Contents

Guidelines for assessing Quality of Life in EORTC clinical trials

1. AIMS OF THE QUALITY OF LIFE GROUP

1. To develop reliable and valid instruments for measuring the quality of life of cancer patients participating in international clinical trials.

2. To advise the EORTC about the assessment of the multidimensional aspects of patients' quality of life as a measurable outcome of cancer treatment. Where appropriate the Quality of Life Group works in collaboration with the Pain and Symptom Control Task Force specifically to advise on evaluation of patients' subjective experience of symptoms.

3. To advise on the design, implementation and analysis of quality of life studies within EORTC trials, in cooperationwith the Quality of Life Unit at the EORTC Data Center.

4. To conduct basic research in quality of life assessment.

5. To contribute to teaching/training initiatives to promote the EORTC approach to quality of life assessment, e.g. through preparation of teaching material, oral presentations, etc.

6. To develop and maintain liaison with other non-EORTC groups conducting quality of life studies in oncology e.g. NCI-Canada Clinical Trials Group.

2. INTRODUCTION

When evaluating the efficacy of medical treatment on cancer, prolongation of life expectancy and tumor shrinkage have traditionally been taken as outcome measures. Despite the substantial side effects and functional impairment often associated with cancer treatment it is only recently that attention has been given to the assessment of quality of life (QoL). This increasing interest in QoL has important implications for clinical trials, as careful planning is required at all stages of a study from protocol design through to reporting of results. The EORTC Quality of Life Group (QLG) and the EORTC Quality of Life Unit (QLU) want to enhance the quality of data that is collected, and this is best achieved through sharing expertise and cooperating and collaborating with the Clinical Groups. It is important that the relative roles of each of the key players are clearly defined and understood. A common set of guidelines allowing a systematic approach across all EORTC clinical trials should further enhance the quality of data that is collected.

This manual aims to provide guidance for standardizing QoL assessment across EORTC randomized clinical trials. The roles of the EORTC QLG, QLU and Clinical Groups are described, along with the current procedures to be followed and the protocol requirements when preparing an EORTC study which includes the evaluation of QoL. Information is given to help decide when QoL assessment is likely to be a relevant and useful endpoint in a clinical trial, and to select appropriate instruments to measure QoL. There is a discussion on when and how often QoL should be assessed along with some practical methods for enhancing compliance and distributing and retrieving questionnaires. Finally there are chapters on data analysis and ethical considerations.

Whilst the guidelines primarily relate to EORTC phase III trials they may be of interest in phase II trials and to other trialists outside the EORTC.

3. THE PLAYERS

3.1. EORTC Quality Of Life Group

The EORTC Quality of Life Group (QLG) was created in 1980. Members of the group include social scientists, clinicians, statisticians, nurses and data managers from both Europe and Canada. Initially the group's activities centered on promoting the relevance of QoL in clinical trials and advocating its measurement, but progressed in 1993 to the development of a valid and reliable instrument for assessing QoL. A modular approach was adopted with a core questionnaire, the QLQ-C30, supplemented by disease and treatment specific questionnaires or "modules" (Aaronson et al., 1993; Bjordal et al., 2000) (Appendix 1). Considerable emphasis was placed on cross cultural applicability and the current version of the core questionnaire is now available in many languages (Appendix 2).

Permission to use the QLQ-C30 may be obtained from the QLU and there is no charge for academic users. A scoring manual is also available (Fayers et al., 2001). (See Appendix 3 for details on how to contact the QLU.) In collaboration with the QLU the QLG has produced a reference manual (Fayers et al., 1998b) in which datasets from various trials are pooled to provide tables and graphs of QoL scores for groups of patients, stratified by well defined variables such as type and stage of cancer, age, gender.
The manual is available in printed form or on a CD-ROM from the QLU.

All modules are developed according to strict guidelines (Blazeby et al., 2001) and are subject to peer review. If a module has been fully validated and published then permission to use it may be obtained from the QLU. Otherwise investigators should contact the QLU for the name and address of the principal coordinator. For a list of currently available modules see Appendix 4.
Translation guidelines have been produced (Cull et al., 1998) and may be obtained from the QLU. As part of the development process all modules are field tested in English and at least three other European languages.
Recently, a database containing all items from the core questionnaire and all existing modules in the available languages has been developed by the QLU (Vachalec et al., 2001). This Item Bank is accessible through the Internet (www.eortc.be/itembank). A user name and password may be requested from the QLU.

Research activities of the group now fall into four main categories:

1.   Further module development.

2.   Joint scientific projects with EORTC Clinical Groups where QoL is an endpoint in a new clinical trial protocol.

3.   Statistical/Methodological issues.

4.   Other areas.

A joint scientific sub-committee has been formed consisting of QLG members and QLU staff, all with expertise in different disease sites in addition to their knowledge of QoL assessment.
The aim of the sub-committee is to be able to offer advice to every EORTC Clinical Group on incorporating measurement of QoL into a clinical trial protocol, and on some of the practical issues associated with implementation. The aim of this manual is to ensure consistency of advice across EORTC Clinical Groups. For a list of joint scientific committee members and their contact address see Appendix 5. Where no suitable contact is listed the QLU will be responsible for all activities.

3.2. EORTC Quality Of Life Unit

The rapid growth in the number of studies assessing QoL emphasizes the need for a coherent policy and a standard approach to conducting this research. It is for this reason that the Quality of Life Unit (QLU) was created in the EORTC Data Center in November 1993. The Unit's main objective is to stimulate, enhance, and coordinate QoL as a treatment outcome in cancer clinical trials. In this context, the principal tasks of the Unit are to establish an adequate infrastructure for the data management of QoL studies; to facilitate the incorporation of QoL data collection into clinical trial protocols, and to support the analysis of QoL data in EORTC clinical trials. Both the QLQ-C30 and the modules are copyrighted instruments developed by the QLG with all rights reserved. Written prior consent of the QLG is therefore required for its use and the administration of the QLQ-C30 is an additional responsibility of the QLU.

Since its creation, the Unit's tasks have expanded, and staff numbers have increased. Staff in the Unit now includes a coordinator, a statistician, a data manager, an administrative assistant and research fellows. The translation coordinator, the module development manager, and the QoL specialist, also working at the QLU perform more specific support tasks.

The QLU is involved in a wide range of studies, across EORTC Clinical Groups, through all phases from the design to the analysis and publication of the results. The Unit has responsibility for supervising the data management for QoL evaluations in EORTC studies, and where possible encourages investigators to adopt a standard approach. This is achieved by involving the QLU in reviewing QoL issues in new protocols before they are submitted to the EORTC Protocol Review Committee. To ensure adequate rates of patient accrual, compliance, and data quality, there is a continuous need to maintain and improve standard data management strategies. Training courses are considered as an important preparation for those responsible for data collection in the clinical setting, and the QLU in collaboration with the QLG actively pursues such activities. A standard list is available for coding information on the reasons for missing data and incomplete forms, and rules have been drawn up for coding missing or ambiguous data (Appendix 6).

In addition the QLU provides Clinical Groups with regular feed-back on compliance figures, which should be prepared by the Clinical Group's data manager or statistician. For a list of all current EORTC studies that include QoL as an outcome measure see Appendix 7.

Statistical research activities at the QLU include collaboration with the QLG on production of the reference values data manual and investigating various methods of analyzing QoL data in cancer clinical trials. Analyzing QoL data may be complicated for several reasons e.g. repeated measures are obtained, data may be collected on ordered categorical response scales, the instrument may have multidimensional scales and complete data may not be available for all patients. In addition, it may be necessary to integrate QoL with length of life. The QLU has made some progress in all of these areas and has published articles in peer reviewed journals in association with members of the QLG and with members of other national cancer research organizations (Curran et al., 1998a; Curran et al., 1998b; Rosendahl et al., 1997; Troxel et al., 1998). The QLU has also developed statistical expertise from analyzing data from various EORTC clinical trials (Curran et al., 1997; Curran et al., 1998c).

3.3. EORTC Clinical Groups

The fundamental structure of the EORTC Treatment Division is based upon 28 Clinical research Groups and four task forces which develop their clinical research through the direct input of their participating scientists. Research is accomplished mainly through the execution of large, prospective, randomized, multinational cancer clinical trials. More than 2,500 clinicians located in 350 medical institutions in 35 countries participate in EORTC protocols. Each year approximately 6,500 new patients are entered into about 100 ongoing studies. Although there are 28 treatment Clinical Groups in the EORTC, not all of them have included QoL as an outcome measure in their trials. However, some groups have a long-standing experience in QoL assessments.

4. PROTOCOL DEVELOPMENT PROCEDURE

4.1. Background

The EORTC New Treatment Committee (NTC) and the EORTC Protocol Review Committee (PRC) are comprised of clinical trial experts including medical doctors, statisticians and QoL researchers. The PRC may also avail itself of external consultants who are specialists in a given field and to whom protocols may be submitted for external review.

The NTC reviews and approves the concept of EORTC trials with non-registered modalities on the basis of their scientific background, interest and feasibility. The PRC performs these same functions for EORTC trials with registered modalities, but additionally reviews the methodology for all studies regardless of modality. It also verifies that important scientific, methodological, collaborative, and administrative issues are in agreement with general EORTC operating procedures. Thus the roles of the NTC and PRC are two-fold:

1. To review and approve all new EORTC protocols with respect to their scientific value, feasibility and relevance within the framework of the EORTC.
2. To assist the Clinical Groups whenever necessary concerning any aspect of the design and implementation of their studies.

