Intensive Care Med (1996) 22:1034–1042


ORIGINAL

R. Rivera Fernandez
J. J. Sanchez Cruz
G. Vazquez Mata

# Validation of a quality of life questionnaire for critically ill patients

Received: 16 August 1995
Accepted: 11 June 1996

This study forms part of the PAEEC (Project for the Epidemiological Analysis of Critical Care Patients), and was supported by a grant from the "Fondo de Investigaciones Sanitarias" (F.I.S.-91/0703), and by the Granada University Research Group (Number 3244)

R. Rivera Fernandez
G. Vazquez Mata (✉)
Servicio de Medicina Intensiva,
CRT Hospital Virgen de las Nieves,
Carretera de Jaen, sn,
E-18013 Granada, Spain
FAX: +34(58)241282

J. J. Sanchez Cruz
Escuela Andaluza de Salud Pública,
Granada, Spain

**Abstract** *Objective:* Development and validation of quality of life questionnaire for critical care patients.
*Design:* Prospective study.
*Setting:* Intensive care unit (ICU) of a general hospital and ICUs of 83 Spanish hospitals.
*Sample:* Patients admitted to the ICU >18 years of age; close family members.
*Method:* A committee of experts designed a questionnaire with characteristics judged essential for intensive care use: easy, quick administration (5–10 min); capable of completion by patient or close family member, by direct or telephone interview. Fifteen items relevant to critical care patients were grouped in three subscales: basic physiological activities, normal daily activities, and emotional state. Reproducibility of interobserver, intraobserver, patient/family member and telephone/direct interviews was analysed and also internal consistency, responsiveness, and main components.
*Results:* Internal consistency (578 patients): Cronbach's alpha coefficient = 0.85. Reproducibility: intraobserver reproducibility ($n$ = 150): Spearman correlation coefficient = 0.92. Interobserver ($n$ = 85): correlation = 0.92. Patient/family member ($n$ = 81): correlation = 0.92. Telephone/direct interview ($n$=54): correlation = 0.96. Validity: factorial analysis confirmed that the three subscales were fundamental questionnaire components. There was good concordance between questionnaire/subscale and Glasgow Outcome Scale (GOS) results. Responsiveness: quality of life score changes between preadmission and 6 months' postdischarge correlated with GOS findings (weighted kappa index = 0.56).
*Conclusions:* Questionnaire meets objectives recommended for critical care use, and fulfills essential requirements of validity and reproducibility when applied to critically ill patients.

**Key words** Health status · Intensive care · Quality of life

## Introduction

Clinical outcomes in intensive care are evaluated in two main ways: by analysis of mortality and of quality of life. The former method has advanced considerably with the development of standardized systems for predicting mortality, and those in most widespread use have been subject to several revisions, as is the case with the Acute Physiology and Chronic Health Evaluation (APACHE) systems [1–3], the Simplified Acute Physiology Score [4, 5] and Mortality Probability Models [6, 7]. By contrast,

quality of life has been less thoroughly studied, despite its importance in outcomes analysis [8–10].

Quality of life is a multidimensional concept [11] and includes aspects concerning functional capacity, physiological functions, affective states and emotional behaviours, social interaction, work activity and economic situation. There is no consensus in the literature about which particular dimensions should be included in quality of life, and the instruments used for its measurement differ considerably in their selection of variables. There is, however, a consensus [12] that evaluation of quality of life must include both objective (functional capacity, physiological functions) and subjective (emotions, states of mind) aspects of the patient's life.

A quality of life survey for critical care patients should have the following characteristics [9]: it can be universally applied to all types of critical care patients; it can be filled in by a close family member when the patient cannot respond, which is a frequent situation in critical care; it can be expressed as a score; and it can be completed easily and rapidly to permit routine use. It should be suitable for use not only in follow-up studies after discharge from the intensive care unit (ICU) but also at the time of admission to the ICU, evaluating a previous stable period and differentiating this from the particular situation prompting admission. It is the comparison between the stable situations before and after the stay in the ICU which reveals the impact of the illness on quality of life.

Our research team had developed a quality of life questionnaire [13, 14] which was properly validated but which ignored subjective aspects. The survey contained only seven items and was simple and quick to complete. This second survey attempts to reflect quality of life in a more complete way, including the subjective dimension. It fulfills the above-mentioned criteria for critical care use. We have also increased the number of items making up the scales, augmenting the range of values of the questionnaire and improving its discriminatory potential.

