel*) and the second (*right panel*) item scaling study.

in each of the 12 categories were spread over a wide range, from 0.23 (BCM) to 0.94 (E). For only 36% of the 136 items the difference in ranking between the two versions was 0, while a difference in rank of ≥2 or ≥3 was observed for 37% and 18% of the items, respectively. For some of the items of the BCM category such a difference was as large as 9 positions. These results led to a major revision of the Italian version. Changes in the translation (Table 1) were necessary for 14 items, and the modified SIP version was evaluated again through the second scaling study.

The overall results of the second scaling study are reported in the right panel of Figure 1. Concordance between the average ranks of the two versions did considerably improve after the revision of the Italian SIP ($r = 0.93$, $p < 0.001$). In the analyses restricted within each category the minimal correlation (0.71) was found for HM. Considering the whole series of 136 items, 45% of their ranks were found to be identical (difference = 0), 25% and 7% of the ranks showed differences ≥1 or ≥2, respectively, and only in 2 cases (1.7%, both in the BCM category) the rank difference was >4. Thus, in general, the rank order rating of dysfunction attributed to the items within each category by Italian and American judges showed scale equivalence.

Effects of Age and Profession

Concordance between the Italian and U.S. versions was substantially independent of age and of the profession of the observer, as shown in the four panels of Fig. 2, where correlations are virtually identical among the four groups of raters stratified by age and professional status.

A further analysis done separately in each category was carried out to quantify the concordance of each Italian observer with the order of severity of the items as reported in the U.S. version. In this analysis, for each individual, full concordance is indicated by a correlation coefficient of 1, while discordance increases as the coefficient approaches 0 and is maximal when it equals −1. The effect of age and profession on the degree of concordance was analyzed by models of factorial analysis of variance, using the individual correlation coefficients as the dependent variables in the models (Table 2).

Single observer concordance was significantly lower among older observers in all categories except for AB and RP, whereas no effect of the type of profession could be detected. The percent distribution of the correlation coefficients (Fig. 3) was consistent with these findings. In fact, in all categories except AB and RP the correlation coefficients were more often in the high range among younger than among older observers. Interestingly, older observers also showed more variable correlation coefficients (Fig. 3), in spite of the close correspondence between their overall evaluation and the order of the reference U.S. version (Figure 2).

DISCUSSION

This study shows that the general psychometric properties of the original version of the SIP are retained in our Italian translation. These results suggest that our translation procedure was effective and that the underlying hierarchical structure of the items in each category is robust enough to withstand adaptation to another language. In general, the results of the item scaling analysis were similar to those reported for the original SIP, and no systematic difference according to the age or to the profession of the observer was found. These findings are consistent with previous reports that demonstrated that, at least in large part, the assessment of a deteriorating health status is not distorted cross-culturally [16–18]. It is interesting to note that these results could be obtained only by means of a complex procedure consisting of a double translation, followed by a careful re-wording of the translated items that was guided by a two-step, dynamic re-scaling process. A substantial improvement in the cross-cultural correspondence of the weights assigned to the single items used to calculate the average SIP score was obtained by even marginal changes in the wording of the questionnaire. These findings further stress the need for a formal process of psychometric evaluation of translated versions of any questionnaire before they can be

