

EORTC QUALITY OF LIFE GROUP

Guidelines for Developing Questionnaire Modules

**Colin Johnson, Neil Aaronson, Jane M Blazeby,
Andrew Bottomley, Peter Fayers, Michael Koller,
Dagmara Kuliś, John Ramage, Mirjam Sprangers,
Galina Velikova, Teresa Young**

April 2011

Fourth Edition

On Behalf of

**EORTC
Quality of Life Group**



CONTENTS

PREFACE	2
1 INTRODUCTION	3
2 OVERVIEW OF MODULE DEVELOPMENT	4
3 DETAILED DESCRIPTION OF MODULE DEVELOPMENT	5
3.1 Phase 1: Generation of QL issues	6
3.2 Phase 2: Construction of the item list	10
3.3 Phase 3: Pre-testing	12
3.4 Phase 4: Field-testing	14
3.5 Minimal requirements for psychometric validation of the module	16
3.6 Description of modules in different phases of development	17
3.7 Co-ordination of module development and Quality Assurance	18
3.8 Preparation of Documents	20
3.9 Translation procedure during module development (Phases 1 to 3)	22
4 UPDATING EORTC QL EXISTING MODULES	24
5 MERGING TWO MODULES	26
6 PUBLICATION OF MODULES	27
7 OWNERSHIP AND USE OF MODULES	28
8 APPENDICES	30
8.1 Matrix for directing patient recruitment	30
8.2 Generation of relevant quality of life issues in Phase 1 Example of a patient review	31
8.3 Generation of relevant quality of life issues in Phase 1 Example of a health care provider interview	34
8.4 Decision rules for selection of QL issues in Phase 1	36
8.5 Example of a patient interview in Phase 3	37
8.6 Decision rules for inclusion or exclusion of items in Phase 3	39
8.7 Debriefing questionnaire in Phase 4	41
8.8 Item Response Theory for scale structure & selection of items in Phase 4	42
8.9 Template for Report on module construction (Phases 1 to 3).	43
8.10 Appendix 9 - Flow chart of module development process	45
9 REFERENCES	46

PREFACE

The EORTC Quality of Life Group (QLG) guidelines for module development have been shown to be a useful tool for questionnaire development. Modules that have been produced following these guidelines have exhibited good levels of psychometric and cross-cultural validity. The guidelines also allow those who use modules (individuals and industry) to understand the rigorous methodology of module development.

Experience with module development since the last revision in 2002 has highlighted areas where the guidelines require further development or refinement. These areas include: (i) alternative methods of identifying relevant quality of life (QL) issues or psychosocial issues (ii) links to the newly developed Item Bank at the EORTC Quality of Life Department at the EORTC Headquarters, (iii) translation of modules and (iv) changes to the methods used to produce Phase 4 modules. We have also added new sections on procedures for updating modules and merging two related modules.

These amendments are included in the current version.

We would like to thank all members of the EORTC QLG who have contributed to this document. We hope that these updated guidelines will continue to ensure uniformly high quality across modules. Users who have comments or questions are encouraged to contact the authors to enable them to further improve the guidelines.

We are particularly grateful to Ann Cull and Mogens Groenvold who contributed to the Third Edition (2002) of the Guidelines, which formed the basis for the current revised version, also to Sheila Scott-Sanderson who proof read and set up the final document.

April, 2011

*Colin Johnson
Neil Aaronson
Jane M Blazeby
Andrew Bottomley
Peter Fayers
Michael Koller
Dagmara Kuliś
John Ramage
Mirjam Sprangers
Galina Velikova
Teresa Young*

1 • INTRODUCTION

An essential aspect of the "modular" approach to QL assessment adopted by the EORTC QLG is the development of modules specific to tumour site, treatment modality, or a QL dimension, to be administered in addition to the core questionnaire (the EORTC QLQ-C30). The modules, like the core questionnaire, are primarily designed for use in cancer clinical trials, but can be used in other research settings as well.

Guidelines are provided to assist Module Developers and to standardise the module development process in order to ensure uniformly high quality across the modules.

Modules may relate to QL issues affecting particular tumour types (e.g., primary site, metastatic site), aspects of care (e.g., patient satisfaction), patients' psychological needs or experiences (e.g., information), or spiritual well being. Modules have been developed for defined patient groups (e.g., elderly; palliative care) and for generic cancer symptoms or treatment side-effects (e.g. fatigue).

Researchers who are considering developing a new module should discuss this in the first instance with the Chair of the Module Development Committee who can advise on procedures and give a preliminary view of the suitability of the proposed module.

2 • OVERVIEW OF MODULE DEVELOPMENT

The module development process consists of four phases: (1) generation of relevant QL issues; (2) conversion of the QL issues into a set of items; (3) pre-testing the item list or preliminary module questionnaire; and (4) large-scale international field-testing.

Module development should be conducted simultaneously in several languages and cultural groups. At each of these steps, the Module Developers should ensure cross-cultural consistency.

Phases 1 and 2 of the construction process should include at least three languages and countries, to include one representing each of the following groupings: (a) English-speaking countries (e.g., Australia, Canada, United Kingdom, United States); (b) Northern Europe (e.g., Austria, Denmark, Germany, The Netherlands, Norway, Sweden); and (c) Southern Europe (e.g., French-speaking part of Belgium, France, Greece, Italy, Spain).

Phase 3 should be conducted in a wider range of countries and regions: it is recommended to use at least six countries and to include, in addition to the three regions described above, at least one country from Eastern Europe and at least one non-European country.

The provisional item list and the provisional module should be initially developed in English. These English versions should be sent to all national co-ordinators for direct feedback on the translatability of the items.

Phase 4 is an international field test. As many countries as practical should be involved (including at least all those participating in Phase 3). Patients complete the EORTC QLQ-C30, the provisional questionnaire and a short debriefing interview. Translations of the Module will be provided through the Translation Unit of the EORTC Quality of Life Department.