4.2. Submission Procedures

Within the EORTC, proposals for conducting a new clinical trial are developed by the Clinical Groups who generally appoint a writing committee to prepare the protocol. One investigator, appointed the Study Coordinator, will actually write the protocol and is responsible for the good conduct of the study within the Clinical Group. The protocol must be written in accordance with EORTC guidelines. Guidelines for submission of outlines and protocols to the EORTC NTC/PRC are provided on the EORTC website at www.eortc.be. For studies which include QoL as an endpoint the protocol writing committee consists of the study coordinator and additional members of the Clinical Group of the particular disease site, the statistician and data manager of the Clinical Group, a liaison person from the QLG and/or staff from the QLU. For all newly proposed phase III trials, a study outline describing the trial must be submitted to the NTC/PRC. The study outline template is available on the EORTC website (www.eortc.be). It is amended periodically. The current outline appears in the format of a standardized questionnaire. One section of the outline is specifically related to QoL as follows:

Do you intend to assess QoL in the study?

Yes/No

If yes,

- Have you contacted the Quality of Life Group (liaison person)?

  Yes / No

- What is the rationale for including quality of life in the study?

- Which QoL instrument(s) will be used in the study?

Issues relating to QoL must be discussed with the liaison person from the QLG or the coordinator of the QLU *before submission*. When the outline is complete it may be submitted electronically to the Data Center. Prior to submission to the NTC/PRC the outline is reviewed internally by the appropriate EORTC Data Center personnel, including the statistician, the medical supervisor and the QLU. After internal review the collective comments of the EORTC Data Center personnel are sent to the study coordinator. When the comments of the EORTC Data Center personnel are taken into account the revised outline is sent to the NTC/PRC.
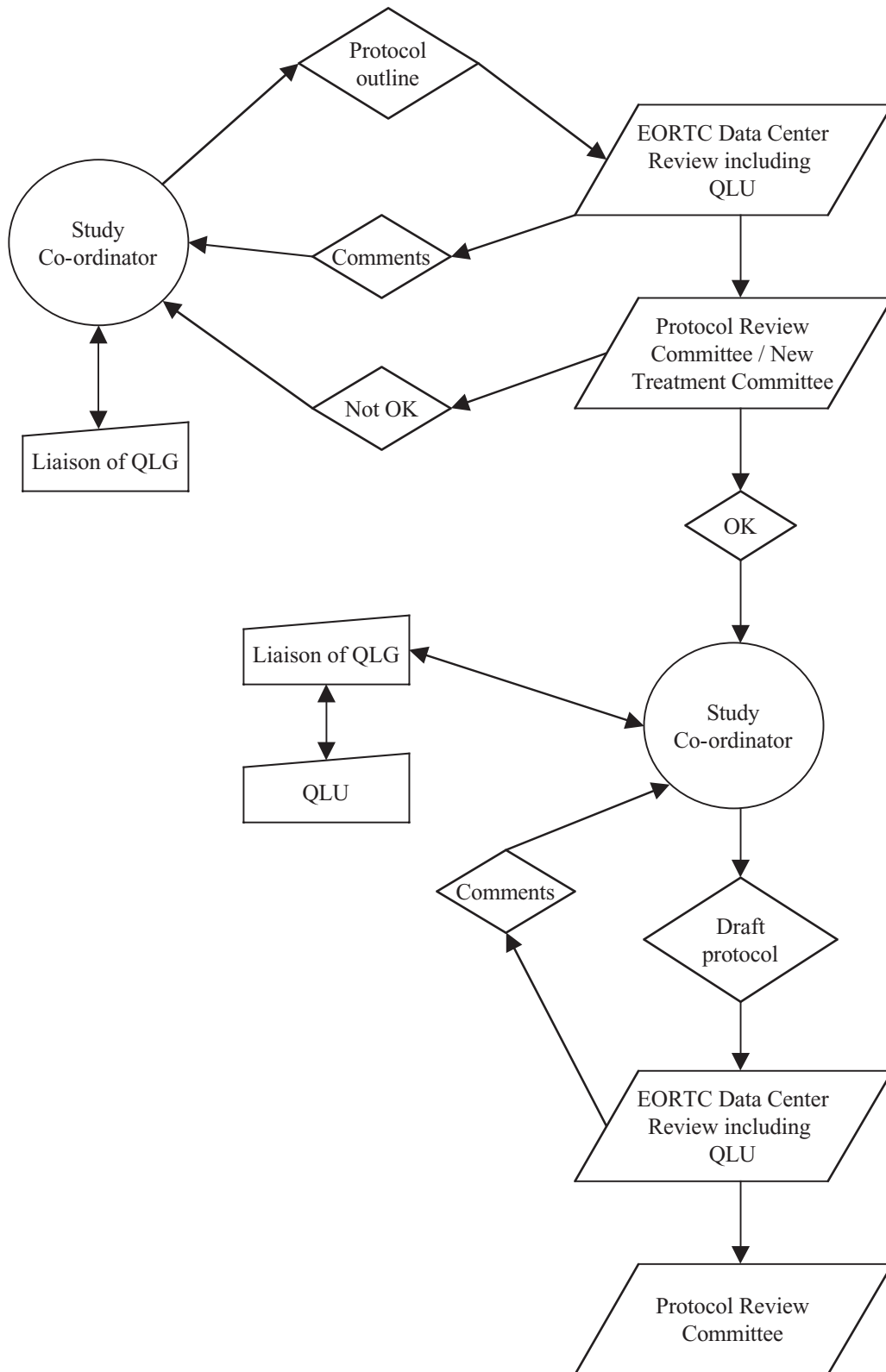
The outline is sent for external review. Afterwards, the study coordinator will receive a letter prepared by the Chairman of the PRC/NTC stating the results of the outline review: accepted, accepted pending modification, to be revised and resubmitted, or rejected. The PRC meets quarterly to discuss "problem projects". In order to have an outline discussed at the PRC meeting, it should be submitted at least six weeks prior to the meeting.

After approval of the basic concept by the PRC/NTC the study coordinator is encouraged to develop the full protocol in association with all members of the writing committee. The liaison representative of the QLG in conjunction with the QLU will draft the sections covering the topics presented in Table 1. After agreement from all parties involved (the liaison person of the QLG, the QLU and the study coordinator) on the content of the sections of the protocol related to QoL assessment, the text is incorporated into the protocol. Prior to submission of the final protocol to the PRC, the version of the protocol to be submitted to the PRC must be approved by the EORTC Data Center.

For phase II protocols written according to a PRC approved master protocol, a "quick procedure" may be employed. This implies that the first step involving developing the two-page outline is bypassed, and the full protocol may be developed in conjunction with the Data Center and the liaison person of the QLG.

Figure 1: Flow Chart Showing Stages in Protocol Development Prior to PRC Submission

Protocol outline

Study Co-ordinator

EORTC Data Center Review including QLU

Comments

Liaison of QLG

Protocol Review Committee / New Treatment Committee

Not OK

OK

Liaison of QLG

QLU

Study Co-ordinator

Comments

Draft protocol

EORTC Data Center Review including QLU

Protocol Review Committee

5. PROTOCOL REQUIREMENTS

The success or failure of a trial may depend on how well the protocol was designed and written. The protocol must be detailed and precisely worded with all the requirements clearly indicated so that the study may be uniformly carried out by all participants (Fayers et al., 1997). Table 1 lists the topics that should be covered in the full EORTC protocol. Discussion of these topics can be found in the relevant chapters.

Table 1: Protocol Contents Relevant to QoL Assessment

Topic	Chapter
1. Description of rationale for measuring quality of life	5
2. Statement of quality of life variables considered relevant (which side effects, late effects, psychological domains, primary or secondary endpoint)	6
3. Detailed description of design of the study	7
4. Patient eligibility	
5. Choice of instrument (which and why)	6
6. Timing of assessments	7
7. Mode of data collection (in person, by mail, etc.)	9
8. Statistical considerations (sample size, hypothesis to test)	10
9. Missing data (importance and methods for enhancing compliance)	8, 9, 10
10. Informed consent procedure	11
11. Appendices (instruments, patient information leaflets, consent form, diary record system)	

6. SELECTING TRIALS IN WHICH QUALITY OF LIFE IS RELEVANT

It would be unrealistic to recommend that QoL should be evaluated in all clinical trials. QoL has considerable resource implications and these should be balanced against the usefulness of the data and its likely impact on recommendations for treatment choice following completion of the trial.

6.1. Phase I & Phase II trials

Most investigators acknowledge that QoL measurement is unnecessary in Phase I and Phase II trials because the primary aim of such studies is to determine anti-cancer activity and toxicity, and patient numbers are usually small. However, it may still be useful in the following circumstances:

- If a QoL instrument needs piloting before being used in a Phase III trial. Investigators can ensure that the instrument covers all the relevant issues, assess reliability and validity and test the infrastructure for future data collection.

