Cancer Treatment Reviews (1996) **22** (*Supplement A*), 75–89

Practical aspects of quality-of-life measurement: design and feasibility study of the quality-of-life recorder and the standardized measurement of quality of life in an outpatient clinic

Jörg Sigle and Franz Porzsolt

Cancer Centre of the University of Ulm, Oberer Eselsberg, 89081 Ulm, Germany

Introduction

Although 'Quality of Life' (QL) has gained importance in recent years, few articles on this subject have been reviewed, showing that there is no broad consensus about the interpretation of the term nor about the methods applied for QL measurement. The present work reports practical experiences made with the QL recorder developed by the authors. This instrument makes validated QL tests available for use as an everyday standardized routine measure. This work aims to support QL measurement as a common element in research and practice for investigation and documentation of the success of these measures.

The development of the QL recorder included four phases: construction and acceptance test of the QL recorder; investigation of acceptance and additional personnel and management requirements; collection of representative data from a large patient sample in a standardized procedure; and statistical analysis of the data obtained.

As a report on the whole project can be found elsewhere (1), this article will focus on Phases III and IV of the project. Theoretical aspects of QL measurement are discussed elsewhere (2–6).

Methods

Collection of representative QL data

Special requirements in the outpatient clinic In a university's outpatient clinic, a large number of patients are distributed to a certain number of physicians and examination facilities. Reliable inclusion of all patients into a study not only requires enough hardware equipment, but also depends on adequate organization of patient recruitment. The new, temporary examination has to be integrated smoothly into established structures and processes.

This supplement was sponsored by Boehringer, Ingelheim.

0305–7372/96/22A075 + 15 $12.00/0

Hardware and software In a designated room located next to the reception desk of the outpatient clinic, six QL recorders were installed using standard PCs in a network under Novell Netware 3.11. Each QL recorder ran MS-DOS, GraTaSim V. 2.91 and QLQC-33 V. 1.2. (1, 7, 8).

Supporting measures In Phase II, the completeness of data in a QL assessment was found to depend primarily on the management of the assessment and on the involved personnel's cooperation. It could not be guaranteed that each patient had visited the desk before each consultation. To cope with this situation, posters were displayed throughout the outpatient clinic asking patients to visit the QL measurement room prior to each consultation. Nurses and physicians were asked to send patients who had no printout of QL measurement to the measurement room. A stock of papers for patient information were provided in the QL measurement room and in the waiting areas. In the QL measurement room, an additional 50 copies of the paper version of the QLQ-C30 (+3) were provided for those patients who would not agree to an electronic assessment but would fill in a paper questionnaire.

Personnel During the sampling period, two persons from 8:00 to 10:00 a.m. and a single person for the rest of the day asked incoming patients to fill in the questionnaire, gave short instructions, and stayed nearby to help if any difficulties should occur. Any problems were documented as well as patients' reasons for refusing to participate. The last attendant of each day performed the measures for quality assurance described below.

Inclusion and exclusion criteria: patient selection The QL measurement was defined as an 'examination essential for each patient who did not meet the exclusion criteria'.

Exclusion criteria were: emergency; patient was bedridden or could not visit the QL measurement room for other reasons; patient did not speak German; visual problems; patient was already hospitalized; patient visited the Automatic-Implantable-Cardioverter/Defibrillator (AICD) clinic only; or patient was a child.

Inpatients would visit the outpatient area for several special examinations, for example for ultrasonic examination of the thyroid gland. All patients visiting the AICD check were excluded *a priori* because the additional QL measurement would disturb the course of the scheduled examinations and this could not be imposed on the patients. Patients who did not agree to participate in the study were asked to present themselves in the QL measurement room for registration of their initials and date of birth. As the QLQ-C30 (+3) defines a time-frame of 1 week, only one measurement was scheduled for each patient during the short assessment period of the study.

Chronological integration of the QL measurement Most patients had to wait for a period of time after presentation at the desk. This period of time was considered optimal for the QL measurement as it would not mean additional length of stay for the patient. When patients entered the outpatient clinic, posters asked them to visit the QL measurement room before their consultation. If they had no QL printout, they were sent to the measurement room from the

reception desk, from the blood-taking room, from several special examinations, or, finally, by the physician after the consultation.

Patients' interaction All patients were asked to complete the QLQ-C30 for study purposes. Participating patients were instructed how to use the QL recorder and completed the electronic version of the EORTC QLQ-C30 (+3). Immediately afterwards, they received a printout showing patient ID, time and duration of the assessment, and the results of the QLQ-C30 (+3), numerically as well as in a simple graph. If a patient declined to take part in the study, his ID and reason for disagreement were recorded. The patient received a short note confirming registration in the QL programme.