The Module Developers should collaborate as a group of independent researchers and should seek to achieve consensus after each crucial step. The Developers should regularly review progress, data collection and analysis of the responses in order to agree upon:

1. The formulation of the QL domains;
2. The list of QL issues derived from the literature to be put to patients and health care professionals in different countries;
3. The final list of QL issues to be included in the provisional item list;
4. Final list of items (provisional module) that will be field-tested;
5. The final validated Module;

Most Module Development Groups find it helpful to meet every six months during the bi-annual QLG meetings. As a minimum the development of a module should:

- involve at least one member of the EORTC QLG
- be carried out according to the Guidelines
- report progress regularly at the bi-annual meetings of the EORTC QLG. Module Development Committee (in writing and preferably in person)

3 • DETAILED DESCRIPTION OF MODULE DEVELOPMENT

Protocol design, approvals and Reports

Module construction begins with a clear description of the research question and the target population for which the module will be designed. The need for a new module should be clearly demonstrated with a written proposal. This should be submitted for approval to the Chair of the MDC before work begins. The developers will be invited to make a presentation of their proposal at the next meeting of the MDC. Consideration should be given to which patients will be appropriate for Phase 1 and Phase 3. Modules may be used in clinical research at any stage of the patient's illness. Module Developers should identify all relevant treatment groups and stages which are to be included in the development. The numbers of patients required in Phases 1 and 3 are different, but in both Phases the sample should be evenly distributed across the relevant categories. A matrix such as the one depicted in Appendix 1 should be developed to guide patient recruitment. Module developers should prepare a detailed Protocol of Phases 1 to 3, which will be peer-reviewed either during the process of grant applications, or by the Module Development Committee.

A Protocol for a Phase 4 study (field study to confirm the psychometric properties of the module) is prepared by representatives of the working group which has been responsible for developing the module. Phase 4 studies may be conducted through the EORTC Headquarters (QL Department), in studies coordinated outside EORTC headquarters, or by analysis of data collected during use of the provisional module in EORTC or other clinical trials. A detailed protocol is required to specify the study aims and procedures for field testing. The protocol is peer-reviewed either during the process of grant applications, or by the Module Development Committee and by the EORTC Protocol Review Committee (PRC) when the study is conducted through EORTC HQ and QL Department.

The study coordinator is responsible for ensuring strict adherence to ethical guidelines, research governance, quality assurance, data management and statistical analysis procedures and their rigorous documentation. Documentary evidence of these aspects should be submitted to the Chair of the MDC with the Phase 1 & 2 Report and with the Phase 3 and Phase 4 Report (which may be a manuscript for publication), to be archived at the QLD in case of future regulatory inspection.

In the following text it is important to note the distinction between an 'issue' and an 'item'. An issue is a neutrally phrased descriptive statement (e.g. 'ability to eat, dress or wash yourself', or 'social contact with friends') whilst an item is a question which can be phrased in a positive, negative or neutral manner (E.g. Do you need help with eating, dressing or washing yourself? Have you had trouble having social contact with friends?).

Language

The working language for module development is English. For presentation to patients, issues and items will be translated into the patient's own language, following EORTC QLG Translation procedures.

3.1 PHASE 1: GENERATION OF QL ISSUES

This phase is aimed at compiling an exhaustive list of relevant QL issues that cover the domain(s) of interest. In the process of compiling this list, three sources are used:

- (i) Literature (including existing questionnaires);*
- (ii) Patients with the relevant condition and all relevant stages of disease and treatment;*
- (iii) Health care professionals (e.g., physicians, nurses, psychologists, dieticians) with clinical expertise in the area of the module.*

The following text provides a guideline for using these resources, but adaptations are permissible provided the following are determined:

- a) Relevance: the extent to which patients have experienced issues on the list including problems, limitations and positive experiences;
- b) Breadth of coverage: that the list includes all significant issues;
- c) Relative importance of issues.

Whilst it is necessary to review the literature before beginning the interviews, the recruitment of patients and health care professionals can proceed simultaneously. In considering the information gathered from interviews, the responses of patients should be given highest priority. QL measures must be derived in a patient-centred way, to ensure greatest content validity (Food and Drug Administration 2009). However if patient burden is a concern, then Module Developers may choose to conduct interviews with health care professionals first. The list may then be adapted. Another option may be to conduct focus groups with relevant patient groups or health care professionals.

It is essential to ensure that high levels of content validity are achieved. Patient interviews are the most important of four steps identified by Rothman et al (2009) to ensure that high content validity is achieved and demonstrated. (The other steps recommended are: good conceptual match between instrument and its purpose, demonstration that the most relevant and important content is included and good documentation of all modifications made to items or modules.)

(i) Literature searches

Literature searches should be conducted on MEDLINE and on other relevant databases (e.g., PSYCHINFO) to ensure that the relevant QL issues have been identified. From this and other sources (such as PROQOLID) existing, relevant questionnaires should be reviewed (i.e., general quality of life questionnaires and disease-specific questionnaires). A list of all questionnaires identified and finally a list of all potentially relevant QL issues should be created.

Module developers should use a systematic approach to identification and review of previous studies and questionnaires. Details of the literature review process (databases, key words, selection criteria for inclusion of papers) should be presented in the Phase 1 report.

Phase I aims at identifying an exhaustive list of issues in order to achieve content validity. The clinical literature around the domain should be reviewed. In the development of new modules or topics that are more abstract than symptoms (e.g. psychological or sociological concepts) careful attention must be given to the theoretical literature. Issues may also need to be selected based on their ability to reflect an accepted theoretical framework.

If, for example, a module to assess coping was to be developed it would have to relate to current theories about coping.

The search may also identify existing questionnaires, or relevant questions may exist within general QL questionnaires. The underlying issues should be extracted from the questions and added to the list derived from the theoretical and clinical literature.

Issues may arise which appear very similar. In the early phases it is better to include all similar issues and present all of these to the patients and health care professionals for evaluation. This is preferable to the researcher making a judgement and excluding some issues without input from patients.

(ii) Interviews with patients

Inclusion of qualitative or semi-structured interviews at the earliest stage of the module development is essential to ensure content validity. This has been a cornerstone of EORTC Module Development for many years and is now endorsed by the Food and Drugs Administration guidance (2009). Patient interviews are the most important of four steps identified by Rothman et al (2009) to ensure that high content validity is achieved and demonstrated. Appropriate methods for developing conceptual issues and frameworks for qualitative interview research, developing the interview discussion guide, reaching saturation, analysis of data, developing a theoretical model, item generation and cognitive debriefing are available (Brod et al 2009).

Patients should be recruited from a variety of locations, including hospital inpatient and outpatient clinics, community settings and self-help groups. The nature of the module and the target population will help identify the most appropriate sources.

Patient selection

Strict eligibility criteria should be adopted to ensure that subjects adequately represent the target population for which the module is being devised. It is recommended that five to 10 patients should be interviewed from each different treatment group or disease stage and similar numbers of patients should be recruited from each country participating in Phase 1. The age and gender distribution of recruited patients should reflect that of the target population. Interviews should continue until no new issues arise.