The main aim of this work was the creation and validation of a quality of life questionnaire for critical care patients.

## Materials and methods

Reliability reflects the degree to which a test score is free from measurement errors [15]. The American Psychological Association considers the following elements to be fundamental to the evaluation of a test: the identification of the main sources of measurement error, the size of such errors, the degree of expected reliability between pairs of scores in particular circumstances, and the generalization of results for items, methods, the evaluators, etc. We have established the reliability of our scale by studying the following areas, associated with possible sources of measurement error: internal consistency, inter- and intraobserver reproducibility and patient/family member and telephone/direct interview reproducibility.

The other important aspect of test evaluation is validation, i.e. that which refers to the appropriateness, significance and utility of specific inferences drawn from test scores. Validation is a process of accumulating evidence about such inferences. There are different ways of gathering this evidence, but validation as such is a unitary concept. Traditionally, it is accepted that a test is valid when it measures what it sets out to measure, and a differentiation is made between validity of content, of criteria and of construct.

Different samples of critical care patients were used to determine the different aspects of reliability and validity, except for internal consistency, factorial analysis and convergent validity, which were all based on the same sample of patients – the first 578 patients interviewed in the Project for the Epidemiological Analysis of Critical Care Patients (PAEEC). PAEEC is a multicentre study to define both the type of patients in Spanish ICUs and the therapeutic activity provided. Eighty-five ICUs from throughout Spain are participating. Details of the samples employed for each aspect are given below.

### Quality of life questionnaire

In 1991 this research team published a quality of life questionnaire specifically for critical care patients, designed to be completed by a close family member when patients were unable to respond (because of coma or the severity of their condition). It had only seven items, to allow for rapid and simple administration, and no subjective items. It evaluated oral communication, pharmacological dependence, tolerance of effort, capacity to perform precise movements, capacity for work or activity appropriate to age, sphincter control and walking.

Our current (second) questionnaire (Appendix) was designed to be an improvement over the first. We have increased the number of items to 15 that are grouped in three subscales that evaluate basic physiological activities (4 items), normal daily activities (8 items) and emotional state (3 items). The index of Independence in Activities of Daily Living, described by Katz et al. [16] was the inspiration in our design of some items.

For the construction of this new version, a committee of experts was convened, who applied the experience gained from their design and use of the previous questionnaire. They selected the aspects which they wished to be evaluated and designed items to explore the facets of quality of life which would describe those aspects. The items had to be easy to read and interpret.

For the basic physiological activities subscale, the same items used in the first questionnaire were selected, with the following modifications: in the first item, which evaluates oral communication, the number of possible responses was reduced to improve the item's reproducibility; the item measuring both diuresis and defecation in the first survey was divided into two separate parts to avoid confusion and to improve clarity. An item evaluating food intake was added to this subscale.

The subscale measuring normal daily activities included all the items from the first version, which evaluated work activity, movements of precision, tolerance of effort and ability to walk. We added further complementary items: tolerance of major effort, mobility, social relationships and ability to dress. The aim was to improve the sensitivity of this subscale and also to increase its proportional weighting in the survey. Experience with the first model had shown that a greater number of patients had problems in the normal daily activities dimension compared with basic physiological activities. The increased number of items enabled us to reflect this situation.

The committee then added a third subscale of only 3 items to analyse emotional state. These were designed to be items that a close family member would know about and they ask about the patient's subjective well-being, state of mind and vitality.

Finally, it was decided not to incorporate the pharmacological dependence item from the first version. Experience with that ver-

sion suggested that the wide range of reasons for drug prescription and use reduced the value of this item, as it reflected very different and heterogeneous patient situations.

The scores of this second questionnaire range between 0 and 29 points: the range of scores for the basic physiological activities subscale is 0–9 points, for the normal daily activities subscale 0–15 points and for the emotional state subscale 0–5 points. A score of 0 always signifies normality, with an increasing score corresponding to a worsening quality of life. Our aim was to preserve the advantages of the first version: (a) it was simple and easy to administer, taking only 5–10 min, and therefore useful in large studies; (b) it could be completed by a close family member when the patient was unable to respond; and (c) it could be administered by telephone to facilitate follow-up. However, we wanted the new questionnaire to evaluate more aspects, including subjective ones, to provide better discrimination and internal consistency. As in the first questionnaire, the second evaluates the baseline situation of the patient over the 2 months prior to admission to the ICU.