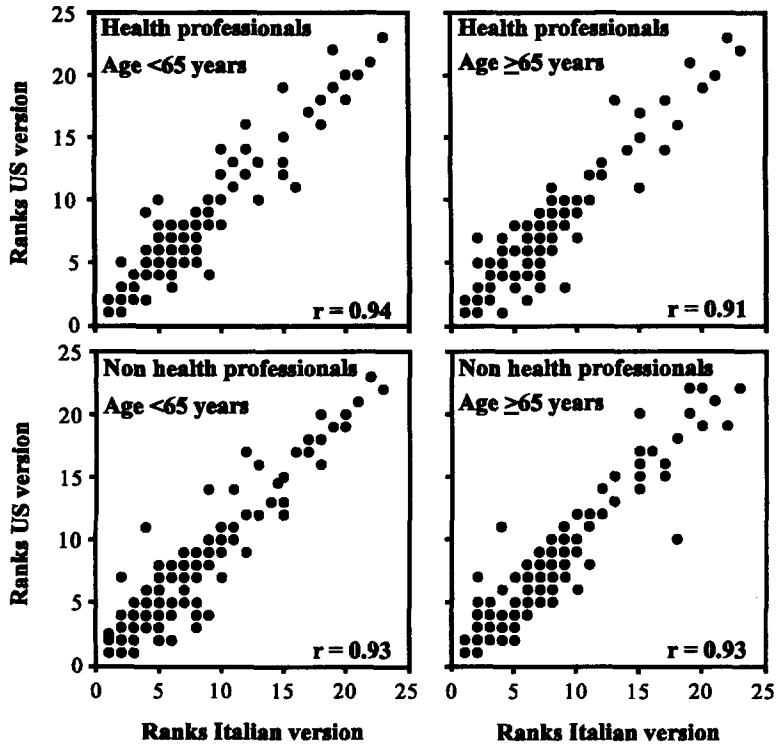

FIGURE 2. Comparison of the average ranks of severity attributed to the 136 items of the Italian SIP with the ranks of the original U.S. version in the second item scaling study. Data are plotted separately for observers under and over the age of 65 years and for health and non health professionals.

reliably used either in clinical or in epidemiological studies. A recommendation that holds particularly true for instruments that are to be specifically used in cross-cultural comparisons.

There are several reasons that might explain the relatively unsatisfactory results of the first item scaling study and the improvement that was observed in the second one. Improvement may have occurred by chance, simply because of sampling variations in the selection of the raters. However, since the average rank order observed in the four subgroups of the second study was substantially the same, we believe that this hypothesis is unlikely to be true. Secondly, some of the large differences in item rating observed in the first study might reflect real cross-cultural differences between Italian and American raters. The improved concordance observed in the second scaling study contrasts with this hypothesis and rather indicates that inaccurate translation in the meaning of several phrases (Table 1) was the most likely cause of the poorer correspondence reported by the first study.

A result of broader interest is that when judgments of older (65+ years) observers are averaged, the rank of the items within each category is quite similar to the average rank order calculated for the younger observers, whereas

TABLE 2. Analysis of the effect of age, profession, and their interaction on the concordance between the hierarchical order expressed by each Italian observer and the rank order of the original SIP version

	Independent variable		
SIP categories	Age (<65 vs. ≥65 years) F (*p*)	Health versus non-health professional F (*p*)	Age × profession F (*p*)
SR	10.99 (<0.001)	0.67 (0.41)	0.59 (0.44)
EB	23.11 (<0.001)	0.07 (0.80)	0.05 (0.83)
BCM	13.71 (<0.001)	0.71 (0.68)	0.26 (0.61)
SI	18.87 (<0.001)	2.57 (0.11)	0.02 (0.87)
A	7.96 (<0.01)	0.03 (0.86)	0.01 (0.91)
HM	12.96 (<0.001)	0.30 (0.59)	0.00 (0.99)
M	4.68 (<0.05)	2.87 (0.09)	2.86 (0.09)
AB	3.82 (0.06)	0.84 (0.36)	0.48 (0.49)
C	13.85 (<0.001)	0.73 (0.39)	0.00 (0.97)
W	5.63 (<0.05)	0.01 (0.92)	0.48 (0.49)
RP	2.84 (0.10)	3.25 (0.07)	0.07 (0.79)
E	14.84 (<0.001)	1.54 (0.22)	0.69 (0.41)