- As an exploratory study to investigate if there are unexpected QoL issues not covered by the questionnaire in use. Interventions may then be required in a phase III study to minimize symptoms and dysfunction.

- If a randomized study is likely to continue as a phase III study in which QoL is considered an important outcome measure.

6.2. Phase III trials

Some advocate that QoL should be measured in all phase III cancer trials (Osoba, 1992), and investigators are required to justify a decision not to assess QoL. Others adopt a more pragmatic approach and select trials where QoL is particularly relevant (Aaronson, 1995). Gotay and colleagues (Gotay et al., 1992) recommend that QoL evaluation is included in the following settings:

- A trial where QoL is considered to be the primary endpoint (e.g. the comparison of two palliative treatments).

- A trial where no significant differences between treatments are expected in terms of cure, disease free survival or overall survival but one arm is expected to be associated with significantly more morbidity. Following the trial the decision as to which treatment to recommend may have to be based on QoL outcomes.

- A trial where survival and disease free survival or cure are expected to differ between the two arms but the advantageous primary outcome is only achieved at the expense of major toxicity, e.g. high dose chemotherapy plus bone marrow transplant versus standard chemotherapy. Here data on QoL assessment can be used to support decision making when the benefits identified in the primary endpoint have to be weighed against a negative outcome in terms of QoL.

- It may also be necessary to assess QoL in studies evaluating cost-effectiveness. Specialist measures or instruments will usually be required and the advice of a health economist through the Health Economics Unit at the EORTC Data Center is recommended.

7. SELECTION OF INSTRUMENTS

Any questionnaire chosen to evaluate QoL should have proven, good psychometric properties with respect to validity, reliability, and responsiveness to change (Ware, 1987). Responsiveness refers to a combination of both reliability (identical scores in stable subjects over time) and sensitivity (the ability to demonstrate changes when the subject's state of health improves or deteriorates, or to detect treatment effects). This latter characteristic is particularly important in a clinical trial setting. The questionnaire should also be simple, brief, and easy to administer. These properties enhance participation and compliance,

and they reduce the burden for both patient and staff. Gelber & Gelber (Gelber & Gelber, 1995) recommend that the selection of instruments should be based on an assessment of the following four issues:

1. The purpose of the clinical trial.
2. The patient population.
3. The treatments and their potential toxicities.
4. The resources of the investigators and the participating clinicians.

In addition the QoL questionnaire should be available in the appropriate languages in relation to potential participants in the clinical trial.

There are two basic types of instruments: generic and disease specific. Generic instruments focus on the main components that constitute QoL, and they are intended to be applied in a wide range of populations and health states across all diseases. Disease specific instruments have been developed especially to detect subtle, disease and/or treatment related effects. There are many excellent validated self-completion questionnaires for cancer patients available e.g. EORTC QoL Questionnaire (EORTC QLQ-C30), Functional Assessment of Cancer Therapy (FACT), Rotterdam Symptom Checklist (RSCL), Functional Living Index-Cancer (FLIC) (Aaronson, 1995; Cella et al., 1993; de Haes et al., 1990; Schipper et al., 1984). All these questionnaires are multidimensional, minimally covering physical, psychological and social domains as well as some overall judgement of the valuation of life or the health condition. It is rarely necessary (or advisable) to develop a new instrument.

7.1. EORTC QLQ-C30

For trials coordinated by the EORTC both the QLG and the QLU recommend that, whenever possible, QLQ-C30 (version 3) should be used in its entirety for a number of reasons:

- The instrument has been carefully developed in a multi-cultural setting.
- Translations are available in 43 languages. If additional translations are required they can be developed using rigorous and standardized translation procedures.
- The instrument has been shown to be valid, reliable and responsive to change.
- Disease specific modules are available to supplement the core questionnaire.
- Study results can be compared across trials.
- Reference data is available for calculating sample sizes.
- The questionnaire is easily understood by most patients and is quick to complete (mean time 11 minutes).

Table 2: Structure of the EORTC QLQ-C30

EORTC QLQ-C30 (30 Questions in total)			
FUNCTIONAL SCALES (16 questions)	SYMPTOM SCALES (6 questions)	SINGLE ITEMS (6 questions)	GLOBAL QUALITY OF LIFE (2 questions)
Physical	Fatigue	Constipation	Global QoL
Role	Pain	Diarrhoea	
Cognitive	Nausea/vomiting	Sleep	
Emotional		Dyspnoea	
Social	Appetite		
		Financial	

Modification of the questionnaire is not permitted without prior consent of the QLG as it is a breach of copyright. Scales that appear irrelevant should only be omitted in exceptional circumstances and individual questions should never be used alone when they form part of a scale. Amongst the reasons for this are:

•   The psychometric performances of individual scales and items when used alone are not known.

•   Less than 5% of patients find any one item upsetting.

•   Having been used in over 600 studies there is no evidence to suggest that patients are bothered by questions relating to symptoms or problems they do not experience.

•   Unexpected and important results may be observed and used to generate hypotheses for future studies.

•   Data can be used for updating the reference manual of normative data.

•   Data can be used for meta analysis.

In collaborative studies with other groups an alternative measure may be proposed. This may occur when industry has sponsored a particular instrument or the other party has had experience with one. In such cases it is recommended that, finance permitting, a head to head comparison with the other instrument is carried out (e.g. FACT, RSCL, SF36).

7.2. Modules

Cancer site or treatment specific modules have been developed by, or in collaboration with, members of the QLG according to strict guidelines drawn up by group members. These modules are in various stages of development and a list is included in Appendix 4 along with contact details for the principal developer. For modules where validation data has been published, permission to use the module may be obtained from the QLU, otherwise permission may be obtained from the principal developer. The modules are intended to supplement the QLQ-C30 and should not normally be used without concurrently administering the QLQ-C30. Where an EORTC module is not available, use of an existing instrument in the area of interest, with known psychometric properties, is preferred to developing an ad hoc questionnaire.

Where a specific research question is posed and use of the QLQ-C30 and a module is insufficient, it is possible to add extra questions as a checklist. The QLU has recently set up the Item Bank in close collaboration with the QLG Item Bank Committee. Validated items from the Item Bank can be used

to supplement the core questionnaire and other modules, if necessary. The use of additional items has to be carefully discussed with and approved by the QLU. Other main aims of the Item Bank are to improve the quality of modules and to standardize the wording of items in existing and future modules, to improve the speed and quality of module development and to improve the speed of translations.

7.3. Other Situations

Whilst the QLQ-C30 is usually the instrument of choice in phase III clinical trials, there are situations where its sole use would be inadequate or inappropriate as it does not address all the research questions of interest. e.g.

- Survivorship studies where the QLQ-C30 does not cover the relevant aspects such as work rehabilitation, relationships, infertility.

- Cost-effectiveness analyses where a single measurement is often required.

- Paediatric studies.

8. WHEN & HOW OFTEN SHOULD QUALITY OF LIFE BE ASSESSED?

Within the context of a clinical trial where it has been decided that it is appropriate to assess QoL there are countless opportunities when patients could be asked to complete a QoL measure. However the burden on the patient and the associated data management and data analysis problems necessitate some limitation. The ideal number and timing will vary from one clinical trial to the next, dependent upon the research hypothesis, but will usually involve assessments before, during and after treatment. To facilitate a sensible and practical interpretation of the results the minimum number of measurements should be used. They should be timed to yield maximum information about changes in QoL due to both treatment and changing disease status. Timing schedules should be similar across all treatment arms.

8.1. Before Treatment

Information on the patient's QoL prior to their diagnosis of cancer is rarely available so a pre-treatment score is usually considered to be their starting point or baseline assessment.

- It allows for comparison between study groups before treatment is initiated. If differences are found they can sometimes be controlled for during subsequent analysis.

- A pre-randomization assessment provides a starting point for assessing changes caused by both treatment and disease status.

- Where follow up data is missing and the patient was known to still be alive it may allow for detection of a systematic bias. For example, patients with a poor QoL at baseline may not always be asked to complete follow up assessments.

- In addition it has also been shown that QoL at baseline may be of use as a prognostic factor for clinical outcomes, including survival, response to treatment and nausea and vomiting (Coates et al., 1997; Osoba et al., 1994; Tannock et al., 1996).

8.2. During Treatment

Choosing a schedule for collecting QoL data during treatment will often involve a compromise. To reduce the administrative burden and thus improve compliance, assessment times should coincide with the clinical care schedule dictated by the trial regimens. However assessments should be timed to reflect the expected profile of treatment burden and toxicity. These requirements often conflict and it is often not possible to recommend a fixed schedule to be used in all trials. Rather it is essential to appreciate the clinical course of the disease and to discuss within the protocol writing committee expectations regarding serious or acute

effects, stable periods and chronic problems. An appropriate research hypothesis can then be formulated and used as a guide to determine appropriate measurement times.

When writing a protocol one should aim for similar (if not identical) schedules in terms of frequency and number of assessments between the two arms.

8.2.1 Anchoring Events

Assessments can be a) time-based - given a set number of days or weeks after randomization, independently of the specific treatment schedules, or b) event based - given to coincide with specific treatment cycles (See figures 2 and 3), or even daily. A decision as to which is the most appropriate approach will depend upon the research question, but careful thought should be given to the consequences of delayed treatments.