## Reliability

The analyses of reproducibility were performed by a physician trained for the purpose, who was joined by another doctor, similarly trained, for the interobserver interviews.

### *Intraobserver reproducibility*

The interviewer administered the questionnaire, on two separate occasions 1 week apart, to 150 consecutive, communicative (able to understand and respond) patients admitted to the ICU of our hospital in the second half of 1992. The mean age of the sample was 65.9±10.48 years; 89 patients (59.3%) were male.

### *Interobserver reproducibility*

Two interviewers administered the questionnaire to another group of individuals with an interval between the interviews of 1 or 2 days. These were 85 consecutive, communicative, critically ill patients admitted to the ICU in the second half of 1992. The mean age of the group was 60.62±15.51 years; 58 (68%) were male.

### *Patient/family member reproducibility*

It is often impossible to administer a questionnaire directly to a critically ill patient because of coma, mechanical ventilation, sedation or other reasons. We therefore correlated responses from the patient with those from a member of the family living with the patient. The same interviewer interviewed the patient and the family member. This study group was made up of 81 consecutive, communicative patients admitted to the ICU our hospital in the second half of 1992. The mean age was 61.85±16.8 years, and 41 (51.2%) were male.

### *Telephone/direct interview reproducibility*

In order to evaluate the use of the telephone for the follow-up of patients confined to their homes, we analysed the correlation between responses to the direct interview and to the telephone interview. The same interviewer performed both types of interview, with a 1-week interval between them. Fifty-four consecutive, communicative patients were studied in the days immediately following their discharge from the ICU, so that a telephone call a week later found them at home. The mean age of this sample was 56.44±17.62 years, and 27 (50%) were male.

### *Internal consistency*

We used indices of internal consistency to establish whether all items measured different aspects of one attribute and not different attributes. The internal consistency of the whole survey and of its subscales was analysed by Cronbach's alpha coefficient [17]. The variation of this coefficient was also measured after the elimination of each of the items of the questionnaire and of its different subscales. This study was performed with the first 578 patients interviewed in the PAEEC multicentre project. The mean age of this sample was 58.08±17.15 years, and 69.2% were men. Their severity score by APACHE II was 15.67±8.86.

## Validity

### *Content validity*

Content validity refers to a study's capacity to measure all the dimensions of the phenomenon under investigation or the degree to which an empirical measurement reflects and includes the specific areas of the problem being studied. The committee of experts was made up of doctors and nurses who specialized in intensive care and of epidemiologists. The committee modified some of the items in our first questionnaire and added subjective items. The areas explored included aspects relevant to quality of life: basic physiological activities, normal daily activities, social activities and subjective well-being. The items were designed so that family members or family doctors as well as the patients themselves could respond.

### *Construct validity*

Construct validity refers to the degree to which a measurement corresponds to theoretical expectations about it. Factorial analysis allowed us to explore the pattern of subscales initially considered for the questionnaire and to establish whether there were factors, or groups of factors, that had not been considered that might improve the questionnaire's structure. The factorial analysis was based on data from the sample used for the internal consistency analysis.

A prerequisite for the separate analysis and interpretation of the different subscales is a degree of independence between them. A high interscale correlation would make it difficult to use the questionnaire to establish different quality of life effects in distinct areas of human activity. As a flexible criterion, we fixed a maximum value of interscale correlation at $\leq 0.60$, which must be below the degreee of internal homogeneity of each subscale, as measured by Cronbach's alpha.

### *Convergent validity*

Convergent validity expresses a measurement's capacity to correlated with an external criterion (another measurement) of the phenomenon under study. We studied the concordance of the questionnaire with the Glasgow Outcome Scale (GOS) [18], which was in-

cluded in the data collection protocol of the PAEEC project for this purpose. The GOS divides quality of life into five categories: normal, limited but self-sufficient, limited and not self-sufficient, vegetative and dead.