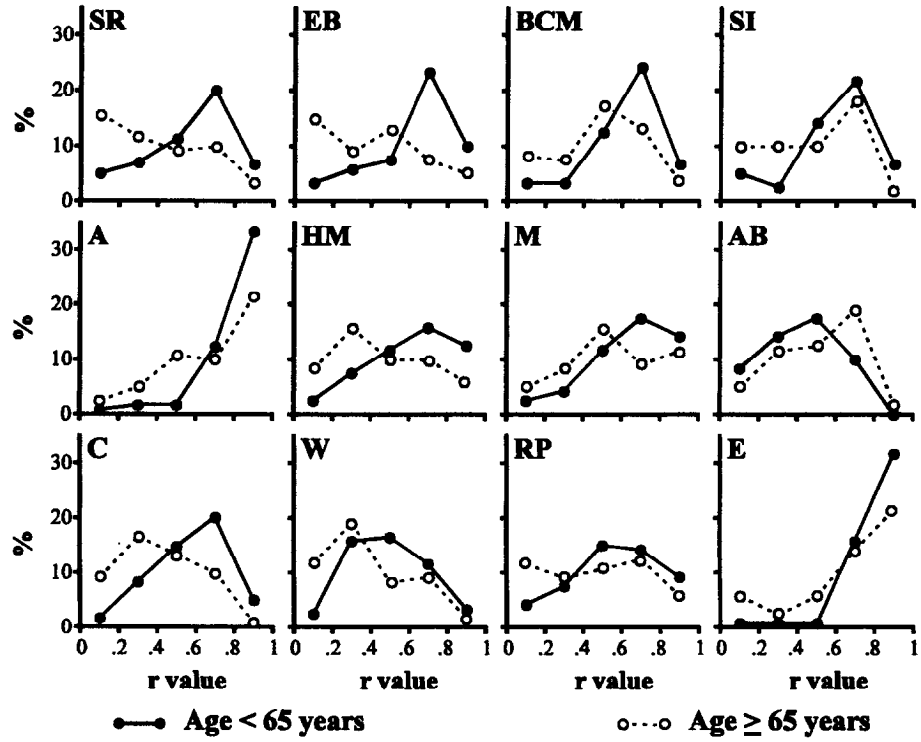

FIGURE 3. Percent distribution of the correlation coefficients between the item rank order reported by each of the 120 Italian observers and the item rank order of the reference US version in the second item scaling study. Data are plotted separately for observers under and over the age of 65 years for each SIP category.

single older observers put the items in a significantly more variable order (compared to U.S. ranks) than younger ones. This finding was in some way unexpected and deserves comment. We are convinced that such a difference stemmed principally from previous or present experience of a disease. Older persons, who are more likely to have experienced some of the specific health-related conditions described in the SIP, may rate those states as more severe, compared to people without comparable experience [7,23]. This hypothesis raises an issue of crucial methodological importance in the use of rater-based scaling techniques. In scaling studies, such as the present one, raters who can weight the severity of hypothesized experience, without being influenced by their own past experience, would be highly desirable. We chose to rescale the SIP also with older raters to approach more closely the experience of patients who are most often evaluated by this instrument. However, older raters are more likely to have had previous health problems. This might raise questions on the generalizability of the weights obtained in a scaling study, when an instrument is thereafter actually used for rating health experiences in real patients. A recent report [24] has shown that current health status has an important effect on the evaluations attached to different health states, but past experience of illness has only a negligible effect on these evaluations. We believe that scaling studies maintain their validity even when they include older raters, since present health experiences in a random sample of raters cover a wide range, and should therefore be reciprocally balanced when the average ranks of weights are considered. Thus, our finding of a larger variability of judgments among older raters does not reduce the validity of the SIP, but suggests only the need for some caution when the SIP is used to compare the QoL between patients at the individual level. The SIP can be reliably used to measure the QoL of groups, even with older patients.

Some limitations of the present study should be noted. First, this item scaling study was carried out with a slightly different method from that used by the original authors [21], who asked the same set of raters to rank the items of each category twice, first from 1 to 11 and later from 1 to 15. Indeed, the original method assumes that judges have a complete understanding of the criteria to be used in the scaling technique and are able to deliver pure quantitative judgments of the severity of a certain condition. In a pilot study, we found this direct scoring method not easily accepted by older patients. On the contrary, the same patients could provide appropriate indication on the hierarchical order of items belonging to the same category through direct comparison. Since one of our purposes was to test the effect of age on the SIP item scaling, we stratified our study population by age, including therefore similar numbers of individuals younger and older than 65 years. Thus, we preferred to simplify the rating operation by asking observers only to order the items according to their severity within each category. For the same reason, we did not attempt to obtain any comparison between the severity of the different categories. Although this is not strictly a scale equivalence, this method has been successfully applied to the French version of the SIP [16].

In spite of these limitations, the results presented here

demonstrate that the Italian version of SIP developed through a careful, systematic research protocol maintains reliable psychometric characteristics. For practical purposes, the Italian and the American version of the SIP should be considered cross-culturally unbiased, and therefore suitable for comparing health-related QoL as determined by chronic diseases in different countries.

A validation study is currently ongoing on large groups of patients affected by several different chronic conditions. The combined results of these research efforts are expected to provide an instrument that can be widely used in clinical settings and research projects.

This work has been supported in part by a grant from the Special Project (40%) on Cardiac Failure, the Italian Ministry of the University and the Scientific and Technological Research, and in part by a grant from the Italian National Research Council (CNR), Special Project on Aging (SP5).