Interview technique

There are two main approaches to gathering information about quality of life issues from patients with a particular condition, or who are undergoing a particular treatment. The researcher may ask the patient to describe their experience and allow the patient to provide information freely, or in response to predetermined questions in a semi-structured interview. Alternatively, the researcher may show the patient existing relevant material to begin the discussion and to prompt the patient's description of significant issues. These two approaches may be used sequentially, that is an open or semi-structured interview, followed by a review of written material such as the EORTC QLQ-C30 and a list of possibly relevant issues, during which the patient is encouraged to comment on the issues and to score each issue for relevance to themselves.

Breadth of coverage

All information provided by the patient should be recorded, preferably on audiotape or a digital audio file and then transcribed for later analysis. This method ensures accuracy of wording as used by patients and reduces any bias that may result from the selective noting of patients' comments by the researcher

During the open or semi-structured interview, the patient should be encouraged to consider all issues which they believe to be relevant to the condition. The interview design and prompts used will be decided by the nature of the module under development. The interviews should continue until no new issues are raised. This requires a constant review of accumulating data, to assess whether new information has been gathered.

Techniques for this and for documenting the process, have been described (Kerr et al 2010). A minimum of 20 patients should be interviewed. Usually no more than 30 are required.

Review of provisional list of issues

When the Module Developers have a provisional list of issues from patient interviews and the literature review, this list together with the EORTC QLQ-C30 are administered to a limited number of patients (usually not more than ten in total), followed by a debriefing interview to determine what the various issues mean to the patient, the extent to which patients have experienced the problems, limitations, or positive experiences during the period of their disease and to check for any significant omissions ("debriefing"). Such cognitive interviews may be critical to refine items and avoid ambiguity or other difficulties in the final module (Fortune-Greeley et al 2009). These interviews may be conducted individually or in a focus group.

Patients should be encouraged to explain their response to each item as they read through the EORTC QLQ-C30 and the list of issues ("think aloud" technique). After completion of the item list, a structured interview should be used to explore additional issues suggested by reflection on the lists provided, or by the patient's own experiences.

Relevance and Importance

When interviewing patients about issues it is usual to ask whether particular issues have ever arisen for them. If an issue is common, it will be retained as an item. It may be appropriate to discover whether an issue matters to the person, which may be different from their experience of that issue. Patients should be asked to rate issues for relevance and for importance.

During the interview to determine relevance it is important to avoid ambiguity when interpreting responses. Some patients may never have considered a particular issue, for example, 'being in control'. The interviewer should ask the patient to consider each issue and to score its relevance to their own situation using a scale of 1 (not relevant) to 4 (very relevant). Appendix 2.

To determine relative importance, patients may be asked to rate each item for importance on a 4 point Likert scale, or to choose a limited number (e.g., 5 to 10) of issues which troubled them most (or caused the greatest problems/nuisance/ distress) or which they valued particularly highly. An example of an interview is provided in Appendix 2. Patients may also be asked to identify issues which they think should definitely be included or definitely excluded.

(iii) Interviews with health care professionals

The provisional list of issues and the core instrument should be presented to health care professionals, for feedback on appropriateness of content and breadth of coverage. At least five health professional should be included; it is usually unnecessary to recruit more than 20 individuals, drawn from all countries represented in Phase1. The health care professionals may be of any relevant discipline and should have experience with treating patients belonging to the target population.

The list of QL issues may be administered in the form of a structured, personal interview in which basically two questions are asked: (a) are issues included which the specialists consider irrelevant for this patient group and if so, why do they consider these issues irrelevant?; and (b) are there issues missing from this list that the specialists consider relevant and if so, why do they consider these issues relevant?

To establish the relative importance of the QL issues, the specialists should be asked to identify a subset (e.g., 5 to 10) of issues that, in their opinion, affect patients' QL most profoundly and which should definitely be included in the final questionnaire.

Appendix 3 provides an example of a detailed interview protocol. The Module Developers will consider the comments of these specialists during selection of items for inclusion in the item list for Phase 3.

Amendments of the list of issues

On the basis of the responses collected in the interviews (or focus groups) the list of QL issues may be amended during Phase 1. The aim of Phase 1 is to develop a comprehensive list of issues and researchers are discouraged from removing issues at this stage; new issues arising during Phase 1 should be added to the list and presented to further patients for evaluation.

Variations in approach

In practice it may be advantageous to elicit patients' and health care professionals' opinions in slightly different ways. The following variations in the conduct of the interviews are acceptable:

1. Order of interviews

Patients and health care professionals may be consulted simultaneously or sequentially. In the development of patient-centred measures it is important to give maximum weight to the views expressed by patients during development. In practice, interviews with health care professionals are likely to function as a means of ensuring that all relevant issues have been considered for inclusion.

2. Interview format

(a) Focus group interviews (Krueger & Casey) may be conducted instead of individual interviews, provided that the researcher has the requisite skills and is able to bring together 10 to 15 patients belonging to the target population. Well conducted focus groups where patients are encouraged to interact with each other rather than just an interviewer can provide a rich source of qualitative data. Seeking patients' and health professionals' views unaided, before imposing an interviewer generated list may help identify missing issues.

(b) In some cases it may be appropriate to interview patients and health care professionals in an open rather than in a pre-set way and ask them to describe their opinion with regard to the relevant dimension, prior to administering the list of QL issues (Groenvold, 1997).

3. Patient groups

In some instances it may be relevant to consult self-help groups in eliciting patients' experiences in addition to or instead of consulting patients in clinical settings.

Selection of issues

The lists of issues from all sources should be reviewed by the Module Developers to produce a single, comprehensive list of issues for formulation into the provisional item list in Phase 2. Module Developers should agree on the decision rules to be used before the selection of items takes place. If there is disagreement between the views of patients and professionals, the views of patients will usually take precedence. In every case, the reasons for inclusion or exclusion of items should be given in the Phase 1 and 2 Report

Decision rules may vary somewhat across modules. Examples of decision rules suitable for adaptation are given in Appendix 4. Issue lists should be reviewed for overlap between issues and care should be taken that potential new issues are not already covered in the core questionnaire. When there is a very large number of issues (e.g., more than 50), most weight should be given to the patient responses during selection of issues.