Quality assurance The first stage of quality assurance was imposed by all personnel; all coworkers were asked to send patients without a QL printout to the QL measurement room. For the second stage, registered patients were compared to those who had been expected, according to the clinic's diary, at the end of each day. If missing patients were reported, the responsible specialty was identified and it was ascertained whether a patient had cancelled his appointment, or whether he had not been sent to the QL measurement facility.

In the third stage of quality assurance, all patient identifications recorded in the QLQ files with redundant entries removed were compared to all entries in the clinic's diary at the end of the study. When this comparison was finished, several QLQ files had no matching entry in the clinic's diary. Therefore, several weeks later, the clinic's diary was checked to see whether matching appointments had been made in the meantime and the comparison was repeated.

Finally, information was received about the number of patients in the outpatient clinic from the clinic's data centre. For each department, the number and proportion of patients reported by the data centre was compared to the number and proportion of registered patients. Moreover, the data centre computed the total number of patients who visited the outpatient clinic during the assessment period. This number was compared with the number of matched QLQ files.

Data protection A study involving electronic data processing is required to meet substantially higher requirements in data protection than everyday clinical tools. Quality-of-life measurement was located in a separate room which was locked in off-service hours, or continuously supervised by at least one staff member. The fileserver was password-protected and did not have a monitor or keyboard for security reasons. Patients' printouts were treated like other examination reports. All information was entered directly into the computer. Access rights were set to write-only except for the required programmes. As no connection to other networks existed, information was sealed in this network as soon as the patient ticked the 'Finished' field on the questionnaire. The enhanced access rights for the daily quality assurance were password-protected and only available in a small time-window from 1 h before to 1 h after the end of service hours.

Staff poll About 4 weeks after completion of the study, a poll was taken to assess the staff's opinion about the additional work load for personnel and patients. Physicians were asked to rate the importance of the QL measurement. The questionnaire included six visual-analogue scales. Results were plotted and examined for correlation using Spearman's Rho.

Organizational experiences and cost calculation Patient flow was examined using data which had been recorded automatically. Results may be important in the design of new QL facilities. Also, costs that might be expected for a routine QL measurement in a patient population comparable to that reported were calculated.

Data analysis

Only those QLQ files which could be matched to an entry in the clinic's diary were used for exploratory data analysis. During the assessment period, a number of students had a medical examination because of entrance to their final year internship (*'Praktisches Jahr', PJ*). They were treated as a control group.

Besides commercially available software (MS Excel and SPSS), software was programmed according to the requirements for the matching of patients, separation of patients outside of SPSS, and for all other special tasks.

Psychometric analysis Values were computed which have already been used in the validation process of the paper version of the QLQ-C30. The results were compared with those found in the literature to search for an obvious influence of the electronic QL recorder on the questionnaire results.

Analysis of distances and cluster analysis As part of the exploratory data analysis, different methods of hierarchical and k-means cluster analysis were applied to check whether or not, and how well, different sub-populations could be distinguished.

Cluster analysis is a method which tries to generate clusters of cases of a sample. Cases within a cluster should be as similar as possible while the clusters should be as different as possible. The first step of a hierarchical cluster analysis is to compute the distances between all patients' results. First generation clusters are made from data sets which are located most closely together. Consecutively, clusters are grouped together until all cases are agglomerated into one cluster. A report of the whole process of clustering can be plotted as a tree diagram which enables the user to identify particularly distinct clusters. A simplified method of cluster analysis is k-means cluster analysis. Here, the number of clusters to be formed is pre-determined. Second, preliminary centres of the final clusters have to be specified. They are modified during the process of the analysis.

With a sample of over a thousand patients, a hierarchical cluster analysis requires a great deal of computation time and excessive storage capacity, while the resulting plot is probably difficult to interpret. For this reason, the clustering

process was not applied completely. Instead, the distances between all multi-dimensional QL vectors of all patients were computed. They were examined in a distance histogram. This method was applied using vectors of all 18 dimensions or using vectors of a few selected dimensions which were expected to distinguish the specialties.

To find whether results from patients of the same specialty would be similar, all distances between all possible pairs of patients' QL vectors (about 624,000) were sorted in ascending order. For the n closest and the n most distant pairs of QL vectors, it was checked whether both patients of the pair belonged to the same specialty.

Characteristics of different specialties To identify differences in QL among outpatients, the results of each specialty were compared with a control group. This control group consisted of all included patients except for the PJ students and the patients of this particular specialty. Observed characteristics of the subgroups were compared with expected clinical pictures. For this analysis, the arithmetic means of the sub-populations were compared, and the two-sided significance level for the difference of means was computed using the non-parametric Mann–Whitney U-test.