- An event-based approach is often logistically easier to manage but it does not provide QoL data between treatments. Assessments are usually scheduled to take place immediately before the next course of treatment. In the case of treatment delays assessments would still fall immediately before a specified course but if the reason for delay was toxicity, information will not be collected on the patient's QoL at the time of the delay. If the "events" happen at different times in the trial arms but the treatments are equivalent, patients in one arm may have more advanced disease because a longer interval has elapsed since randomization (e.g. figure 2, 3rd QoL assessment). The number of assessments is independent of the duration of treatment, which may simplify analysis.

- Whilst it may be more difficult to facilitate reliable data collection in a time-based approach, if the times are carefully chosen data on QoL between treatments can be made available. Assessments will always be at the same time interval in all arms relative to randomization. The number of assessments is only known if treatment is not delayed and analysis of data may become more complex.

- Sometimes a combination of the two is appropriate. For instance recommending that QoL should be assessed at day 21 of each cycle before the next treatment is administered ensures only one assessment per course, but should be sensitive to changes in QoL at a time of treatment delay if this is necessary.

- Daily assessments may uncover details that would be missed by less regular assessments, but there may be difficulties with compliance. A large volume of data is generated and a corresponding increase in time must be allowed for input and analysis. The use of daily or weekly measurements should be very limited.
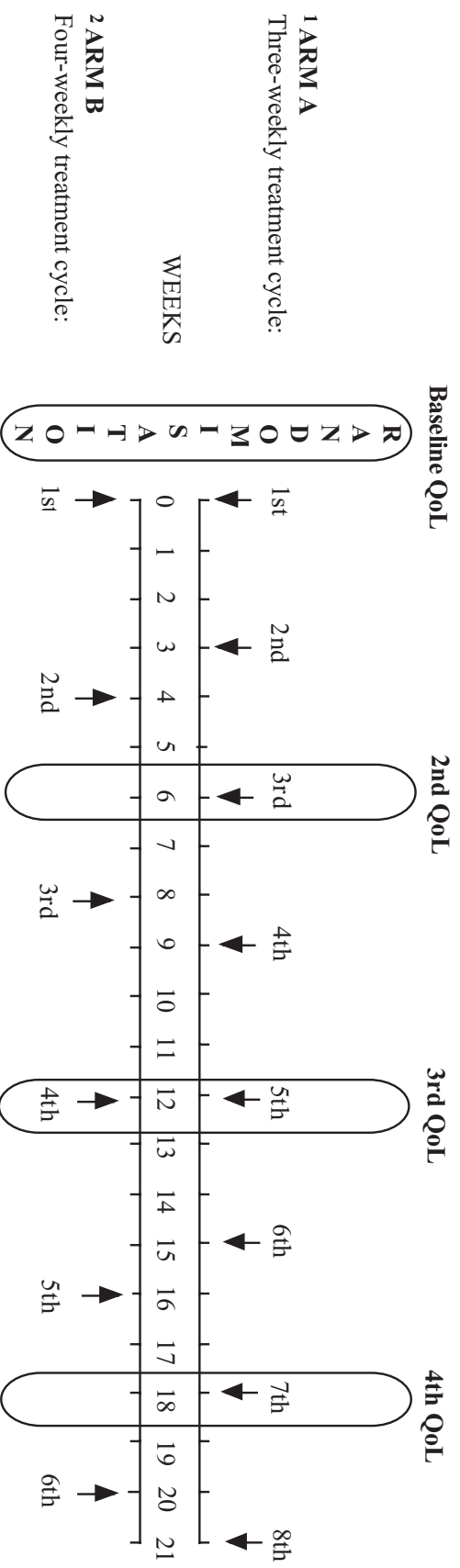
8.2.2 Time Scale

The time scale of the chosen questionnaire needs to be considered. Some questionnaires refer to symptoms and QoL during the previous week (e.g. EORTC QLQ-C30) whilst others ask about current status. Where acute side effects are expected within a few days of treatment it may be inappropriate to collect QoL data three weeks later on the patient's next visit.

8.3. After Treatment

Once treatment is completed the number of QoL assessments required and their frequency depends upon the research hypothesis and whether QoL was specified as the primary endpoint. As before, a compromise will be needed to balance "exploratory" requirements with more pragmatic considerations. To eliminate bias, assessments should occur at equal times in each arm relative to randomization and not to end of treatment.

- In clinical trials where the patients have a poor prognosis data may be lost if the interval between assessments is too long. Care should be taken not to overburden patients in the last few months of their lives.

Figure 2: *"Time" Based quality of life evaluation e.g. 6 weekly assessment*

1 ARM A
Three-weekly treatment cycle:

2 ARM B
Four-weekly treatment cycle:

WEEKS

RANDOMISATION

Baseline QoL — 2nd QoL — 3rd QoL — 4th QoL

Weeks: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

Arm A assessments: 1st, 2nd, 3rd, 4th, 5th, 6th, 7th, 8th

Arm B assessments: 1st, 2nd, 3rd, 4th, 5th, 6th

In both arms the baseline assessment is carried out before randomisation. Treatment should start as soon as possible after randomisation. Thereafter assessments are at 6 weekly intervals relative to the randomisation date.

1 In Arm A treatment is three-weekly so QoL assessment always coincides with the start of a new cycle.
2 In Arm B treatment is four-weekly so QoL assessment sometimes falls between cycles and sometimes it coincides with the start of a new cycle.

Figure 3: "Event" Based quality of life evaluation e.g. Assessment every 2nd cycle
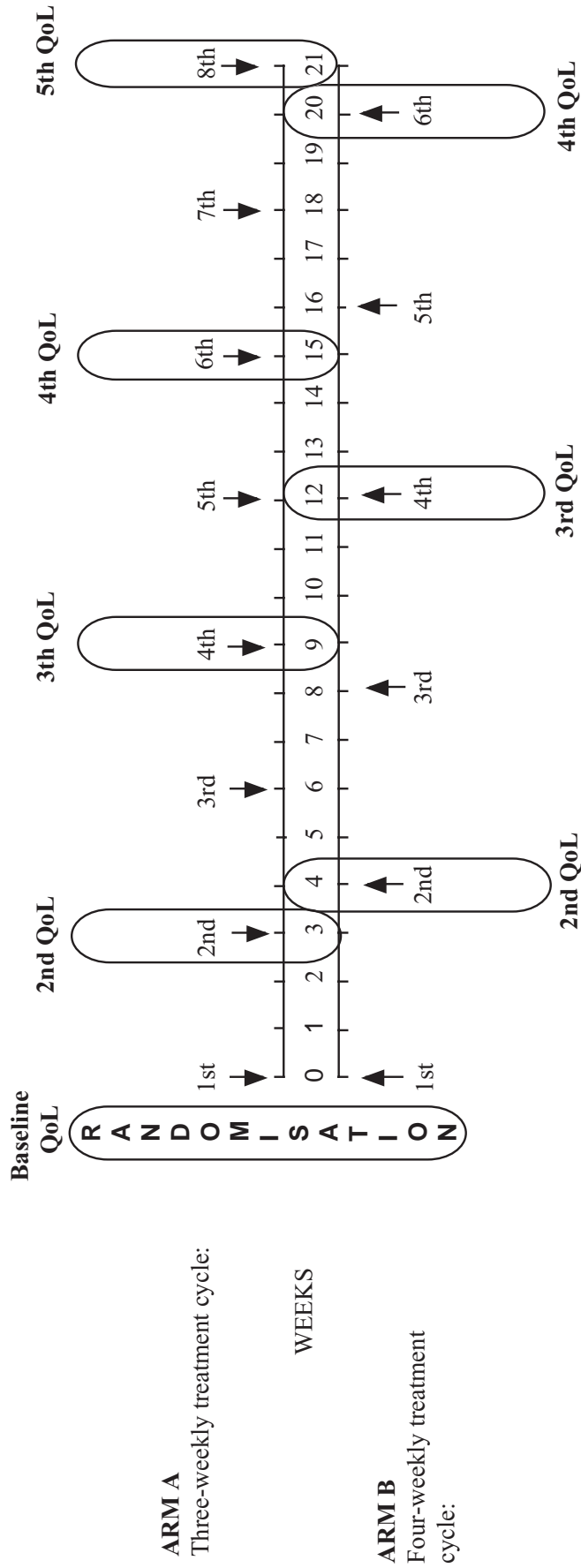
QoL is assessed at baseline (before randomisation) in both arms. Treatment should start as soon as possible after randomisation and QoL is then assessed after the 2nd, 4th, 6th (and 8th) cycles.

In both arms QoL assessment always coincides with the start of a new cycle but in arm A there are more assessments and assessment is also more frequent.

- For patients receiving radical treatment where long-term survival benefits are expected intervals between assessments can be extended.

- If QoL data is required at the time of relapse local institutions should consider how they will achieve this. The patient will be withdrawn from the study at an unknown point in time and the appropriate questionnaire may not be to hand, but compliant patients may be upset if, as they relapse, an assessment of their QoL is seen as unimportant. Once the patient has relapsed subsequent QoL data collection may be hampered by the infrequent, if any, visits, of patients to the study hospital.