### Responsiveness

The capacity to detect changes, along with reliability and validity, is a standard requisite of instruments able to measure subjective phenomena. The instrument must detect changes within a clinically meaningful time. To this end, we measured the level of quality of life of patients over the 2 months prior to the episode prompting their ICU admission and again 6 months after ICU discharge. The measurements were obtained using our quality of life questionnaire and the GOS. The GOS provided a point of reference for the changes detected by our questionnaire and allowed us to establish a measurement of responsiveness. We created two new variables, the comparison between preadmission and 6-month post-discharge quality of life scores by our questionnaire and the same comparison as measured by GOS. Both variables have three levels: (1) same quality of life prior to admission as 6 months after ICU discharge; (2) a better quality of life 6 months after discharge; (3) a worse quality of life 6 months after discharge. The concordance in the classification made by the two methods was established by the weighted kappa index, using quadratic weights.

### Statistical analysis

The Kolmogorov-Smirnov test of adjustment to normality revealed asymmetry in the distributions corresponding to the scores of the questionnaire's global scale and those of its different subscales. For this reason we used non-parametric methods. The comparisons between paired data were based on the Wilcoxon signed ranks test. To compare three or more independent samples we used the Kruskal-Wallis analysis of variance. For correlations between orginal data for larger samples, Spearman's rank correlation was used. Cronbach's alpha was used to determined the internal consistency of the questionnaire and its subscales. Concordance was determined by the weighted kappa index using quadratic weights. For factorial analysis, the varimax procedure was used, and statistical analysis was carried out using the SPSS/PC+ 4.0 package (SPSS, Chicago, Ill., USA).

## Results

In the sample of 578 patients studied for internal consistency and factorial analysis (the first 578 patients interviewed in the PAEEC multicentre project), the score for the quality of life questionnaire was $4.11 \pm 0.19$, and the scores on the three subscales were $0.32 \pm 0.04$ for basic physiological activities, $2.73 \pm 0.13$ for physical capacities and $1.05 \pm 0.06$ for emotional state.

### Reliability

#### *Internal consistency*

Cronbach's alpha coefficient, both non-standardized and standardized, was calculated for the global scale and for each of the subscales. The alpha coefficient was 0.85 (standardized = 0.87) for the global scale, 0.66 (standardized = 0.70) for the basic physiological activities subscale, 0.81 (standardized = 0.86) for the physical capacities subscale and 0.82 (standardized = 0.83) for the emotional state subscale.

The values of the alpha coefficient for the global scale, when each item was removed from the scale, varied between 0.824 and 0.852. There was no item whose removal caused an appreciable modification of the global internal consistency. This analysis was performed for each of the subscales, producing an identical result with respect to the internal consistency of the three subscales.

#### *Interobserver reproducibility*

Table 1 shows the means and standard deviations for each of the two interviewers in the global scale and the three subscales. There were no statistically significant differences between the means corresponding to the two interviewers, in either the global scale or any of the subscales, after the application of the Wilcoxon signed ranks test for paired samples. Spearman correlation coefficients for the two interviewers are also reported in the global scale and the subscales. All were statistically significant. There was a high correlation, $>0.90$, for the global scale and the physical capacities subscale. The emotional state subscale had a correlation of 0.77, and the basic physiological activities subscale 0.61, both statistically significant.

#### *Intraobserver reproducibility*

Table 1 also shows the data for intraobserver reproducibility. Again, there were no significant differences between the mean scores of the two questionnaires administered by the same interviewer, in either the global scale or any of the subscales. The highest correlations were again in the global scale and in the normal daily activities subscale, both of which were $>0.90$ and statistically significant. The correlations in the basic physiological activities and emotional state subscales were also statistically significant at 0.84 and 0.93, respectively.

#### *Patient/family member reproducibility*

Data for patient/family member reproducibility are also set out in Table 1, and reveal the absence of statistically

**Table 1** Analysis of interobserver, intraobserver (test-retest), patient/family member and telephone/direct interview reproducibilities with the Wilcoxon test and Spearman correlation