Results

Descriptive statistics

Assessment period, number of patients, assessment rates Patients were included into the Ambu2-study for 19 days. During this period of time, 1315 patients were recorded in the clinic's diary, including appointments that had been entered up to 4 weeks after the end of the study. Of these, 101 patients had to be excluded and 33 were inpatients. The resulting number of patients in the target group was 1181.

During the same time period, 1156 electronically completed non-redundant QLQs and one QLQ completed on paper were obtained. Of these, 1120 had a matching entry in the outpatient clinic's diary. Thirteen patients did not agree to take part in the study.

In the worst case, completely ignoring the QLQ files that could not be reliably matched with entries in the clinic's diary, the resulting registered proportion of the target group was 1133/1181 (95.9%) and the proportion of the target group represented by completed questionnaires was 1120/1181 (94.8%).

The data of the 1120 patients with a matching QLQ file and entry in the clinic's diary served as database for further analysis.

Age distribution The age distribution of the patients in the study resembles the German demographic curve except for very young children (Figure 1). Children are only admitted from the childrens' hospital to the authors' hospital for special diagnostic procedures. Besides, the QLQ-C30 was not designed for young children.

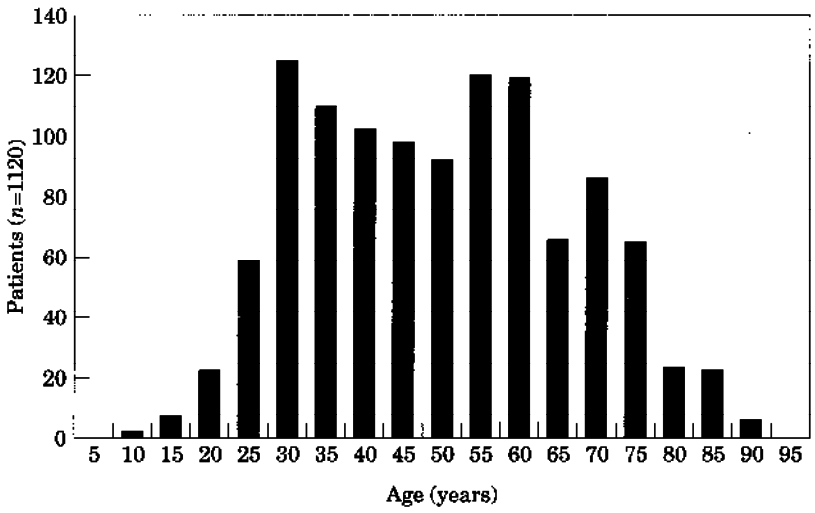

Figure 1. Age distribution of included patients. This histogram shows that the age distribution of the included patients resembles the German demographic curve. Typical incisions can be seen at 50, 65 and 80 years in intervals of about 15 years. Very young children were excluded.

Compliance and completeness of the questionnaires Of 1133 registered patients, only 13 did not agree to take part in the study while 1120 agreed. The resulting compliance was 98.9%. Generally, patients who did not agree to participate, did not even visit the room of QL measurement. The most frequent reason for not participating was being afraid of delaying one's appointment. An elderly patient was afraid of disadvantages resulting from giving 'wrong' answers. One patient who did not want to participate because of a 'general suspicion against questionnaires', changed his mind and completed all questions very carefully. He returned 30 min later and wanted to correct some of his answers after he had studied the printout.

The completeness of questionnaires was greater than 99.96%. Missing data did not result from overlooked questions, but some patients said they could not answer single questions.

Distribution of excluded patients The frequencies of patients meeting the different exclusion criteria are shown in Table 1. Almost 70% of the exclusions

Table 1. Reasons for patient exclusion

Reason	Absolute	Relative (%)
Language problem	44	38.60
Automatic-Implantable-Cardioverter/Defibrillator	35	30.70
Patient is bedridden/in too bad condition to fill in the questionnaire	11	9.60
Forgotten glasses/visual problem	10	8.80
Emergency case	5	4.40
Children	5	4.40
Other reason	4	3.50
Total	114	100