- In some studies it may be necessary to collect QoL data until death. For example studies comparing immediate versus delayed treatment, or studies when a treatment is expected to prolong time-to-progression or disease-free interval but no difference in duration of survival is expected. In these situations it is important to illustrate the benefits of extending the time to an event versus the side-effects of the treatment.

- If questionnaires are to be handed out in person then the assessment schedule needs to coincide with the follow-up schedule. However if questionnaires are to be mailed any appropriate time schedule can be used.

9. ENHANCING COMPLIANCE

Unless or until collecting QoL data is seen as part of routine clinical practice it will be necessary to implement specific measures for each clinical trial because QoL data cannot be collected retrospectively. Missing data makes analysis very complicated and results difficult to interpret. If QoL is specified as an important primary or secondary endpoint in a multicenter clinical trial protocol then it must be mandatory in all participating centers. If QoL assessment is left as optional and restricted to centers that have an infrastructure to facilitate it, the patients may not be representative of the wider sample drawn from all participating centers. QoL assessment in a subgroup of patients would still require a large number of patients in order to satisfy the statistical power of the study. There would be a crucial need for maximal compliance in the subgroup and an unforeseen number of drop-outs might render the study unevaluable.

Both staff and patients should be provided with the necessary resources for optimal data collection at the appropriate times. Collaboration between the study coordinator, the data center responsible for the day to day administration of the study and the individual investigators can lead to organizational improvements. More specific measures can be targeted at the patient, the physician and the data manager or the research nurse, though there is considerable overlap.

9.1. Organizational Issues

Good organization and forward planning ensures that all those involved know their respective roles.

- A small local pilot study may be organized prior to commencement of the main study, followed by a debriefing meeting. This will enable each center to make a realistic assessment of the number of patients they are likely to recruit and the time, space, personnel and financial resources that will be required.

- Where the task of collecting QoL data is shared amongst a number of people, one should be appointed as the local coordinator. The coordinating data center could request details of this individual at the same time as they verify ethical committee approval and collect data on laboratory normal values. The individual is then responsible to the coordinating data center and any queries can be directed through them.

- Baseline QoL assessment should be one of the eligibility criteria. Completion of a QoL questionnaire would then be included on the checklist which has to be completed before randomization can take place.

- The procedures for collecting QoL data at each center should be documented and include names and contacts for all those involved. Then in the event of staff absences everyone is aware of where to find the relevant paperwork and who has what responsibility.

- Recruitment and compliance figures are prepared by the data manager and statistician of the Clinical Group and can be presented by the QLG liaison person or a representative of the QLU every six months at the Clinical Group meetings. This allows feedback and discussion if there are any problems. Procedures in place within the EORTC to monitor data timeliness apply to QoL data as well as clinical data. Investigators with patients who are not evaluable due to missing QoL data will be notified so that policies can be set in place within their institution to rectify the problems and not jeopardize the study.

- For some trials the QLU has developed a schedule for QoL assessment for each patient starting from the date of randomization. This is sent to the attending physician and kept in the patient files. This is useful for the medical staff to check when the patient should complete the QoL assessment. It also allows the study monitor to check if QoL assessments have been done according to the schedule provided in the protocol.

- A spare copy of the QoL questionnaire should be kept in the patient's clinical file; if the questionnaire is lost a backup copy will then be available.

- The following set of questions are included on the clinical Case Report Forms as a reminder that QoL questionnaires should be completed and to aid in determining the reasons for missing questionnaires:

  Has the patient filled in the current QoL questionnaires, 0 = no, 1 = yes
  If no, please state the main reason

  1 = patient felt too ill
  2 = clinician or nurse felt the patient was too ill
  3 = patient felt it was inconvenient, takes too much time
  4 = patient felt it was a violation of privacy
  5 = patient didn't understand the actual language / illiterate
  6 = administrative failure to distribute the questionnaire to the patient
  7 = not required at this time point
  8 = other, please specify

9.2. The Patient

Most patients are willing to complete QoL questionnaires. Specific measures to improve compliance include:

- Providing a clear explanation of the reason for collecting QoL data in the context of the rest of the study. This information should be given verbally and supported by a written information sheet.

- Providing information on when questionnaires will be due e.g. a copy of the reporting schedule mentioned above.

- Informing patients what will happen to their completed questionnaires, e.g. they will not be stored in the patient's clinical notes and will remain confidential. (Within the field of clinical trials there are major discussions about individual and collective ethics. In many studies completed QoL questionnaires from trial patients are not made available to the treating clinician during their consultations and patients should therefore be made aware of this at the time they consent.)

- Ensuring the questionnaire itself is not too long and contains questions that appear relevant to the patient and are easily understandable. If more than one questionnaire is used care should be taken to avoid duplication of issues. The format and layout should be clear and include written instructions.

- Providing a private and comfortable environment for completing the questionnaire.

- Providing help if necessary for patients who are unable to complete the questionnaire unaided for whatever reason (e.g. poor comprehension, no glasses, etc.).

- Showing appreciation once the questionnaire is completed and expressing an interest in any concerns the patient may raise.

9.3. The Physician

Some clinicians are unconvinced of the scientific validity of QoL assessment. They may be sceptical about the value of measuring QoL within a clinical trial and therefore have difficulty in seeking the cooperation of all patients. The following measures may have a positive influence on their opinion of QoL assessment:

- QoL data collection should not be presented as an "optional extra" in a trial but rather seen as a mandatory and integral part of the study.

- Published work where QoL data has made a significant contribution to the scientific validity of a study should be presented and promoted.

- The study coordinators should be seen to be convinced of the value of QoL measurement. They should also be clear as to the rationale for collecting QoL data in their particular trial and should use this information to motivate the study investigators.

- Clinical considerations suggested during discussions between the coordinator and the investigators, should, where possible, be taken into account when designing the QoL component of the trial.

- Investigators should receive feedback regarding the QoL data collection in much the same way as they receive recruitment updates and preliminary results from the clinical component of the study.

9.4. The Data Manager/Nurse

Responsibility for distributing QoL questionnaires is often allocated to a data manager or research nurse. A distinction should be made between someone who is available in the clinic to attend to the patient personally, and someone who visits the treatment center at regular intervals, but can only leave the questionnaire in a prominent place with a reminder that it be given to the patient on their next appropriate visit and collected later. The latter have a limited role to play.

- If data managers or nurses are expected to distribute questionnaires personally they should be well informed so that they can answer questions. Trial specific workshops held prior to the commencement of a study have been advocated. The rationale for collecting QoL data can be explained and the practical procedures to be followed discussed in detail. In a multi-national setting this may not be practical or cost-effective. As an alternative, national training courses in data management could be encouraged to broaden their coverage of general issues surrounding the collection of QoL data, which would then be applicable to a wide variety of trials.

- Regular contact between the data manager/nurse and the study investigator should increase motivation and enhance compliance.

- When the investigator receives feedback on the center's compliance they can convey this information to the data manager/nurse.

For some patients, completing a QoL questionnaire may prompt them to seek more information or support. It is then important that the data manager/nurse is competent to deal with any issues that may arise, or knows where to refer the patient for appropriate help. Relevant training should be arranged along with an awareness of the resources available.

10. PRACTICAL PROCEDURES FOR DATA COLLECTION

The EORTC QLU have produced a two page information sheet "EORTC Guidelines for administration of QoL questionnaires". (Appendix 8)

10.1. Mode Of Delivery

The two most common modes of administration are handing questionnaires to the patient in person whilst they attend a clinic or mailing them to their home address. For baseline and on-treatment assessments the QLG and QLU recommend handing out questionnaires in person because:

- The first time patients are asked to complete a questionnaire they may not understand the instructions or may find some questions confusing. Available staff can give a verbal explanation.

- Where patients are unable to fill in the questionnaire themselves for practical reasons (e.g. forgot glasses, too frail) the member of staff may choose to read out the questions and fill in the questionnaire on the patient's behalf. This should then be recorded on the form.

- There is an opportunity to check questionnaires for missing data and ascertain whether this is accidental or deliberate. In the former case patients can be asked to complete the missing questions whilst in the latter case the questionnaire should be marked that the patient did not wish to answer particular questions.

- Some patients are unable or refuse to complete a whole questionnaire; the reason for this can then be ascertained and recorded.

- When necessary one can tactfully try to discourage relatives from answering the questions on behalf of the patient.

Patients should be discouraged from taking questionnaires home and returning them either by post or in person on their next visit. (In such cases the investigator has no control over the exact completion date or whether the patient's answers were influenced by family members or friends.)

If questionnaires are mailed to patients it will be necessary to check on their survival status beforehand to avoid distressing relatives of patients who have died. Reply paid envelopes should be provided.

10.2. Time Of Delivery

- It is normally recommended that baseline data is collected before randomization, so that completion can then be made an eligibility criterion and the outcome of randomization cannot influence any of the domains in the QoL score.

- As it is preferable to reduce all sources of potential bias it is recommended that questionnaires are completed prior to seeing the physician. This has the advantage that it may prompt the patient to discuss any worrying symptoms.