| | Total questionnaire | Subscale 1 (basic physiological) | Subscale 2 (normal daily) | Subscale 3 (emotional state) |
|---|---|---|---|---|
| *Interobserver reproducibility* ($n = 85$) | | | | |
| Observer 1 | 3.83 ± 4.06 | 0.38 ± 0.62 | 2.29 ± 2.87 | 1.16 ± 1.5 |
| Observer 2 | 3.95 ± 4.34 | 0.36 ± 0.59 | 2.34 ± 2.99 | 1.25 ± 1.56 |
| Wilcoxon test *p* | 0.44 | 0.84 | 0.67 | 0.48 |
| Correlation *r* | 0.92 | 0.61 | 0.91 | 0.77 |
| *Intraobserver reproducibility* ($n = 150$) | | | | |
| Test | 5.28 ± 4.1 | 0.44 ± 0.74 | 3.57 ± 2.8 | 1.28 ± 1.44 |
| Retest | 5.28 ± 4 | 0.44 ± 0.72 | 3.56 ± 2.73 | 1.29 ± 1.43 |
| Wilcoxon test *p* | 0.94 | 1 | 0.87 | 0.98 |
| Correlation *r* | 0.92 | 0.84 | 0.94 | 0.73 |
| *Patient/family member reproducibility* ($n = 81$) | | | | |
| Patient | 6.08 ± 4.74 | 0.66 ± 0.88 | 4.05 ± 3.29 | 1.37 ± 1.42 |
| Family member | 6.38 ± 5.23 | 0.76 ± 0.89 | 4.22 ± 3.67 | 1.39 ± 1.47 |
| Wilcoxon test *p* | 0.21 | 0.13 | 0.3 | 0.78 |
| Correlation *r* | 0.92 | 0.82 | 0.91 | 0.76 |
| *Telephone/direct interview reproducibility* ($n = 54$) | | | | |
| Personal interview | 4.72 ± 3.95 | 0.42 ± 0.63 | 3.18 ± 2.74 | 1.11 ± 1.38 |
| Telephone | 4.92 ± 4.25 | 0.42 ± 0.63 | 3.31 ± 2.90 | 1.18 ± 1.41 |
| Wilcoxon test *p* | 0.29 | 1 | 0.26 | 0.48 |
| Correlation *r* | 0.96 | 0.86 | 0.97 | 0.81 |

significant differences between the score for the patient and that for the family member, in either the global scale or the subscales. As in the two previous analyses, the highest correlations, >0.90, were in the global scale and the physical capacities subscale. The correlation coefficients in the basic physiological activities and emotional state subscales were 0.82 and 0.76, respectively. All the coefficients were statistically significant.

### *Telephone/direct interview reproducibility*

There were no statistically significant differences between the mean scores in the global scale or any of the subscales when direct personal interviews were compared with telephone interviews (Table 1). Once again we found that the global scale and physical capacities subscale had the highest correlations, both >0.95. The correlations in the basic physiological activities and emotional state subscales were 0.86 and 0.81, respectively. All the coefficients were statistically significant.

## Validity

### *Construct validity*

Factorial analysis confirmed the existence of three factors, which together explained 59% of total variability: the first factor explained 36.5%, the second 14.4% and the third 7.9%. The first factor related to the items on the normal daily activities subscale. The second factor related to the items on the emotional state subscale and the third to the items on the basic physiological activities subscale. Table 2 shows the factorial loading of all the items for the three factors. None of the 15 items presented secondary factorial loadings above 0.50. Secondary loadings were above 0.40 only in these three items: "work activities of those appropriate to age", with a factorial loading in factor 2 of 0.45; "dressing", factor 3: 0.49; and "movements of precision", factor 3: 0.45.

**Table 2** Factorial analysis: loading of different items on the three principal factors ($n = 578$)

| Items | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| *Normal daily activities* | | | |
| Mobility | 0.79 | 0.28 | 0.19 |
| Tolerance of effort | 0.75 | 0.28 | 0.02 |
| Walking | 0.67 | −0.005 | 0.37 |
| Dressing | 0.62 | 0.008 | 0.49 |
| Major effort | 0.62 | 0.29 | −0.04 |
| Work activities | 0.60 | 0.45 | 0.11 |
| Movements of precision | 0.53 | −0.09 | 0.45 |
| Social relationships | 0.52 | 0.34 | 0.29 |
| *Emotional state* | | | |
| State of mind | 0.17 | 0.84 | 0.1 |
| Vitality | 0.15 | 0.83 | 0.04 |
| Subjective well-being | 0.23 | 0.81 | 0.02 |
| *Basic physiological activities* | | | |
| Urination control | 0.003 | 0.12 | 0.77 |
| Defecation control | 0.02 | 0.04 | 0.76 |
| Oral communication | 0.26 | 0.02 | 0.65 |
| Intake of food | 0.31 | 0.07 | 0.56 |