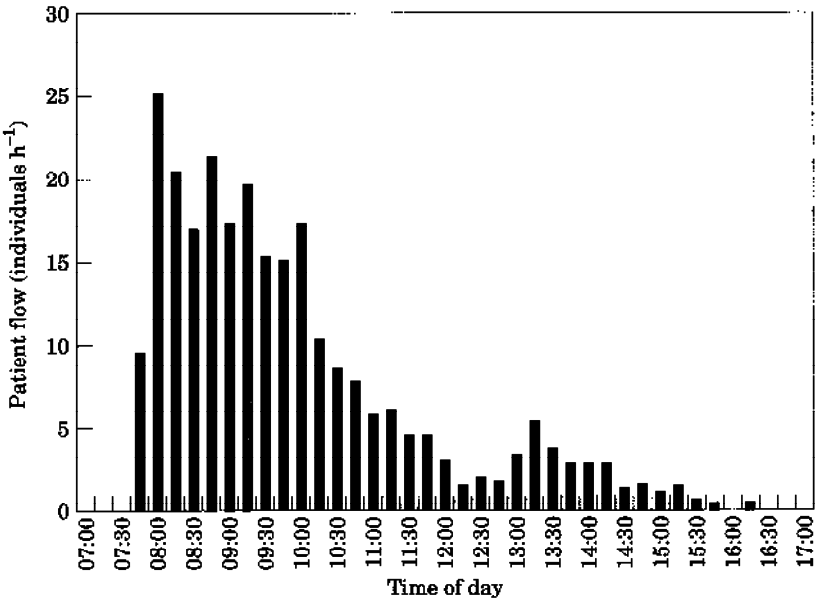

Figure 2. Patient flow *vs* time of day. Patient flow given as individuals per hour.

occurred due to language problems or because the infrastructure of the AICD clinic, which they visited, was not compatible with the integration of QL measurement. As a single patient may meet more than one criterion, the total number of exclusions was 114.

Evaluation of the staff poll The staff poll showed that most staff members thought they were well informed. They thought that QL assessment was important, that the measurement was no major burden for the patient and that their own additional work load was rather small. The authors found that people who thought they were well informed usually talked to patients more ($\rho = 0.63$, $p = 0.003$) and thought that QL measurement was more important ($\rho = 0.44$, $p = 0.045$) than those who thought they were not well informed. Interestingly, the physicians' answer 'I paid attention to the QL results in the patient file' was found to correlate with 'I think QL measurement is important' ($\rho = 0.75$, $p = 0.048$) and correlated negatively with 'I had additional work load from the study' ($\rho = -0.72$, $p = 0.055$).

Integration of the study in the outpatient clinic's environment

Patient flow varies over the day The frequency of patient visits to the QL measurement room *vs* time of day is shown in Figure 2. It shows that there are peak values of 25 patients per hour in the morning and a decrease to only 1/5 to 1/10 of this flow in the afternoon. This phenomenon influenced the QL recorders' system load; while four to six systems were needed around

Table 2. Costs of quality-of-life measurement

Item	Price	Factor	Total costs	
Investment costs for hard- and software equipment				
QLQ recorder, digitizer+Gra TaSim-software	2.200 DM	6	13.200 dm	$9.429
PC, at least 80286, network adapter, mono monitor	1.500 DM	6	9.000 DM	$6.429
Fileserver, 80486, HDU, network adapter, Novell Netware	2.500 DM	1	2.500 DM	$1.786
Ink-jet printer	500 DM	1	500 DM	$357
Total investment costs			25.200 DM	$18.000
Required material per month				
Ink-refills	50 DM	4	200 DM	$143
Paper, 500 pages	10 DM	3	30 DM	$21
Electric energy	50 DM	1	50 DM	$36
Manpower per month				
Clinical Data Manager	3.000 DM	1	3.000 DM	$2.143
Regular costs per month			3.280 DM	$2.343
Total costs over a period of 5 years			222.000 DM	$158.571
Possible QL measurements (at 1315/19 days) during 5 years				90,000
Upper limit for the costs of a single QL-measurement			2.47 DM	$1.76

The calculation includes a network with six QL recorders and all required PC equipment. Infrastructure like room and furniture is excluded.

8:00–10:00 a.m., a single system was usually sufficient in the afternoon. With an average completion time of 5.5 min, only very small delays occurred.

Cost and personnel requirements A cost calculation for 5 years of QL measurement is shown in Table 2. Considering 90,000 QL measurements during a period of 5 years, the total costs for a single assessment were calculated to be below US $2.

This calculation does not include costs for planning a study or for evaluation of the results.

Data analysis

Descriptive statistics, reliability and inter-scale correlations All EORTC-C30 (+3) dimensions yield results of 0–100, but with different resolutions.

The results of descriptive statistics and reliability for the complete sample are shown in Table 3. The results use the complete possible range in all dimensions. The quartiles show that dimensions assessing function are left-skewed while symptoms are right-skewed. Cronbach's α is greater than 0.7 in most scales and still around 0.6 in KF (0.66), physical function (0.62) and nausea/vomiting (0.57). Only the role function scale has an α of 0.36.