3.2 PHASE 2: CONSTRUCTION OF THE ITEM LIST

The list of QL issues is converted into questions with the format and time frame compatible with the EORTC QLQ-C30. That is, questions refer to the patient's experience *during the last week* and the response is recorded on a 4 point Likert scale. Exceptions to this one week timeframe may be acceptable. For example, if the issue or problem area is unlikely to be captured with a one week timeframe, it can be extended. Any proposed change in timeframe needs to be justified.

The QL Group has considerable expertise in module development and construction of new items. It is strongly recommended that the Module Development group should include at least one individual with experience in questionnaire construction.

The EORTC Quality of Life Group Item Bank

At this stage, it is important to avoid duplication of effort and to ensure uniformity across modules. The wording of new items should, as far as possible, be consistent with existing EORTC QLG modules. Existing items should be used unless there are strong arguments not to do so (e.g., when these items appear to perform poorly). This maintains consistency and reduces the requirement for translation, as existing translations are available in the Item Bank.

The Item Bank maintained by the EORTC QL Department should be searched for items related to the issues from the Phase 1 list. The Item Bank is regularly updated and receives new translations every month. The Item Bank may yield several items that cover the same issue. The most appropriate item for the module under development should be chosen. If several items addressing similar issues are identified, it may be necessary to test these in samples of patients from the target population. If the modules from which items are extracted are still under development, then the explicit permission and co-operation of the author(s) should be obtained.

Further information about the purpose and methods of using the Item Bank may be accessed online (www.eortc.be/itembank2). A password may be obtained from the Quality of Life Department (www.eortc.be/qol)

Other resources

In order to save time and effort, existing questionnaires, developed by other research groups, may be consulted for their wording. Subscriptions to the PROQOLID database are available from the Chair of the Module Development Committee to assist in searches. This database is a comprehensive, searchable record of quality of life and other patient reported outcome measures. The format of "existing" questionnaire items may require adaptation to achieve consistency within the module. The explicit consent of the questionnaire constructors should be obtained prior to including the items in the module.

Item construction

If a new item is needed for the new module, it is important to be aware of the major methodological considerations in item construction. The question should be clear, brief and unambiguous. Conditional questions should be broken into their component parts, for example “1. Do you have a stoma? (yes/no). 2. If yes, how much have you been troubled by leakage from the stoma appliance?” When the module elements of interest cover positive issues, the resulting items should be positively phrased (i.e., in terms of abilities, capacities and positive experiences). Other items should be negatively worded (i.e., in terms of problems, limits in functioning and negative experiences).

Module Developers should be alert in order to avoid possible confusion and biased responses due to differences in the orientation of items (negative versus positive). Patients' attention can be drawn to these differences for example, by highlighting or underlining. Items of similar orientation should be grouped together in the item list. Further advice on item construction can be found in standard textbooks (Converse and Presser 1986; Fayers and Machin 2007; Streiner and Norman, 2003).

Scale structure

The forming of multiple item scales should be anticipated by including several items relative to similar constructs. Scoring will be simplified if all items in a scale are negatively or positively phrased. However, if this is not feasible, it should be noted that the EORTC approach to item and scale *scoring* requires that all items and scale relating to functioning be scored in a positive direction (which may require recoding of negatively worded items), while all items and scales related to symptoms and side effects of treatment be scored in a negative direction.

Conditional questions

If the question relates to the impact of a certain symptom, intervention or side effect, consideration should be given to how patients who are not experiencing that particular issue will answer the question (e.g., if asking whether pain medication helps, how will patients who do not take pain medication respond?). Similarly, responses about the impact of a patient's ability or capacity may depend on whether the patient uses that ability (e.g., sexual functioning). Conditional (Yes/No) questions may be considered in this context. “For example: Do you take pain medication? If yes, please answer the following question”

Consultation of health care professionals

The resulting provisional list of items should be reviewed for clarity and overlap by persons with expertise or knowledge of questionnaire development or of the target population, other than those who were involved in step 1 (e.g. colleagues, patients). It may be advisable to present the provisional module to one or two health care professionals (who may have been involved in Phase 1) for review and to consult the Chair of the MDC or of the relevant EORTC Tumour Group, to ensure breadth of coverage. On the basis of these final comments the list of items may require further adaptation, before it is administered to patients in the pre-testing phase.

3.3 PHASE 3: PRE-TESTING

Aim

The aim of pre-testing the module is to identify and solve potential problems in its administration (e.g., the phrasing of questions, the sequence of questions) and to identify missing or redundant issues.

Even if all items are from existing questionnaires, there is still an obligation to pre-test the module, because:

- 1) The meaning of questions can be affected by the context of the neighbouring questions;
- 2) Items may require adaptation when used in different languages and cultural settings than those of the initial development (that is in Phases 1&2);
- 3) Questions developed originally for a particular target group may perform differently when applied in a new setting.

Pre-testing consists of:

1. *Administering the EORTC QLQ-C30 and the provisional module to patients belonging to the target population, however were not involved in Phase 1, to obtain a response score for each item, together with rating of relevance and importance;*
2. *Conducting structured interviews with each patient after completion of the module to ensure completeness and acceptability of the items in the list.*

Patient sample

Strict eligibility criteria should be defined to ensure that subjects adequately represent the target population for which the module is being devised. A sample matrix should be drawn up to include all relevant treatments (e.g. surgery, chemotherapy, radiotherapy) and patient groups. The Module Developers may choose to group patients by treatment stage (pre-treatment, during treatment, post-treatment) or by disease stage (localized (curable), locally advanced (incurable), or metastatic) as appropriate for the module.

In Phase 3, each cell of the sample matrix should contain at least 15 patients.

Examples of possible sample matrices are shown in Appendix 1.

Administration of provisional module/item list

Each patient should complete the EORTC QLQ-C30 and the new module/item list. The responses will be considered in the final analysis of items. In addition, each item of the new module/item list should be rated by each patient for “importance” and “relevance” to that individual. Importance and relevance may be scored as yes/no, on a 4 point Likert scale or by selection or ranking of the most important/relevant items.

The structured interview

The interview should, in principle, be directed to each item separately and should invite further comments about:

- (1) *The particular experience to which the item refers (e.g. is this experience related to your disease or treatment?);*
- (2) *The wording of the item itself (e.g. was the item difficult to respond to? was the item annoying, confusing or upsetting? And how would you have asked this question?).*

If there is a large number of items (e.g. more than 20), the time involved in inquiring about each individual item may be prohibitive. In those cases the questions may be directed towards the entire module (e.g. *were there questions that you found difficult to answer? were there questions that you found annoying, confusing or upsetting? and do you have other comments about these questions?*).