10.3. Missing Data

Protocols should contain explicit instructions for normal practice and what to do in the event of a protocol deviation. If a questionnaire is missed the protocol should be clear on the practice to follow. Should the data be accepted as missing and only a reason recorded (e.g refusal, nurse forgot, etc.) or should the patient be contacted? Options include mailing the questionnaire and a reply paid envelope or telephoning. Whichever method is chosen it is important to establish beforehand an acceptable time delay or "window" during which the questionnaire must be completed. During treatment, windows should be narrow to evaluate short term toxicity timing (e.g. +/- 1 week) but during follow up wider windows may be acceptable.

10.4. Proxy Ratings

It is generally agreed that patients are the best raters of their own QoL (Slevin et al., 1990). There are circumstances in which it is difficult or even impossible for patients to rate their own QoL (e.g. patients who are cognitively impaired due to their cancer, patients who are terminally ill and children). In these circumstances their QoL may be assessed by a third person. This can be a family member (e.g. partner or a parent) or the care taker (e.g. physician or nurse).

In two samples of cancer patients, Sneeuw (Sneeuw et al., 1997,1998) examined the level and pattern of agreement between ratings provided by patients and their significant others on the EORTC QLQ-C30. At the individual patient level, more than 90% of scores were within one response category of difference, and correlations for the several dimensions were moderate to good (between 0.40 and 0.80). At the group level, significant others tended to rate the patients as having a lower quality of life than the patients themselves, but this bias was of a limited magnitude.

Sneeuw (Sneeuw et al., 1997) reported very similar findings when comparing ratings provided by cancer patients, significant others and physicians on the COOP/WONCA charts, assessing several quality of life dimensions at a generic level by means of seven single questions. Lower levels of agreement were noted for more private domains, such as feelings, social function, and overall quality of life. The level of agreement between patients and their physicians was only slightly lower than that observed between patients and their significant others. Physicians tended to underrate patients' pain severity.

Stephens (Stephens et al., 1997) investigated the concordance between ratings provided by lung cancer patients and their physicians on eleven symptoms derived from the Rotterdam Symptom Checklist. Of all comparisons made, 78% showed exact agreement between doctor and patient, 18% disagreement by one, 4% by two, and 1% by three grades. However, there was increasing disagreement with increasing symptom severity, and a consistent bias towards doctors underestimating symptom severity. Importantly, physician compliance was higher than patient compliance, and the between-treatment comparisons reached the same conclusions regardless of whether the data was patient-based or physician-based.

If it is anticipated that an increasing percentage of the patient population under study will be unable to complete questionnaires during the course of the trial (e.g. due to neuro-psychological deficits or a seriously deteriorating physical condition) proxy respondents could be considered from the trial outset.

11. DATA ANALYSIS

Analysis of QoL data raises a number of contentious issues, and we outline some of the main ones. Three points are of particular note.

1. The analysis of completed trials will be simpler and more convincing if the principal hypotheses have been specified a priori. Both the hypotheses and the QoL outcomes to which they relate should be specified in detail in the protocol.

2. The definition of "clinically important differences" should be considered at the time of writing the protocol, and should be specified. (This will also be necessary when sample size estimation is based upon QoL endpoints.)

3. Since missing data (non-returned QoL forms) raises major questions about bias and poses severe analytical problems, every attempt should be made to ensure high compliance.

11.1. Simple Comparisons

Many of the complications in analysis arise because studies which assess QoL usually assess each patient at multiple time points. When cross-sectional analyses are carried out (for example, all patients at the pre-randomization time point), many of the problems disappear. Sometimes simple t-tests may be appropriate (for example, when comparing global health status across two treatment arms).

Often non-parametric tests, such as Wilcoxon or Mann-Whitney tests, may be more appropriate because many of the single items and some of the scales have asymmetric distributions. It should also be noted that the single items are mostly 4-point scales, and so ordered logistic regression may be appropriate if one wants to use regression techniques to examine the effect of prognostic variables upon QoL outcomes.

An alternative approach, which may be especially suitable for single items, is to consider percentages rather than means or averages. For example, instead of estimating the average vomiting score for each group of patients, one can calculate the percentage of patients in each group who report "quite a bit" or "very much" vomiting. Many readers may find percentages more intuitive and easier to understand than average levels. For example, the statement "24% of patients reported vomiting at least "quite a bit" has a more obvious interpretation than reports such as "the average level of vomiting was 58.2". When percentages are used, the analyses often reduce to comparisoıns of binomial proportions or possibly chi-squared tests.

11.2. Multiplicity Of Outcomes

The core QLQ-C30 contains 30 items and a number of scales (five functioning scales, one global health status, and three symptom scales), with the supplementary modules containing additional items and scales. Thus there are potentially many pairwise statistical comparisons that might be made. As is well known, for every 100 independent statistical tests that are carried out, even if we assume there is no treatment effect, we would expect approximately five comparisons to be statistically significant at $p<0.05$. Therefore, when making multiple significance tests, we are likely to obtain about 5% of results as false positives.

There are three main methods of making allowance for this. First and foremost, the study protocol should identify one or two QoL outcomes as being of principal interest. These few outcomes will be the main focus of the analysis, and therefore there will be no problems of multiple testing. It is important that these outcomes are listed in the protocol, to avoid it being suggested that the investigators "cheated" and inspected the data before determining which variables are of most interest. All other analyses may then be regarded as primarily hypothesis generating, and will be regarded more critically.

The second method, which is sometimes used in conjunction with the first, is to adopt "conservative p-values." If many statistical tests are being performed, it is possible to use $p<0.01$ as indicating statistical significance, thereby reducing the rate of false positives. In extreme cases, $p<0.001$ could be used. Rather related to this, "Bonferroni corrections" are often used. The principle underlying this is that in theory one should not use a fixed but arbitrary $p<0.01$ irrespective of the total number of statistical tests. Instead, if it is planned to make N statistical tests, one can estimate the equivalent p-value that will maintain overall significance at, say, 5%. For an overall p-value of a, the Bonferroni method indicates that one should use p-values of a/N for the individual tests (Bland & Altman, 1995). The third method is either to use some form of global multivariate test, or alternatively to reduce the items to a few summary scores. This method has not often been used in QoL studies. Tandon describes applications of global statistics in analyzing QoL data (Tandon, 1990).

11.3. Repeated Measurements

There are various methods available for data description and statistical significance testing when repeated measurements are available for the comparison of two or more treatments. One of the simplest approaches is to use graphical displays and accompany these by cross-sectional analyses at a few specific time points. Ideally, the study protocol will have pre-specified that the analysis will focus upon QoL at these particular time points, with the additional measurements being regarded as of secondary importance. For example, a chemotherapy protocol might specify that differences in QoL at the time of the third course, and also at one month after completion of chemotherapy will be tested for statistical significance. These tests could be accompanied by graphical displays showing the average levels of QoL for the treatment arms and possibly for various patient subgroups.

A second approach is to condense the repeated measurements for each individual into a few summary statistics. For example, one could estimate the average level of each QoL scale, taken over the on-treatment period. This would reduce the repeated on-treatment measurements for each patient to a single score. Other summary statistics that are frequently employed include (a) the overall average QoL for each patient, (b) average QoL after completion of therapy, (c) the worst QoL experienced during therapy (or highest levels of toxicity) and (d) the "area under the curve" (AUC), which is equivalent to the average if the time points are at equal intervals. The analyses can then compare and test the summary statistics. The application of these methods is described by Matthews et al (Matthews et al., 1990).

Finally, some sophisticated statistical methods are available for the analysis of repeated measurement data. Mostly, these methods involve fitting a mathematical model to the data. Since repeated measurements on any one individual are likely to be correlated, the model must allow for the auto-correlation between values at successive time points. The main methods are multivariate analysis of variance (MANOVA) for repeated measures, hierarchical models (multilevel models) and generalized estimating equations (GEE) (Diggle et al., 1994; Goldstein, 1995; Hand & Crowder, 1996; Lindsey, 1993).

11.4. Missing Data

Two types of missing data may be distinguished. First, patients may fail to compete all items on a form, possibly accidentally. The EORTC QLQ-C30 Scoring Manual describes an elementary method of calculating scale-scores when there are a few missing values for some items.

The second, and usually far more serious problem, arises when whole forms are missing. In particular, it is often difficult to know whether patients do not return forms because they feel too ill, or whether the reason is that they feel fine and see little point in replying. Thus one can never be confident that the observed QoL data is representative of all the patients in the study. Sometimes there may be serious bias. Missing data is often a particular problem when carrying out longitudinal (repeated measurements) analysis. However, it should be emphasized that whenever there are many patients with missing data the results of any analysis, cross-sectional or longitudinal, may be suspect. How can we be sure that those patients with data are truly representative of the total sample recruited to the study? Hence, are the results biased?

When data is missing, there is no easy solution for eliminating bias. Therefore, emphasis must always be placed upon avoiding the problems by ensuring optimal compliance with assessment. This cannot be stated too strongly. Any form of correction to the analysis will always be regarded with suspicion by readers, and the study results will only be convincing if compliance is high and missing data is kept to a minimum.

Analytical methods tend to be complex, and are controversial. A special issue of Statistics in Medicine (1998, Volume 17) is devoted to this topic, and contains contributions made on behalf of the EORTC QoL Study Group (Curran et al., 1998b; Fayers et al., 1998a). More recently, Fayers and Machin also published a book on assessment, analysis and interpretation of quality of life data (Fayers and Machin, 2000).