For a comparison with the literature (2), inter-scale correlations according to Pearson were calculated. The following results were found (all ρ values <0.005). Correlations between function scales ranging from 0.27–0.7. The strongest correlations were found between role function and new role function as well as social function and new role function. Correlations between global measures

Table 3. Descriptive statistics and reliability analysis

Scale	*n*	Mean	S.D.	Min	25%	50%	75%	Max	Items	Cronbachs α
PF	1120	83.02	21.29	0	60	100	100	100	1–5	0.62
RF	1120	77.93	30.79	0	50	100	100	100	6, 7	0.36
NRF	1120	70.98	32.56	0	50	83	100	100	26, 27	0.87
EF	1119	63.74	27.28	0	42	67	83	100	21–24	0.85
CF	1118	79.51	24.16	0	67	83	100	100	20, 25	0.66
SF	1118	73.51	30.19	0	50	83	100	100	23, 29	0.84
QL	1119	58.67	24.53	0	42	58	83	100	31, 33	0.89
NQL	1119	56.61	25.88	0	33	58	75	100	32, 33	0.91
FA	1120	37.82	29.32	0	11	33	56	100	18, 12, 10	0.86
NV	1120	9.88	18.28	0	0	0	17	100	14, 15	0.57
PA	1120	32.93	33.48	0	0	17	67	100	9, 19	0.88
DY	1119	23.33	29.10	0	0	0	33	100	8	
SD	1120	32.59	35.42	0	0	33	67	100	11	
AL	1120	13.94	26.02	0	0	0	33	100	13	
OB	1120	9.05	21.78	0	0	0	0	100	16	
DI	1120	13.52	25.90	0	0	0	33	100	17	
FI	1120	16.08	29.64	0	0	0	33	100	30	

S.D., standard deviation; 25%, 50%, 75%, quartiles. EORTC QLQ-C30 (+3) Dimensions: PF, physical function; RF, role function; NRF, new role function; EF, emotional function; CF, cognitive function; SF, social function; QL, general health/global quality of life; NQL, new general health/global quality of life; FA, fatigue; NV, nausea and vomiting; PA, pain; DY, dyspnoea; SD, sleep disturbance; AL, appetite loss; OB, obstipation; DI, diarrhoea; FI, financial impact.

ranged from 0.29 (with symptoms) to 0.65 (with functions) with the strongest correlations between general health and new general health. Correlations between symptom scales ranged from 0.04 to 0.52, preferring the area around 0.3. Fatigue showed the strongest correlations with symptoms (around 0.3) as well as with functions (around 0.6). Obstipation, diarrhoea and financial impact only showed correlations in the range of 0.1–0.2.

Analysis of distances and cluster analysis The distances between PJ students were short even when all 18 dimensions were included, and in a distance histogram they appeared as a sharp peak. Excluding the PJ students, other subgroups of patients only produced broader peaks when vectors constituent of a small number of dimensions were used. If additional dimensions were added, these peaks melted into a normal distribution. For example, a peak resulting mainly from cardiac and pulmonologic patients could be distinguished well using vectors of age, physical function and dyspnoea, while it disappeared in a distance histogram of vectors of all 18 dimensions.

These results indicated that the postulated clusters would be distributed over all the space available with large overlapping. This was expected as the resolution of the dimensions of the QLQ-C30 (+3) is quite low. Because of this, in 1120 cases there is a high probability that each sub-population's results cover the whole available range in each dimension.

Using several values for n in the range of 10–1000, it was observed that in every case there were more pairs of patients from the same specialty among the n closest pairs of vectors than among the n most distant pairs of vectors.

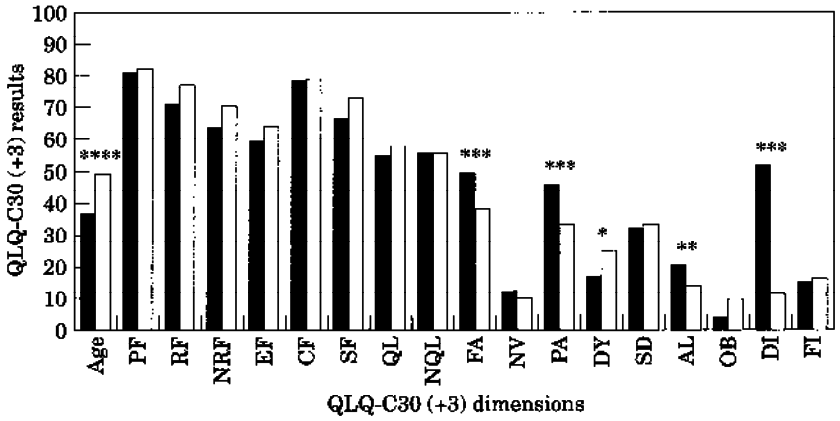

Figure 3. QLQ-C30 (+3) results (see Table 3 for definitions) of M. Crohn/Colitis Ulcerosa patients. Significance level of Mann–Whitney test: *$p<0.05$; **$p<0.01$; ***$p<0.005$; ****$p<0.001$. $n=59$ *vs* 1008. Solid bars, M. Crohn/Colitis Ulcerosa patients; open bars, other patients.