These general questions may then be supplemented by the further probing of selected module items that are expected to cause some difficulty or items that appear to be troublesome during the interview.

The pre-testing interview should be completed with two questions directed to the entire questionnaire (i.e., core questionnaire plus module):

- (1) *Were there questions that you found irrelevant?*
- (2) *Can you think of additional issues that are relevant for you but are not included in this questionnaire?*

On the basis of this pre-testing phase, the provisional questionnaire may require adaptation. Examples of a detailed interview protocol as well as decision rules for deletion, addition and rewording of items are provided in Appendix 5.

Analysis and retention/deletion of items

Any difficulties arising in the wording or translation of items should be corrected. This may require changes to only one language, or to all languages. Each item should be considered for retention or rejection according to any comments made by patients: items which viewed as irrelevant by a substantial number of patients should be considered for rejection. If many items are designated as important, they should be ranked to assess the most important items. Items which are upsetting may benefit from modification but should not be rejected outright.

Clear decision rules should be defined by the Module Developers before analysis of the Phase 3 responses. Examples are shown in Appendix 6. Although the provisional item list may be long, to ensure all issues are considered, the threshold for retention used in Phase 3 should be set relatively high (taking account of all the features described above), to retain only those items that are essential, thereby minimising respondent burden.

Preliminary testing of hypothesized scale structure

It is possible to carry out some preliminary, albeit limited testing of the psychometric properties of the provisional module in Phase 3. It is likely that, during Phase 2, a number of items were generated that are hypothesized to form a multi-item scale. These scales will be tested fully in Phase 4, but it may be appropriate to conduct preliminary testing of the hypothesized scales in Phase 3, provided an adequate number of patients (sample size) are recruited to support such analyses. Reliability of hypothesized scales may be tested using Cronbach's alpha coefficient and simple correlations or more complex methods (e.g., factor analysis) may be used to investigate the hypothesized scale structure. Depending on numbers, some form of validity testing (e.g., known group comparisons) can be done (e.g., patients on and off treatment). Final testing of scale structure, reliability, validity and responsiveness to change over time requires larger numbers of patients and is carried out in Phase 4.

3.4 PHASE 4: FIELD-TESTING

Aim

The module and its scale structure should be field-tested in a large, international group of patients in order to determine its acceptability, reliability, validity, responsiveness and cross-cultural applicability.

It is necessary to field test the module because: 1) the sample size needed to carry out the requisite psychometric evaluation is substantially larger than that used typically in Phase 3; 2) completion of the module in Phase 3 is typically done in the presence of a researcher and the questionnaire may perform differently when completed without such supervision; 2) items may require adaptation when used in different languages and cultural settings than those of the initial development (that is in Phases 1 and 3).

Field-testing consists of:

1. Administering the EORTC QLQ-C30 and the provisional module to patients belonging to the target population, but who were not involved in Phases 1 or 3; and
2. Completion of a debriefing questionnaire by each patient after completion of the module.

Patient sample

Subjects should represent all groups in the target population for which the module is being devised. A sample matrix should be drawn up to include all relevant treatments (e.g. surgery, chemotherapy, radiotherapy) and patient groups. The Module Developers may choose to group patients by treatment stage (before, during, after, palliative) or by disease stage (localized (curable) locally advanced (incurable), or metastatic) as appropriate for the module. The sample matrix may be similar to that used in Phase 3 (Appendix 1), but module developers will need to take account of planned known group comparisons, accessibility of patients in different treatment groups or stages and the subject matter of the module when planning Phase 4 recruitment. Sample size will depend, in part, on the number of items in the Module (see below); there should be adequate numbers in each of the cells of the sample matrix.

In order to determine the **acceptability** of the module, patients should respond to the debriefing questions: (a) *How long did it take you to complete the questionnaire?* (b) *Did anyone help you to complete the questionnaire and, if so, what kind of help and how much help was provided?* (c) *Were there questions that you found confusing or difficult to answer?* (d) *Were there questions that you found upsetting?* and (e) *Please use the space below if you have other comments about the questionnaire* (Appendix 7).

Scale structure and reliability

It is advantageous to combine items into scales dealing with different domains of QL, when items are related to the same clinical or psychosocial concept. A range of analyses can be

conducted to test empirically the module's hypothesised scale structure and to establish scale reliability. For example, multi-trait scaling analysis and exploratory or confirmatory factor analysis can be used to examine the extent to which the items of the module can be combined into the hypothesised multi-item scales (Fayers and Machin 2007,).

The *internal consistency* of the multi-item scales can be assessed by Cronbach's alpha coefficient. Reliability of a magnitude of 0.70 or greater is desirable for group level data (Fayers and Machin 2007,). However, aggregating symptoms or side effects (so called 'causal' indicators in relation to overall quality of life) into a summated scale should be done with greater caution than other aspects, such as depression (for which items may be 'effect' indicators in relation to overall quality of life) (Fayers and Hand, 1997; Fayers et al., 1997).

More recent approaches to scale construction could also be adopted, including those based on item-response theory (Nunnally and Bernstein, 1994) and differential item functioning or item bias analysis (Nunnally and Bernstein, 1994; Groenvold et al. 1995; Fayers and Machin 2007,) (See *Minimal Requirements for Psychometric evaluation* below).

If the design allows for assessing the module's test-retest reliability or stability, intra-class correlation coefficients can be calculated between the two assessments. Finally, score distributions (i.e.skewness, floor and ceiling effects) of the multi-item scales and single items can be examined.

Validity

For the purpose of external validation of the module, additional information should be collected. Dependent on the QL dimensions assessed, this information could include socio-demographic data, clinical data and additional instruments assessing relevant QL dimensions. Since the module will contain items specific to certain groups of patients and/or QL dimensions, external validation criteria should be specific to the patient groups concerned (e.g., breast conserving therapy versus mastectomy to validate a body image scale included in a breast cancer module). The relevant patient groups and the corresponding comparisons should be identified before starting so that the required data can be collected and an adequate analysis plan can be set up.

In addition, the assessment of the module questionnaire at more than one point in time will permit the evaluation of its responsiveness to changes in clinical status over time.

A range of analyses is available to evaluate the validity and responsiveness of the questionnaire scales and single items. For example, known-groups comparison (Fayers and Machin 2007) can be used to evaluate the extent to which the module is able to discriminate between subgroups of patients with different disease stages, current symptoms and/or performance status. Analysis of variance can be used to test for the statistical significance of group differences.