References

1. Applegate WB, Phillips HL, Schnaper H, Shepherd AM, Schocken D, Luhr JC, Koch GG, Park GD. A randomized controlled trial of the effects of three antihypertensive agents on blood pressure control and quality of life in older women. **Arch Intern Med** 1991; 151: 1817–1823.
2. Balducci L. Perspectives on quality of life of older patients with cancer. **Drugs Aging** 1994; 4: 313–324.
3. Pahor M, Guralnik JM, Gambassi G Jr, Bernabei R, Carosella L, Carbonin PU. The impact of age on the risk of adverse drug reactions to digoxin. **J Clin Epidemiol** 1993; 46: 1305–1314.
4. Goldman L, Caldera DL, Nussbaum SR. Multifactorial index of cardiac risk in non-cardiac surgical procedures. **N Engl J Med** 1977; 29: 845–855.
5. Gorbien MJ, Bishop J, Beers MH, Norman D, Osterweil D, Rubenstein LZ. Iatrogenic illness in hospitalized elderly people. **J Am Geriatr Soc** 1992; 40: 1031–1042.
6. Schipper H, Clinch J, Powell W. Definition and conceptual issues. In: Spilker B, Ed. **Quality of Life Assessment in Clinical Trials**. New York: Raven Press; 1990, pp. 11–24.
7. Gill TM, Feinstein AR. A critical appraisal of the quality-of-life measurements. **JAMA** 1994; 272: 619–626.
8. McSweeny AJ, Creer TL. Health-related quality-of-life assessment in medical care. **Dis Mon** 1995; 41: 1–71.
9. Osoba D. Lessons learned from measuring health-related quality of life in oncology. **J Clin Oncol** 1994; 12: 608–616.
10. Wilkin D. Longitudinal studies of patients with chronic illness: Problems of linking process and outcome. **Scand J Prim Health Care Suppl** 1993; 2: 54–57.
11. de Haan RJ, Limburg M, Van der Meulen JH, Jacobs HM, Aaronson NK. Quality of life after stroke. Impact of stroke type and lesion location. **Stroke** 1995; 26: 402–408.
12. Spilker B. Introduction. In: Spilker B, Ed. **Quality of Life Assessment in Clinical Trials**. New York: Raven Press; 1990: 3–10.
13. Bergner M, Bobbit RA, Pollard WE, Martin DP, Gilson BS. The Sickness Impact Profile: Validation of a health status measure. **Med Care** 1976; 14: 57–67.
14. Bergner M, Bobbit RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. **Med Care** 1981; 19: 787–805.
15. McDowell I, Newell C. General health measurements. In: McDowell I, Newell C, Eds. **Measuring Health. A Guide to Rating Scales and Questionnaires**. New York: Oxford University Press; 1987: 269–321.
16. Chwalow AJ, Lurie A, Bean K. A French version of the Sickness Impact Profile (SIP): Stages in the cross cultural validation of a generic quality of life scale. **Fundam Clin Pharmacol** 1992; 6: 319–326.
17. Gibson BS, Erickson D, Chavez CT, Bobbit RA, Bergner M, Carter WB. A Chicano version of the Sickness Impact Profile. **Cult Med Psychiatry** 1980; 4: 137–150.
18. Badia X, Alonso J. Re-scaling the Spanish version of the Sickness Impact Profile: An opportunity for the assessment of cross-cultural equivalence. **J Clin Epidemiol** 1995; 48: 949–957.
19. Patrick DL, Sittampalam Y, Somerville SM, Carter WB, Bergner M. A cross cultural comparison of health status values. **Am J Public Health** 1985; 75: 1402–1407.
20. Badoux A, Mendelsohn GA. Subjective well-being in French and American samples: Scale development and comparative data. **Qual Life Res** 1994; 3: 395–401.
21. Carter B, Bobbit RA, Bergner M, Gilson BS. Validation of an interval scaling: The Sickness Impact Profile. **Health Serv Res** 1976; 4: 516–528.
22. SAS Institute Inc. **SAS Procedures Guide, Version 6.** Cary, NC: SAS Institute, Inc.; 1993.
23. Kind P, Dolan P. The effect of past and present illness experience on the evaluations of health states. **Med Care** 1995; 33: AS255–263.
24. Dolan P. The effect of experience of illness on health state valuations. **J Clin Epidemiol** 1996; 49: 551–564.