As a hierarchical cluster analysis requires a great deal of computation time and excessive storage capacity, and as the results of the grouping process in such dimensions were not expected to be interpretable, a complete automatic grouping was not run. Instead, a k-means cluster analysis was performed. This shortened method, depending on the initial values, could identify clusters, but it could not generate any clusters which exactly resembled groups of patients from different specialties.

Search for characteristics of groups sorted by specialties

For some selected specialties, the following diagrams show how their results differ from all other patients.

M. Crohn/Colitis Ulcerosa This sub-population (Figure 3) differs from other patients by lower average age, more fatigue, more pain, more appetite loss and more diarrhoea. These characteristics fit the clinical picture of both diseases, i.e. an age peak between 20 and 40 years, inflammatory disease of the intestine with abdominal pain, diarrhoea and fatigue.

Cardiologic patients This sub-population (Figure 4) differs from other patients by higher average age (around 60 years), worse physical function, lower global quality of life and more dyspnoea. These findings reproduce the clinical picture of cardiologic patients.

Pulmonologic patients This sub-population (Figure 5) differs from all other patients by slightly higher average age, reduced emotional function, reduced social function, fatigue, appetite loss and obstipation. Differences in physical function, dyspnoea and pain are not significant, but a larger sample might lead to a better characterization. The reduction of emotional function could be explained by anxiety caused by episodes of asthma. Appetite loss is a common

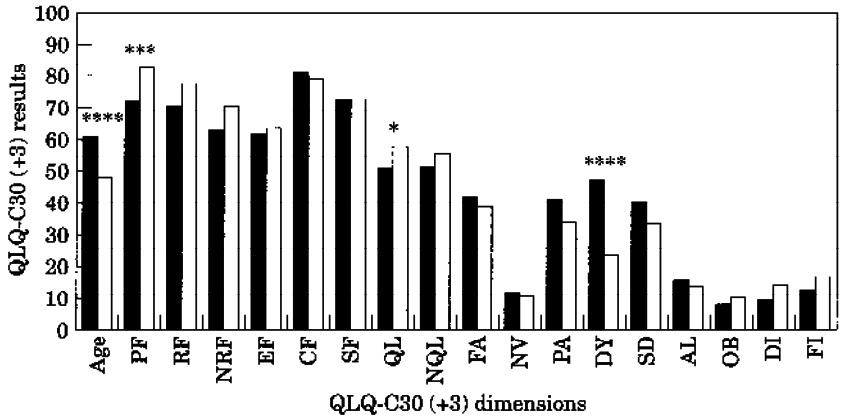

Figure 4. QLQ-C30 (+3) results (see Table 3 for definitions) of cardiology patients. Significance level of Mann–Whitney test: *$p<0.05$; **$p<0.01$; ***$p<0.005$; ****$p<0.001$. $n=46$ *vs* 1021. Solid bars, cardiology patients; open bars, other patients.

symptom in pulmonary diseases. Obstipation is one of the most obvious differences between pulmonologic and cardiologic patients. It might be explained as a side-effect of the medication for gastric ulcer prophylaxis during corticoid therapy.

Bone marrow transplant patients (BMT) This sub-population (Figure 6) differs from all other patients by lower average age (around 37 years), reduced role function, better emotional function, better cognitive function, better global quality of life and more financial difficulties. The lower average age could be explained by the lower transplantation risk of younger patients which enables these patients to receive this treatment more frequently. Differences in emotional function, cognitive function and global quality of life might be caused by the selection process associated with the procedure of BMT and/or by special emotional support received by these patients.

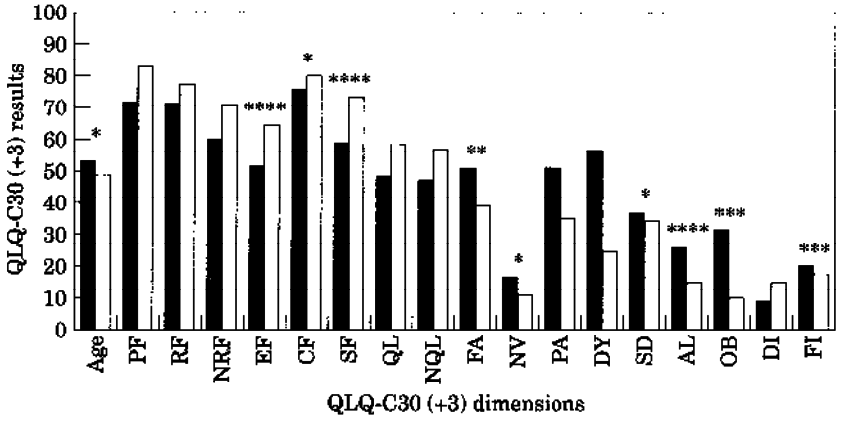

Figure 5. QLQ-C30 (+3) results (see Table 3 for definitions) of pulmonary patients. Significance level of Mann–Whitney test: *$p<0.05$; **$p<0.01$; ***$p<0.005$; ****$p<0.001$. $n=12$ *vs* 1055. Solid bars, pulmonology patients; open bars, other patients.