The responsiveness of the module can be evaluated by examining differences in scores at different times during the course of the disease or treatment, for example, comparing scores before and during chemotherapy. Changes in scores over time may also be examined in relationship to changes in a criterion parameter such as performance status.

Apart from statistical significance, attention should also be paid to magnitude and precision of the constructed differences. These should be reported via the estimated score differences and their respective confidence intervals and effect sizes.

Published reports of international validation studies are listed on the QLG website (www.eortc.be/qol).

Item reduction

Since the number of patients consulted during Phases 1 (generation of QL issues) and 3 (pre-testing the provisional module questionnaire) may be relatively small, the data for informing decisions about removal of items may be limited. The module to be field-tested may therefore contain more items than is desirable. This problem may be avoided if adequate numbers are recruited in Phase 3 and Module Developers apply appropriate thresholds for inclusion of items. Nevertheless, on the basis of the data collected in the Phase 4 sample of patients, elimination of some items may be warranted on psychometric grounds.

3.5 MINIMAL REQUIREMENTS FOR PSYCHOMETRIC VALIDATION OF THE MODULE

Sample size

The sample size required will depend not only on the number of items, the number of scales and the magnitude of the correlations, but also on the heterogeneity of the sample. It is less demanding to evaluate a single scale comprised of a few items and applicable to patients with a single, clearly defined cancer site/stage/histology. It is also crucially important that the patients sampled be representative of the full range of outcomes – a large sample in which nearly all patients make more-or-less the same responses are clearly uninformative despite its size. Fayers and Machin 2007 discuss various rules-of-thumb and suggest that, for an instrument of 30 items and five or more dimensions a minimum of a few hundred patients is required. It would usually be reasonable to aim to recruit a minimum of 10 patients per item in the Module. If it is planned to use IRT in the analysis, at least 400 patients will be needed. Module Developers should obtain statistical advice before finalising their sample size.

Test-retest reliability

A module should yield repeatable scores when applied to a patient whose condition is stable. Test-retest repeatability should be formally assessed, generally by intra-class correlations. Thresholds are controversial, but for comparing groups of patients, many investigators regard correlations of at least 0.70 as “acceptable” and those that exceed 0.80 as “good”; higher standards are normally required in an instrument intended for individual-patient monitoring and management. Sample size determines the certainty of the estimates and this determines the confidence intervals. If a test-retest correlation of 0.85 is observed with a sample size of 100, the 95% confidence interval is 0.78 – 0.90, while a sample size of 150 would narrow this to 0.80 – 0.89.

Item Response Theory (IRT)

IRT may be a useful tool to apply in the selection of items for inclusion or exclusion during Phase 4. IRT is particularly suitable for reducing the number of items to be included, for example if it is desirable to produce a shorter module, or when merging two similar modules (see section 5). IRT requires substantial numbers of patients, typically at least 400 (Appendix 8).

3.6 DESCRIPTION OF MODULES IN DIFFERENT PHASES OF DEVELOPMENT

Some confusion has arisen over descriptions of modules in development; therefore the following definitions should be used. Modules in different stages of development are referred to by the phase they have **successfully completed**. Completion of Phases 2, 3 and 4 occurs when the relevant Report is approved by the Chair of the Module Development Committee.

Phase 1: The term “Phase 1 module” describes modules for which a list of QL issues is being generated. A proposal to develop the module should detail the research question and the target population and must be approved by the Module Development Committee, to ensure that there is a need for such a module and that there is no overlap with existing modules

Phase 2: A Module will be considered to have completed Phase 2 if it has completed all steps required for Phases 1 and 2 as described in the Guidelines). This includes approval by the Chair of the MDC of a Report of Phases 1 & 2 which describes the development process and records the patient derived data on which issue selection was based. A full description of the selection of issues is required. For each item a clear justification should be provided for its selection or deletion from the proposed module.

Phase 3: A module which has completed Phase 3 as described in the Guidelines and has received the formal approval of the Phase 3 Report by the Chair of the MDC is described as “*completed Phase 3*”. Such modules may be used in clinical trials with the permission of the Module Developer. Although they have been carefully developed and tested for acceptability with patients, they have not undergone psychometric testing in a large international group of patients. Therefore the suggested subscales for those modules are hypothetical and may change after psychometric analysis.

Users of “completed Phase 3” modules are advised to perform psychometric analysis of their data prior to undertaking the analysis of their main study data, for example, calculating Cronbach's alpha coefficients to ensure that the questionnaire is performing as expected.

Phase 4: (*international field-testing*): When a module has completed Phase 4 successfully and has received formal approval based on review by the Executive Committee and other peer reviewers selected by the Chair of the MDC, it is considered to be validated. A module which has completed Phase 3 and is undergoing validation testing may be described as “*in Phase 4 testing*”.

When a module has completed Phase 4, it will be made available for general use. A description of the development and validation will usually have been published either as internal reports or in peer-reviewed journals.

The term “*EORTC Module*” is reserved for modules which fulfil these criteria. Publication of the development process should include in the title reference to the phase of development being reported, for example: “*Phase 1 to 3 testing of an EORTC Module... (Specify purpose or tumour type)*”.

Researchers who develop modules to supplement the EORTC QLQ-C30 that do not meet these criteria, are not permitted to use the term “*EORTC Module*” and should explicitly state that the resulting module cannot be regarded as an official EORTC module.

Naming modules

Modules that have completed Phase 3 should be referred to in a standard way. The module name will be ‘QLQ’, followed by two or three letters that will denote the relevant tumour site (e.g., BR for breast cancer, OES for oesophageal cancer), treatment modality (e.g., RT for radiotherapy, CT for chemotherapy), or QL dimension (e.g., BI for body image and SX for sexuality) followed by 1 or 2 integers that denote the number of items included (e.g., the QLQ-BR23, the QLQ-OES24).

3.7 CO-ORDINATION OF MODULE DEVELOPMENT AND QUALITY ASSURANCE

Aims:

To ensure uniformly high quality in questionnaire modules, the entire development process is subject to monitoring, peer review and quality assurance within the EORTC QLG and QL Department. The purpose of these activities is:

- (1) *To ensure the highest scientific standards in module development*
- (2) *To avoid unintended duplication of effort*
- (3) *To avoid variation between modules;*
- (4) *To monitor the progress made;*
- (5) *To provide advice whenever needed during the course of module development and translation;*
- (6) *To evaluate the process, i.e., whether any deviations from the standard procedures were justified and whether any alternative procedures followed were sufficient to meet the standards set;*
- (7) *To evaluate the quality, suitability and compatibility of the provisional and final questionnaire modules and their translated versions.*

Monitoring:

The Module Development Committee

The members of the Module Development Committee (MDC) are the Lead Developers of modules in development and any other Group member who has participated in module development and who wishes to contribute to the MDC. The Chair of the MDC is responsible for the co-ordination of Module Development.