11.5. Interpretation & Clinical Significance

It is relatively easy to obtain a feeling for percentages (for example, "30% of patients reported quite a bit of problem with tiredness"), but many of the items on the QLQ-C30 contribute to multi-item scales which are scored from 0 to 100. Most users are unfamiliar with these particular scales, and do not know how to interpret the mean scores. Also, in a two-arm clinical trial, what interpretation should be given to, for example, a difference between emotional functioning of 58 in one treatment group and 66 in the other? Statistical significance tells us whether the observed data can be explained by chance fluctuations (such as selection of patients), but says nothing about clinical significance. Is a difference of 8 (i.e. 66 - 58) large enough to be important? If a patient's score changes by 8 points, would they even notice the change? Osoba et al. (Osoba et al., 1998) asked patients to complete the QLQ-C30 on repeated occasions, and the patients also rated their perception of change since the previous time they completed the QLQ-C30. Physical functioning, emotional functioning, social functioning and global QoL scales were evaluated. It was found that when these scale scores changed by 5 to 10 points (on the 0-100 scale), patients described

their condition as "a little" better (or worse). A change of 10 to 20 was described as a "moderate" change. A change greater than 20 was "very much" better (or worse).

King (King, 1996) used a very different approach, based upon "known groups" who were expected to differ in terms of QoL scores, such as limited disease patients and those with advanced disease. Data was collated from fourteen published studies. She concluded that for most scales a difference of 5 or less is a "small" difference, but the definition of a "large" difference varied for each scale: for example, it was 16 for global QoL, 27 for physical functioning, and 7 for emotional functioning.

Hjermstad et al. (Hjermstad et al., 1998) report normative data for the QLQ-C30 in a randomly selected sample of 3000 people from the Norwegian population, aged between 18 and 93 years. Data was available for 1965 individuals. Results are presented for the functioning scales, the global QoL scale and the single items. The results are tabulated by age and sex. These normative data may serve as a guideline when interpreting QoL in groups of cancer patients.

For individual patient sub-groups, the EORTC QLG has produced a manual of reference data (Fayers et al., 1998b). Members of the QLG contributed data from their studies, which was pooled for the tables. The manual tabulates the values for QLQ-C30 and its scales according to the main cancer sites divided by stage of disease (early or limited, versus advanced or extensive). Age and gender-specific values are given. This enables investigators to contrast their results with those that have been found in comparable groups of patients.

In summary, the interpretation of results remains essentially qualitative. Clinical significance is subjective, and is a matter of opinion. The values and opinions of individual patients will differ, as will the opinions of the treating clinician and those of society in general. Thus, for a QoL measurement scale, it is unlikely that a single threshold value will be universally accepted as a cut-off point that separates clinically important changes from trivial and unimportant ones. However, many investigators are finding that, for a variety of scales assessing overall QoL and some of its dimensions, changes of between 5% and 10% (that is, between 5 and 10 points on the 1 to 100 scales of the QLQ-C30) are noticed by patients and are regarded by them as "significant changes".

When QoL is a major outcome measure for a clinical trial, it will be necessary to estimate the required sample size to detect the differences in QoL that are of interest. Methods for doing this are described in the manual of reference data (Fayers et al., 1998b). Before the calculation can be performed, the magnitude of the target difference must be specified. This will be based upon consideration of clinically important differences; prior information regarding plausible treatment differences; and an assessment of the feasibility of accruing the desired number of patients.

Other general references particularly worth consulting are Olschewski et al., 1994; Staquet et al., 1998 and Zee, 1991.

12. ETHICAL ISSUES

Collecting QoL data has ethical implications for both investigator and patient. It is important that patients are fully informed about the reasons for collecting QoL data. They should also be clear about the distinction between their entitlement to a professional concern about their symptoms and QoL, and their participation in "research".

12.1. Altruism

Participation in QoL studies often has no benefits for the patients themselves but is in the interest of future patients. Their results may be used to improve care and treatment in the future but those who participate in the study often do not benefit personally from participation. The aim of the QoL assessment must therefore be made clear to the patient before inclusion in the study. A separate sheet for informed consent regarding the QoL study may be of great help to the patient.

12.2. Confidentiality & Disclosure

In clinical trials, it is usually recommended that patients' completed questionnaires regarding their QoL are not shown to their physician or other personnel responsible for their treatment. If this is the case, it should be emphasized to the patient at the time of seeking informed consent that it is their responsibility to communicate any problems or symptoms to their doctor. They should be reminded of this throughout the trial. Occasionally patients may indicate such severe levels of symptoms in response to the items in a questionnaire that it should be considered an adverse event. This may give rise to a dilemma between patient safety and patient confidentiality. In this instance the data collector should return to the patient and suggest that they report this symptom to the physician responsible for their treatment. No intervention can be offered to patients who only disclose their symptoms by completing questionnaires (e.g. medication for constipation can only be prescribed if the patient tells their doctor). Patients may also use the opportunity of completing a QoL assessment to talk about and discuss other problems which may be unrelated to their treatment. This may put the person responsible for data collection in a difficult situation. It is important to ensure that the data collector has the opportunity to discuss these issues with others without breaching the patient's confidentiality.

12.3. Eligibility Criteria For Participation

Ideally completing a baseline QoL assessment should be one of the eligibility criteria for a clinical trial. Patients who have good personal reasons for not wishing to participate in a QoL study are then excluded from the study and may be unable to receive the new treatment. A thorough explanation of the aims of the QoL assessments and an assurance of anonymity may overcome these difficulties. Other patients may be unable to read or write but with appropriate assistance may still be able to participate. The patient's decision not to participate, for whatever reason, must be respected. The decision to make participation in the QoL study mandatory must be based on considerations of the nature of the research question e.g. Is it the primary endpoint? Is the study's integrity at risk if QoL assessment is missing?

12.4. Selection Bias

In clinical trials where QoL assessment is relevant, it is important that all eligible patients are included, otherwise the study may not be evaluable and those included may have participated without cause, making the study unethical. However it may also be considered unethical to coerce patients to participate in a QoL study. The problem may be diminished by explaining the rationale for including as many patients as possible in the QoL study, and reinforcing the principle of anonymity.

12.5. End Of Study Assessment

Before the study starts a decision should be made about when QoL assessment will be discontinued. During the trial a number of patients will relapse or a decision will be made that "treatment has failed". The time interval to these events is often one of the endpoints of the trial and no further clinical data is collected, only survival data. The value of QoL assessments beyond these events is debatable. QoL will be overestimated when only those whose treatment is successful remain in the study, but patients should not be burdened with excessive QoL assessments during the last few months of their lives. Some patients may feel discouraged if no one appears to take an interest in their QoL once their treatment has "failed", whilst others may be reluctant to continue participating. The study coordinator and the data manager must be informed about relapses so that proper respect can be shown when sending out reminders to patients with relapse.

12.6. Long Term Follow-up

There are a number of potential problems in studies where long-term postal assessment is planned - especially when the questionnaires are mailed from one central office which relies upon regular updates from the local centers:

- Patients may change address. Not only could this result in missing data but also in a breach of the patient's confidentiality if the new occupant opens the mail.

- Questionnaires may be sent to the home of a patient who has died, which may be distressing for their relatives.

- For some patients who have been in remission for a number of years it becomes distressing to be reminded at regular intervals that they have cancer.

- In all cases it is important that the responsible physician informs the data center about any change in the patient's circumstances on a routine basis and without delay. It is also important that the patients are informed of the long-term nature of the study. The patients can then be considered to have consented to receive questionnaires for a long time.

Even with these measures it is not possible to ensure that questionnaires will never be sent to those who have died during the course of the study. If it does happen then a letter of condolence and an apology should always be sent. Sometimes it may also help to describe the steps that have been taken to avoid the mistake and to explain to the relative the aim of the QoL study and that the patient had consented to participate.

References

Aaronson, N.K. (1995). Quality of Life and clinical trials. *Lance*t, **346**, 1-2.

Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filberti, A., Flechtner, H., Fleishman, S.B., de Haes, J., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P., Schraub, S., Sneeuw, K., Sullivan, M., Takeda, F. for EORTC Study Group on Quality of Life. (1993). The European Organisation for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *Journal of the National Cancer Institute*, **85**, 365-376.

Bjordal, K., de Graeff, A., Fayers, P.M., Hammerlid, E., van Pottelsberghe, C., Curran, D., Ahlner-Elmqvist, M., Maher, E.J., Meyza, J.W., Brédart, A, Söderholm, A.L., Arrarras, J.J., Feine, J.S., Abendstein, J.S., Morton, R.P., Pignon, T., Huguenin, P., Bottomley, A., Kaasa, S. on behalf of the EORTC Quality of Life Group (2000). A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. *European Journal of Cancer*, **36**, 1796-1807.

Bland, J.M. & Altman, D.G. (1995). Multiple significance tests: The Bonferroni method. *British Medical Journal*, **310**, 170.

Blazeby, J., Sprangers, M., Cull, A., Groenvold, M. & Bottomley, A. on behalf of the Quality of Life Group (2001). Guidelines for *Developing Questionnaire Modules*, (*3rd edition*). Brussels: EORTC.