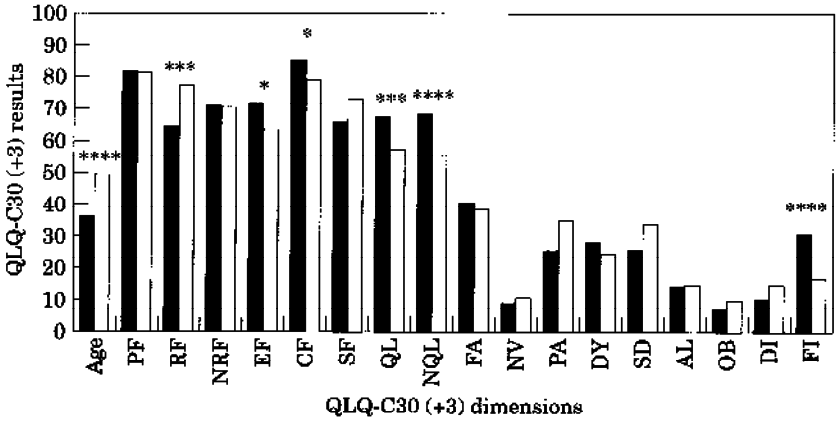

Figure 6. QLQ-C30 (+3) results (see Table 3 for definitions) of bone marrow transplant patients. Significance level of Mann–Whitney test: *$p<0.05$; **$p<0.01$; ***$p<0.005$; ****$p<0.001$. $n=48$ *vs* 1019. Solid bars, bone marrow transplant patients; open bars, other patients.

Discussion

Experiences made in the course of the present study

Advantages and disadvantages of the system Even during the first observations of patients' interaction with the QL recorder, it appeared that virtually no patients have any problems in its handling. While a patient is filling in a questionnaire, the PC remains completely in the background. The patient focuses on the questionnaire which he is familiar with. Only after the patient has finished the questionnaire, does the computer turn up again; either indicating missing answers or confirming the completion of the test.

The QL recorder has a number of disadvantages as well as advantages. Disadvantages are, for example, the size of the tablet which makes transportation difficult. It could be substituted by a laptop computer with a digitizer integrated in the screen. While multiple page questionnaires require a turning of pages on the tablet, a screen's contents could change considerably faster. On the other hand, the large size of the tablet guarantees that the questionnaire is clearly readable even by patients with visual problems. The QL recorder saves the patient from having to learn how to handle a new data entry tool (like a mouse) and from problems in adapting to continuously changing information on a screen. The most obvious advantage is that patients do not feel they are using a computer but that they are completing a 'paper-like' questionnaire.

Compared to more expensive systems, for example pen-based laptop computers or personal digital assistants, with their requested hardware and software-infrastructure, the low costs, the use of fully-developed techniques and standard hardware, and the ability to integrate in current electronic data processing environments are other advantages of the QL recorder.

Experiences gained from the project The goal of the project was to integrate QL measurement into the daily work of the whole outpatient clinic with its

different subdivisions. The authors wanted to assess all patients and were using a new method. Consequently, cooperation with nursing personnel and physicians was absolutely indispensable. The widely varying patient flow could be handled very well, and the time required for the completion of a questionnaire was about half as long as that reported in the literature (2) for the paper version. Over 4 weeks, a data set was generated which is representative for the outpatients. Patient compliance was excellent. The age distribution of the registered patients shows that even elderly patients could use the QL recorder without difficulties.

Results of the staff poll suggest that interest and motivation of the involved personnel influence their perception of additional work load, their estimation of stress induced to patients and their estimation of the importance of QL measurement. Answers supporting QL measurement were predominant, indicating its acceptance by the involved personnel. During the study, no severe problems occurred with the electronic questionnaire itself.

These findings demonstrate that the QL recorder can be integrated into clinical practice.

Results from further data analysis The results of descriptive statistics, of reliability analysis and of inter-scale correlation resemble the data from the validation process of the QLQ-C30 as reported in the literature, without being identical (2). As the observed differences were non-systematic, there was no evidence for a change in the questionnaires' characteristics induced by the electronic instrument.