**Table 3** Convergent validity. Relationship between the Glasgow Outcome Scale categories and the total score/score of the different subscales ($n = 578$)

| | No. | Score total | Subscale 1 (basic physiological) | Subscale 2 (normal daily) | Subscale 3 (emotional state) |
|---|---|---|---|---|---|
| Normal | 295 | 0.95 ± 1.62 | 0.07 ± 0.3 | 0.58 ± 1.23 | 0.29 ± 0.75 |
| Limited self-sufficient | 183 | 5.28 ± 2.73 | 0.26 ± 0.61 | 3.42 ± 1.98 | 1.60 ± 1.37 |
| Not self-sufficient | 100 | 11.29 ± 4.38 | 1.18 ± 1.71 | 7.81 ± 3.04 | 2.30 ± 1.77 |
| $p$* | | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |

*$p < 0.0001$ (statistically significant differences between the GOS categories and the total score/score of each of the subscales. Kruskal-Wallis test)

*Convergence validity*

Convergence validity was studied by analysing the relationship of the scores in the total questionnaires and in the different subscales with the classifications assigned by the GOS. Table 3 shows the results, and it can be seen that as the quality of life measured by GOS worsened, so did quality of life as measured by the questionnaire and by its subscales. The differences between the scores were statistically significant in all cases.

Responsiveness

Concerning the capacity to detect changes, both the GOS and our questionnaire revealed a worsening of quality of life in patients 6 months after ICU discharge, in the global scale and in each of the subscales. When the magnitude of the changes detected by our questionnaire were compared with the GOS results, we found a weighted kappa index of 0.56, statistically significant at $p < 0.001$.

## Discussion

We have validated a quality of life questionnaire for critically ill hospital patients based on a previous questionnaire created by this research team [13, 14]. Quality of life is a complex, multidimensional concept which includes different facets such as functional capacity, physiological functions and subjective aspects. We included the principal aspects of quality of life [19] grouped into three dimensions: basic physiological activities, normal daily activities and emotional state.

A number of methods have been used in intensive care medicine to measure quality of life [1, 18, 20–22]. Most are simple systems for classifying patients into a few categories (from 3 to 6) between normality and total incapacity. The Glasgow Outcome Scale [18] and the APACHE classification by Knauss et al. [1] are two examples of such systems.

The fundamental problem with these instruments is the low sensitivity derived from the small number of categories. Another key point for any quality of life instrument is the time needed for its completion, which should be as short as possible. Questionnaires requiring considerable time to complete, or which only the patient can respond to, are not suitable for use at the time of ICU admission, when medical staff are busy and patients may be unable to communicate because of the severity of their illness. It is also very possible that the patients' state on admission is very different from their prior baseline condition. Our questionnaire is specifically designed for use in follow-up studies and in the first 24 h after admission; it requires little time for completion, and evaluates a patient's stable situation over the previous 2 months; it can be administered to a close family member and by telephone if necessary.

Reliability

We used a multiple approach to establish reliability, studying both internal consistency and reproducibility, measured for direct personal interview of the patient, telephone interview and surrogate interview. We also determined intraobserver and interobserver reproducibility. The questionnaire was designed so that items are completed as responses to questions posed by the doctor, either at the hospital bedside or by telephone to facilitate patient follow-up at home. We tested the correlation between patient interview and surrogate interview scores, as patients are sometimes unable to communicate.

The questionnaire as a whole demonstrated a high degree of reproducibility, and the correlation coefficient was $> 0.90$ in all the procedures used to complete it. The questionnaire's global scale had high internal consistency. When we analysed the influence of each item, none significantly modified the global consistency.

We tested the reliability of each of the subscales in the same way. In the normal daily activities subscale, all reproducibility modes had a consistently high correlation index of $> 0.90$, and internal consistency was also high. The correlation in the basic physiological activities scale was always $< 0.82$, except for interobserver reliability, which had a moderate value of 0.61. The internal con-

sistency of this scale was acceptable. The emotional state subscale had a range of coefficients between 0.73 and 0.81, which represents moderate to high reliability, meeting our objective that these items could be answered by a family member. This subscale also showed an acceptable level of internal consistency. As emotional state items cannot be evaluated in vegetative patients, we assigned them maximum scores (worst quality of life).

### Validity

For the questionnaire's validity, we distinguished between validity of content, of criteria and of construct. These are not different types of validity but constitute distinct dimensions of the unitary concept of validity. The validation of a questionnaire is a continuous process incorporating additional evidence either in favour of or against the validity of the questionnaire. It is not the result of complex analysis, but comes from the use of the questionnaire in a large number of patients in different situations with comparable results.