The Chair of the MDC coordinates module development through:

- 1) *Reviewing proposals for modules to be developed (to avoid duplication), reviewing items generated in Phase 2 (to avoid unintended variation across modules) and reviewing written documentation describing Phases 1 through 4. In these tasks the MDC Chair may invite other members of the MDC to provide written reviews; this is mandatory for the reports submitted after Phases 2, 3 and 4*
- 2) *Being available throughout the process for advice;*
- 3) *Keeping the guidelines for module development up to date by making revisions when needed.*

All Module Developers are advised to contact the Chair of the MDC at an early stage. Developers are strongly recommended to consult the Chair of the MDC if deviations from the Guidelines are anticipated or are being made during Module development.

Reviewer:

The Chair of the MDC will obtain reviews of Reports of the module development (after Phases 1&2, Phase 3 and Phase 4) from at least two members of the EORTC QLG.

The reviewers should not have been involved in the module development, although there may be circumstances in which a reviewer may have been previously consulted for advice.

It may be appropriate to consult other individuals as well, e.g., the Chair of the relevant EORTC Tumour Group.

Manuscripts for publication (the author list should conclude with “*on behalf of the EORTC Quality of Life Group*”), may be submitted for approval in place of Phase 3 and Phase 4 reports. These are also reviewed by all members of the Executive Committee.

Dependent on the nature and scope of the comments made by reviewers, the Report, manuscript for publication or the new module may need revision. If the Module Developers feel that revision is not appropriate, they should respond in writing to the Chair of the Module Development Committee to answer the issues raised by the reviewers.

Review of the translation processes and resulting translated versions is co-ordinated by the EORTC QL Department. The module developer and Translation Team Leader at the EORTC QL Department are required to review the pilot testing results of all translations to ensure the appropriateness of the translated version (see below for more information on translation issues).

Quality Assurance

The EORTC QLG is currently developing quality assurance procedures for development of new modules. Full details will be made available when procedures are established.

3.8 PREPARATION OF DOCUMENTS

The availability of detailed documentation relating to module development serves two purposes:

- 1) *To inform all interested members of the QLG*
- 2) *To provide a record of independent peer review*

Module construction is a sequential, step-wise process in which a new phase cannot be entered into unless the previous phase has been successfully completed. Permission to proceed to the next step is based on the approval of the previous steps, for which several documents need to be prepared.

All documents should be written in English.

Proposal

Before initiation of a module development project, the MDC Chair should review and approve the project on the basis of a proposal including the objectives of the planned module and the multi-disciplinary and multi-cultural involvement of contributors. This proposal should also be discussed at the subsequent meeting of the EORTC QLG to ascertain that it does not overlap with other module development projects. If there is overlap, the work needs to be co-ordinated to avoid duplication of effort.

The multi-national, cross-cultural and multidisciplinary composition of the EORTC QLG and the experience in questionnaire development accumulated by its group members enables important scientific and cultural input to the development of new modules.

Progress reports

For each bi-annual meeting of the EORTC QLG, a brief written report of module development is required, describing the progress since the last meeting, possible deviations from the guidelines and the problems that may have been encountered. This report will be reviewed by the Chair of the MDC and may be discussed at the MDC meeting, or briefly presented in the plenary meeting of the EORTC QLG. The written report will be made available to Group Members on the QLG website and a summary of the discussions will be circulated in the minutes of the EORTC QLG's meetings.

Phase 1 & 2 Report

After completion of Phase 2, a Phase 1&2 report must be submitted to the MDC for review. This will ensure that the development process has been conducted satisfactorily and that identical wording is used in newly proposed modules for those items that are similar in content. The Phase 1&2 report will contain detailed information on literature searching, qualitative interviews and the rationale for selection of the draft list of issues for presentation to patients in Phase 3 (Appendix 9).

Phase 3 Report

After completion of the first three phases (generation of QL issues, creation of a provisional item list and pre-testing) a Phase 3 report is required. This will describe the results of pre-testing and will outline the issues covered in the draft module. The report may be submitted in the form of a paper prepared for publication (with the draft module and any supporting data too detailed to be included in a publication submitted to the MDC as an Appendix), but the draft module should not be published in full at this stage.

Deviations from the Guidelines and the reasons for deviations should also be reported. Appendix 9 provides a detailed list of the topics to be included in a Phase 3 report and its Appendices.

Publications should include in the author list *“on behalf of the EORTC Quality of Life Group”*. Review of draft publications should be completed within four weeks. Developers should not submit their paper for publication until it has been approved by the MDC Chair and the Executive Committee.

Phase 4 Report

The final international field-testing (Phase 4) may be written up in a report for the MDC and/or as a paper to be submitted to a peer-reviewed journal. Phase 4 reports usually take the form of a paper for publication. The variations possible in a field study and in the evaluation of scale structure make it difficult to be prescriptive about the requirements for a Phase 4 report.

Module Developers must submit the report (draft manuscript) to the Chair of MDC and to the EORTC QL Group Executive Committee for review before submission for publication, in order to benefit from rapid constructive comments from the Group. Publications should include in the author list *“on behalf of the EORTC Quality of Life Group”*. Review of draft Phase 4 publications will be completed within two weeks. The MDC requires notification of published papers for addition to the list of QLG publications on the website.

Summary

Reports or papers need to include information about the sample (inclusion and exclusion criteria, recruitment procedures), data collection procedure and results (e.g., scale structure, internal consistency reliability, stability, clinical validity and responsiveness).

The documentation of the entire module development process for each module will accumulate during the development process and will include the following documents each submitted at the appropriate time:

- *a proposal of the planned module including its objectives and the multi-disciplinary and multi-cultural involvement of contributors;*
- *a brief written report for each bi-annual meeting of the EORTC QLG describing the progress since the last meeting;*
- *the provisional module after completion of Phase 2;*
- *two reports on the construction process Phase 1 & 2 and Phase 3, the latter may be a paper for publication);*
- *reports on the translation and pilot-testing of the module in each language separately;*
- *a report on the procedures and results of large-scale field-testing (Phase 4; usually as a scientific paper);*
- *reviewers' comments on each of the three module development reports and translation processes and the co-ordinators' replies.*

All steps of the module development process are described in a flow chart in Appendix 10.