The low Cronbach's α for the role function scale reproduces the findings reported in the literature. The low value of nausea and vomiting might be influenced by the fact that this symptom was quite rare in the study population (average = 9.8; 75% reported values below 17). Physical function and KF approach 0.7 while all other scales exceed it. The very high values of 0.89 and 0.91 for the general health and new general health scales might indicate that the people in this study population identify quality of life with their physical condition or general health constitution to a high degree.

The experiences with new role function and new general health reported at the November 1994 meeting of the EORTC Study Group on Quality of Life in Trondheim could be reproduced; both scales show better reliability than their predecessors.

In the comparison between characteristic results of sub-populations and their clinical pictures, every characteristic result corresponded with the clinical picture of its sub-population. This finding supports the reliability of the questionnaire results. In cluster analysis, the results of pair analysis of distance vectors support the hypothesis that patients from the same specialty have similar results with a higher probability than patients from different specialties. Nevertheless, the questionnaire cannot separate the specialties distinctly throughout all dimensions, as the k-means cluster analysis showed. These observations suggest that similar clinical pictures may cause similar symptoms or functional impairment, which results in better correlation among results from patients of the same specialty *vs* results from patients of different specialties.

Conclusion

The QL recorder can be used to collect representative QL data from patients in an outpatient clinic. No evidence was found for a significant influence of the QL recorder on the psychometric parameters of the EORTC QLQ-C30 (+3). Quality-of-life measurement was well accepted by staff and patients; patient compliance was high (98.9%). The QL recorder can help to save time and money, and improves data quality. It proved to be a useful tool for scientific use as well as for daily clinical practice.

Some distinctive differences between QL results of several sub-populations of outpatients were found. The characteristics of each sub-population corresponded with its predominant clinical picture. This finding and results of cluster analysis supported the hypothesis that although it may be impossible to identify reliably the underlying disease from a single QL measurement, its results have clinical validity and value.

Acknowledgements

The authors would like to thank: M. Fink, B. Mihanovic, G. Lutz and B. Schäffler for taking good care of the patients. Mr. Rettenberger and the staff of the outpatient clinic of the Medizinische Klinik der Universität Ulm for their extremely fine cooperation, W. Voigt and Mr. Ziesel for additional technical equipment, Dr. Kuhn who provided data from the clinical data center, I. Brand and S. Goldmann for their work as secretaries of the Cancer Centre of the University of Ulm, Fa. SPSS GmbH Munich for providing the SPSS-Software in time, A. Coates, W. Gaus, H. Heimpel and D. Machin for their advice in statistical questions. W. Streibl for early support of computing ambitions. A. Piffel for criticism and suggestions and H. Sigle for improving proofreading.

The reported work is part of a doctoral thesis prepared by cand. med. Jörg M. Sigle at the University of Ulm, Germany under the supervision of F. Porzsolt.

References

1. Sigle, J. M. (1995) Praktische Aspekte der Lebensqualitäts-Messung: Die standardisierte Messung der Lebensqualität bei Ambulanzpatienten mit einem elektronischen Lebensqualitäts-Recorder, Dissertation zur Erlangung des Doktorgrades der Medizin der Medizinischen Fakultät der Universität Ulm.
2. Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., Filiberti, A., Flechtner, H., Fleishman, S. B., de Haes, J. C. J. M., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P. B., Schraub, S., Sneeuw, K., Sullivan, M. & Takeda, F. (1993) EORTC Study Group on Quality of Life: The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J. Nat. Canc. Inst.* **85**: 365–376.
3. Dawes, R. M. (1979) The robust beauty of improper linear models. *Am. Psychol.* **34**: 571–582.
4. Osoba, D., Aaronson, N. K. & Till, J. E. (1991) A practical guide for selecting quality-of-life measures in clinical trials and practice. In: Osoba, D., ed. *Effect of Cancer on Quality of Life.* Boca Raton: CRC Press, pp. 89–104.

5. Till, J. E. (1991) Uses (and some possible abuses) of quality-of-life measures. In: Osoba, D., ed. *Effect of Cancer on Quality of Life*. Boca Raton: CRC Press, pp. 137–154.
6. Ware, J. E. & Sherbourne, C. D. (1992) The MOS 36-item Short-Form Healthy Survey (SF-36). I. Conceptual framework and item selection. *Med. Care* **30**: 473–483.
7. Sigle, J. M. SC-JMS GraTaGen and SC-JMS GraTaSim software manuals, Jörg M. Sigle, Kunstvolle EDV & Elektronik 1990–1995.
8. Sigle, J. M. SC-JMS GraTaSim/Lebensqualität—Ein angenehmes System zur computergestützten Erfassung der Lebensqualität, QL-Recorder software manual, Jörg M. Sigle, Kunstvolle EDV & Elektronik 1993–1995.