Cella, D.F., Tulsky, D.S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., Silberman, M., Yellen, S.B., Winicour, P., Brannon, J., Eckberg, K., Lloyd, S., Purl, S., Blendowski, C., Goodman, M., Barnicle, M., Stewart, I., McHale, M., Bonomi, P., Kaplan, E., Taylor, S., Thomas, C.T. & Harris, J. (1993). The Functional Assessment of Cancer Therapy Scale: Development and Validation of the General Measure. *Journal of Clinical Oncology*, **11**, 570-579.

Coates, A., Porzsolt, F. & Osoba, D. (1997). Quality of Life in Oncology Practice: Prognostic Value of EORTC QLQ-C30 Scores in Patients with Advanced Malignancy. *European Journal of Cancer*, **33**, 1025-1030.

Cull, A., Sprangers, M.A.G. , Bjordal, K.& Aaronson, N. on behalf of the EORTC Quality of Life Study Group. (1998). *EORTC Quality of Life Study Group Translation Procedure*. Brussels: EORTC.

Curran, D., Bacchi, M., Hsu Schmitz, S.F., Molenberghs, G. & Sylvester, R.J. (1998a). Identifying The Types of Missingness in Quality of Life Data From Clinical Trials. *Statistics in Medicine*, **17**, 739-756.

Curran, D., Fossa, S., Aaronson, N., Kiebert, G., Keupens, F. & Hall, R. (1997). Baseline quality of life of patients with advanced prostate cancer. *European Journal of Cancer*, **33**, 1809-1814.

Curran, D., Molenberghs, G., Fayers, P. & Machin, D. (1998b). Aspects of incomplete quality of life data in randomised trials: II: Missing forms. *Statistics in Medicine*, **17**, 697-709.

Curran, D., van Dongen, J.P., Aaronson, N., Kiebert, G., Fentiman, I.S., Mignolet, F. & Bartelink, H. (1998c). Quality of life of early breast cancer patients treated with mastectomy or breast conserving procedures: Results of EORTC trial 10801. *European Journal of Cancer*, **34**, 307-314.

de Haes, J.C.J.M., van Knippenberg, F.C.E. & Neijt, J.P. (1990). Measuring psychological and physical distress in cancer patients: Structure and application of the Rotterdam Symptom Checklist. *British Journal of Cancer*, **62**, 1034-1038.

Diggle, P.J., Liang, K.-Y. & Zeger, S.L. (1994). *Analysis of Longitudinal Data*. Oxford University Press: Oxford.

Fayers, P., Curran, D. & Machin, D. (1998a). Incomplete quality of life data in randomized trials: *Missing items. Statistics in Medicine*, **17**, 679-696

Fayers, P., Weeden, S. & Curran, D. on behalf of the EORTC Quality of Life Study Group. (1998b). *EORTC QLQ-C30 Reference Values*. Brussels: EORTC.

Fayers, P.M., Aaronson, N.K., Bjordal, K.,, Groenvold, M., Curran, D. & Bottomley, A. on behalf of EORTC Quality of Life Study Group. (2001). *The EORTC QLQ-C30 Scoring Manual(3rd edition)*. Brussels: EORTC.

Fayers, P.M., Hopwood, P., Harvey, A., Girling, D.J., Machin, D. & Stephens, R. (1997). Quality of life assessment in clinical trials - guidelines and a checklist for protocol writers: The UK Medical Research Council experience. *European Journal of Cancer*, **33**, 20-28.

Fayers, P.M. & Machin, D. (2000). *Quality of life – assessment, analysis and interpretation*. John Wiley & Sons Ltd: Chichester.

Gelber, R.D. & Gelber, S. (1995). *Quality of life assessment in clinical trials*. Recent advances in clinical trial design and analysis. Kluwer Academic Publishers: Boston.

Goldstein, H. (1995). *Multilevel Statistical Models*. Edward Arnold: London.

Gotay, C.C., Korn, E.L., McCabe, M.S., Moore, T.D. & Cheson, B.D. (1992). Quality-of-life assessment in cancer treatment protocols: Research issues in protocol development. *Journal of the National Cancer Institute*, **84**, 575-579.

Hand, D.J. & Crowder, M. (1996). *Practical Longitudinal Data Analysis*. Chapman & Hall: London.

Hjermstad, M.J., Fayers, P.M., Bjordal, K. & Kaasa, S. (1998). Health-related quality of life in the general Norwegian population assessed by the EORTC core Quality of Life Questionnaire - The QLQ-C30 (+3). *Journal of Clinical Oncology*, **16**, 1188-1196.

King, M.T. (1996). The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Quality of Life Research*, **5**, 555-567.

Lindsey, J.K. (1993). *Models for Repeated Measurements*. Oxford University Press: Oxford.

Matthews, J.N.S., Altman, D.G., Campbell, M.J. & Royston, P. (1990). Analysis of serial measurements in medical research. *British Medical Journal*, **300**, 230-235.

Olschewski, M., Schulgen, G., Schumacher, M. & Altman, D.G. (1994). Quality of life assessment in clinical cancer research. *British Journal of Cancer*, **70**, 1-5.

Osoba, D. (1992). The Quality of Life Committee of the Clinical Trials Group of the National Cancer Institute of Canada: Organisation and Functions. *Quality of Life Research*, **1**, 211-218.

Osoba, D., Pater, J.L. & Zee, B. (1994). Effective anti-emetic therapy improves quality of life (QoL) after moderately emetogenic chemotherapy (MEC). *Quality of Life Research*, **4**, 467-468.

Osoba, D., Rodrigues, G., Myles, J., Zee, B. & Pater, J. (1998). Interpreting the significance of changes in health-related quality of life scores. *Journal of Clinical Oncology*, **16**, 139-144.

Rosendahl, K.I., Curran, D., Kiebert, G., Cole, B., Weeks, J.C., Denis, L.J. & Hall, R.R. (1997). A quality-adjusted survival (Q-Twist) analysis of EORTC trial 30853 comparing maximal androgen blockade (MAB) with orchidectomy in patients with metastatic prostate cancer. In *American Society of Clinical Oncology* (ASCO), Vol. 16. pp. 312 (A1113).

Schipper, H., Clinch, J., McMurray, A. & Levitt, M. (1984). Measuring the quality of life of cancer patients: The Functional Living Index-Cancer: Development and Validation. *Journal of Clinical Oncology*, **2**, 472-483.

Slevin, M.L., Stubbs, L., Plant, H.J., Wilson, P., Gregory, W.M., Armes, P.J. & Downer, S.M. (1990). Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses and general public. *British Medical Journal*, **300**, 1458 -1460.

Sneeuw, K.C.A., Aaronson, N.K., Sprangers, M.A.G., Detmar, S.B., Wever, L.D.V., Schornagel, J.H. (1997). The value of caregiver ratings in evaluating the quality of life of patients with cancer. *Journal of Clinical Oncology*, **15**,1206-1217.

Sneeuw, K.C.A., Aaronson, N.K., Sprangers, M.A.G., Detmar, S.B., Wever, L.D.V., Schornagel, J.H. (1998). Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *Journal of Clinical Epidemiology*, **51**, 617-631.

Staquet, M.J., Hays, R.D. & Fayers, P.M. (1998). *Quality of Life Assessment in Clinical Trials: Methods and Practice.* Oxford University Press: Oxford.

Stephens, R.J., Hopwood, P., Girling, D.J. & Machin, D. (1997). Randomized trials with quality of life endpoints: Are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? *Quality of Life Research*, **6**, 225-236.

Tandon, P.K. (1990). Applications of global statistics in analysing quality of life data. *Statistics of Medicine*, **9**, 819-827.

Tannock, I.F., Osoba, D. & Stockler, M.R. (1996). Chemotherapy with mitoxantrone plus prednisone or prednisone alone for symptomatic hormone-resistant prostate cancer: A Canadian randomized trial with palliative endpoints. *Journal of Clinical Oncology*, **14**, 1756-1764.

Troxel, A.B., Fairclough, D.L., Curran, D. & Hahn, E.A. (1998). Statistical Analysis of Quality of Life Data in Cancer Clinical Trials. *Statistics in Medicine*, **17**, 653-666.

Vachalec, S., Bjordal, K., Bottomley, A., Blazeby, J., Flechtner, H. & Ruyskart, P. on behalf of the EORTC Quality of Life Group (2001). *Item Bank Guidelines*. Brussels: EORTC.

Ware, J.E. (1987). Standards for validating health measures: definition and content. *Journal of Chronic Diseases*, **40**, 473.

Zee, B. (1991). In: In *Effects of Cancer on Quality of Life*, Osoba, D. (ed). CRC Press.

Appendices

1. QLQ-C30 v3

2. List of available QLQ-C30 translations

3. QLU staff and contact addresses

4. List of available modules, translations and contact names

5. List of joint scientific committee members and contact addresses

6. Codes for missing data

7. List of all EORTC studies assessing QoL

8. EORTC Guidelines for administration of QoL questionnaires

Requests for permission to use the EORTC QLQ-C30 and the modules, or to reproduce or quote materials contained in this manual, should be addressed to:

<div align="center">

Quality of Life Unit,
EORTC Data Center
Avenue E. Mounier 83 – Bte 11,
1200 Brussels
BELGIUM

Tel: +32 2 774 1680
Fax: +32 2 779 4568

</div>