In this first study, the information is favourable to the validity of our questionnaire. Factorial analysis confirmed the existence of three subscales that are the fundamental components of this questionnaire, consistent with the areas theoretically proposed from the outset and from which the items were designed and selected. None of the 15 items presented secondary factorial loadings above 0.50 and only three items had secondary loadings above 0.40.

There is no gold standard for evaluating the goodness of fit of quality of life scales, because of the nature of the reality under study. The validity of the criteria was established by using the GOS [18]. We discovered a good, positive and significant association between GOS scores and our questionnaire scores. This association was also found with the different subscales.

### Responsiveness

The assessment of the questionnaire's capacity to detect clinically relevant changes in levels of quality of life over time (responsiveness) is also a compelx process requiring evidence from multiple studies. This first study provides information in favour of our instrument's capacity to detect such changes. This evidence is critically dependent on the capacity of the GOS to detect changes, as this is the instrument used as our reference.

## Conclusion

Our study has demonstrated that our questionnaire fulfills the essential requisites of validity and reproducibility when applied to critical care patients admitted to ICUs. This is the population for which it was created and validated. The questionnaire responds to the specific conditions of this area of medicine. It permits the retrospective evaluation of the patients' stable situation. It is designed to be completed easily and quickly, despite evaluating various aspects of quality of life, and some of the items form part of the questions routinely asked of any critically ill patient for their case history. This facilitates its use with large numbers of patients, such as in the PAEEC project and in major surveys to be carried out in the future.

**Acknowledgement** The authors thank Richard Davies for his translation of the text into English.

## References

1. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE (1981) APACHE – Acute Physiology and Chronic Health Evaluation: a physiologically based classification system. Crit Care Med 9:591–597
2. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) APACHE II: a severity of disease classification system. Crit Care Med 13:818–829
3. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A, Harrell FE (1991) The APACHE II prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 100:1619–1636
4. Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Villers D (1984) A simplified acute physiology score for ICU patients. Crit Care Med 12: 975–977
5. Le Gall JR, Lemeshow S, Saulnier F (1993) A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA 270:2957–2963
6. Teres D, Lemeshow S, Aurunin JS, Pastides H (1987) Validation of the mortality prediction model for ICU patients. Crit Care Med 15:208–212
7. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J (1993) Mortality probability models (MPM II) based on an international cohort of intensive care unit patients. JAMA 270: 2478–2486
8. 2nd European Consensus Conference on Intensive Care Medicine (1994) Predicting outcome in ICU patients. Intensive Care Med 20:390–397
9. Spanish Group for Epidemiological Analysis of Critical Patients (PAEEC) (1994) Quality of life: a tool for decision-making in the ICU. Intensive Care Med 20:251–252
10. Ikegami N (1995) Functional assessment and its place in health care. N Engl J Med 322:598–599

11. Katz S (1987) The science of quality of life. J Chron Dis 40:459–463
12. Bowling A (1992) Measuring health: a review of quality of life measurement scales. Open University Press, Philadelphia
13. Vázquez Mata R, Rivera Fernández R, González Carmona A, Delgado-Rodriguez M, Torres Ruiz JM, Raya Pugnaire A, Aguayo de Hoyos E (1992) Factors related to quality of life 12 months after discharge from an intensive care unit. Crit Care Med 20: 1257–1262
14. Rivera Fernández R, Vázquez Mata G, González Carmona A, Torres Ruiz JM, Delgado-Rodríguez M, Abizanda Campos R, Raya Puignaire A (1991) Descripción de una encuesta de calidad de vida en Medicina Intensiva. Med Intensiva 15:313–318
15. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1985) Standards for educational and psychological testing. American Psychological Association, Washington, DC
16. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW (1963) Studies of illness in the aged. JAMA 185:914–919
17. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. Psychometrika 16:297–334
18. Jenett B, Bond M (1975) Assessment of outcome after severe brain damage. Lancet I:480–484
19. Walker SR, Rosser RM (1988) Quality of life: assessment and application. MTP Press, Lancaster, UK
20. Bergner M, Bobbit RA, Carter MB, Gilson BS (1982) The sickness impact profile: development and final revision of a health status measure. Med Care 19:787–805
21. Patrick DL, Danis M, Southerland LI, Hong G (1988) Quality of life following intensive care. J Gen Intern Med 3: 218–223
22. Sage WM, Myer HR, Silverman JF (1986) Is intensive care worth it? An assessment of input and outcome for the critically ill. Crit Care Med 14:777–782

## Appendix: Quality of life questionnaire, second version

Is the patient orientated in time and space? Yes/No
If the patient cannot respond, who responds? .....

Subscale 1: Basic physiological activities

*Item 1. Oral communication*

Is there difficulty with speech?
0– No
1– Yes, a dialogue can be held, but with difficulty, due to: softness of voice or faltering speech; short phrases or single words (monosyllables)
2– Yes, incoherent speech, either sporadic or permanent
3– Yes, no oral communication

*Item 2. Urination control*

Is there difficulty in controlling urination?
0– No
1– Yes, sporadically, due to: urinary retention; incontinence
2– Yes, permanently, due to: need for urinary catheter or bags

*Item 3. Defecation control*

Is there difficulty in controlling defecation?
0– No
1– Yes, due to: ileostomy/colostomy; sporadic incontinence; persistent constipation requiring regular enemas
2– Yes, permanent incontinence

*Item 4. Intake of foods*

Is there difficulty in ingesting food?
0– No
1– Yes. Although feeding is by the mouth, there are difficulties due to: difficulty in swallowing; difficulty in handling cutlery; lack of appetite, rejection of food
2– Yes. No feeding by mouth and requires: nutrition by nasogastric tube; parenteral nutrition

Subscale 2: Normal daily activities

*Item 5. Movements of precision*

Can the following movements be carried out?

| | | |
|---|---|---|
| Pouring a glass of water | Yes | No |
| Using cutlery | Yes | No |
| Using a key | Yes | No |
| Shaving | Yes | No |
| Sewing | Yes | No |

0– Can carry out more than 3 activities
1– Cannot carry out more than 3 activities

*Item 6. Tolerance of minor effort*

Can the following activities be carried out?

| | | |
|---|---|---|
| Go up one floor | Yes | No |
| Walk two blocks | Yes | No |
| Carry a full shopping bag, a full brief-case or similar | Yes | No |

0– Carries out these activities in the same way as before
1– Cannot carry out one or more of these activities now, but could before

*Item 7. Tolerance of major effort*

Can the following activities be carried out?

| | | |
|---|---|---|
| Walking 5 kilometres | Yes | No |
| Running 1 kilometre | Yes | No |
| Going up 4 floors without shopping | Yes | No |
| Practising a sport requiring a high level of physical effort such as football, tennis, swimming, or similar | Yes | No |

0– Can carry out at least one of these activities
1– Can carry out none of these activities

*Item 8. Walking*

Is there difficulty with walking?
0– No
1– Yes, walks with help (crutch or people)
2– Yes, does not walk, and uses wheelchair
3– Yes, is permanently bedridden and depends on others

*Item 9. Mobility*

Is there difficulty in making normal journeys?
0– No, can make all the normal journeys
1– Yes, only moves about the immediate locality
2– Yes, only moves about the house
3– Yes remains in his/her room

*Item 10. Dressing*

Does the patient have difficulty in getting dressed?
0– No
1– Yes, need some help
2– Yes, and is totally dependent on others

*Item 11. Work activities or activities appropriate to age*

Are there difficulties with the patient's work?
0– No
1– Yes. He/she has difficulties, but works as before
2– Yes. Works only part-time or has changed to a job requiring minimum effort
3– Does not work because of his/her condition

For patients retired due to age, the question is: are there difficulties with the patient's activities as a retired person?
0– No, continues with regular scheduled activities
1– Yes, continues with regular scheduled activities but with difficulty
2– Yes, activities are no longer regular, and are only sporadic, or have been changed for alternatives requiring less activity
3– Yes, and has completely abandoned them

*Item 12. Social relationships*

Are there difficulties with social relationships?
0– No, they continue normally within and outside the family
1– Yes, and relationships only continue within the family environment

Subscale 3: Emotional state

*Item 13. Subjective well-being*

How do you normally feel?
0– Fine
1– Not too bad
2– Bad

*Item 14. State of mind*

How do you normally feel?
0– I am happy
1– I am sad or I feel like crying some days every week (less than 3 days)
2– I am sad or I feel like crying every day of the week (4 or more days)

*Item 15. Vitality*

Do you feel like initiating activities?
0– I feel like doing things/starting activities
1– I don't feel like doing anything/nothing interests me