3.9 TRANSLATION PROCEDURE DURING MODULE DEVELOPMENT - (PHASES 1 TO 3)

The questionnaire modules should undergo a rigorous translation process, based on iterative forward-backward procedures. The process is described in detail in "EORTC Quality of Life Group Translation Procedure" (Dewolf et al., 2009) available from the QL Department. Further discussion is published (Koller et al., 2007). In case of any questions or problems, developers can contact the Translation Unit at the EORTC QoL Department. (www.eortc.be/qol).

The Translation Unit is currently reviewing existing translations of Modules in order to improve consistency of translation. Some variations have arisen as a result of separate translations of the same item in different modules. This work will be made available online for consultation by Module Developers. For a new module, the translations existing in the Item Bank should be used whenever possible. In any case of doubt, the Developer should consult the Translation Unit.

The aim of translation is to produce modules which are clear, expressed in language of common use and conceptually equivalent to the original module. The English version should be used as the standard from which all other translations are prepared.

Module developers should consult the Translation Unit before starting any translation work

Phase 1 (Creating a list of issue)

For the collection of issues, developers can consult the Item Bank. If translations of issues are required, these are prepared by the developers – there is no involvement of the Translation Unit.

Phase 2 (Transforming issues into items)

For the phrasing of issues into items, developers can consult the Item Bank. There they will find not only suitable formulations of items in English, but also translations in a number of international languages.

In the case of a completely new item, developers are encouraged to provide a description of the content of the item together with the formulation of the item in English. Here is an example:

Item	Description of the content
Did food and drink taste different from usual	Has the way food and drink tastes changed? "From usual" refers to the time before you had the condition/embarked on treatment. "Taste" is the sensation of flavour perceived in the mouth on contact with a substance. Examples of taste: sweet, sour, salty, bitter.

This will help avoid ambiguities or misinterpretations and will considerably enhance the consistency of translations across different languages.

Module developers should send items plus descriptions of their content to the Translation Unit. Staff members of the Translation Unit will coordinate the translations. By default, new items or modules will be translated into the core languages Danish, Dutch, French, German, Italian, Norwegian, Spanish and Swedish. A forward-backward translation procedure will be applied as specified in the Translation Manual.

Phase 3 (testing the new module in patients; interviewing patients regarding critical items)

The preliminary module that has been generated in Phase 2 will undergo a first test in patients in Phase 3. Patients complete the questionnaire and are then interviewed in depth, to determine items that are difficult to understand, embarrassing or not necessary in the context of their health condition. Thus, Phase 3 can be regarded as the pilot test that is required as an integral part of the EORTC translation algorithm.

The entire development and translation process has to be properly documented and the documents have to be sent to the Translation Unit for review.

Translation procedure after Phase 3 and/or Phase 4 have been completed

Modules that have reached these stages have to be translated according to the guidelines described in the Translation Manual. Elements of this process include iterative forward-backward translations, pilot-testing, full documentation and coordination and review by the Translation Unit in Brussels.

Module developers should consult the Translation Unit before starting any translation work

4 • UPDATING EORTC QL EXISTING MODULES

Introduction

Questionnaire modules to supplement the EORTC QLQ-C30 are widely used to assess QL in clinical trials in oncology. Modules are developed according to EORTC Quality of Life Group guidelines. They contain scales and items addressing disease and treatment specific functional aspects of health and symptoms. In clinical oncology, there are important changes and advances being made in cancer treatment and its evaluation. The introduction of new chemotherapeutic drugs, biological agents, radiotherapy protocols and changes in surgical approach means that modules may become partially obsolete or that they may require additional items to fully cover side effects or benefits associated with new treatments.

The widespread use of a module in clinical trials and other research settings may also identify psychometrically weak items or scales in existing modules. A module may be updated to ensure that the module addresses key quality of life issues relevant to new treatments and to update scales or items in the original module with weak clinical or psychometric properties. The EORTC Quality of Life Group therefore proposes the following methodology that can be employed to update existing EORTC Quality of Life Group Modules.

The module update should start with a clearly defined research question and the target population for which the module will be updated. A list of new treatments introduced since the original module development will be produced. The actual process of updating the module consists of four phases: (1) generation of new issues related to the new treatment and identification of problematic items and/or scales in the original module, (2) creation of a revised item list by conversion of new issues into items and changing the wording of problematic items, (3) pre-testing the new module and (4) international validation field testing. It is advisable to consult the Chair of the Module Development Committee before starting this process.

Methods

Phase 1:

This phase is aimed at compiling an exhaustive list of relevant quality of life issues that covers the new treatments identified in the research question. In the process of compiling the list, three sources are used:

- (1) *Literature, including “grey literature” from the pharmaceutical industry*
- (2) *Patients*
- (3) *Health care professionals*

In addition, any problems that have arisen in the use of the original module, which may require modification, should be identified.

Literature

Two separate literature searches should be conducted to update the literature review from the original Phase 1 development. The first is designed to identify studies that report the potential QL issues associated with new treatments and to provide a list of additional new issues.

The second literature search should identify all studies that have used the EORTC module. Tables are created to summarise the studies and potential methodological problems with the module.

Tables may include information about which questionnaires are included in the study (EORTC QLQ-C30 and module, additional questionnaires), which scales and items have been reported and which scales and items show clinically significant differences. It is desirable to tabulate data concerning score distribution, validation of module or scales and any qualitative information reported. Information about missing data from particular scales and items and details of how long the questionnaires took to complete and patients' responses to them may be useful. The tables may also include information from the publications about reported problems experienced by users of the module. The update report should report which scales are used in each publication as a measure of which scales are well accepted. This information will be useful for later decisions about which scales need to be changed. Another table should be prepared containing information about the internal consistency of the separate scales of the module. All methodological problems and suggestions for new items reported in the literature should be used in the development of the updated issues list.

From the literature review, a list of new issues will emerge, containing additional issues relevant for new treatment strategies and additional issues suggested by other authors not included in the existing module

Interviews

Interviews with patients (at least 20 who were not involved in development of the first version of the module) and health care professionals (at least five, with a majority who were not involved in the first version) should be undertaken, to discuss and consider the potential new issues suggested from the literature and to discuss potential changes to existing scales and items. HCPs and patients will receive a list of issues, combining the two issue lists, the "old" module and the "new" issues.

Patients may be interviewed before, during or after treatment. Investigators should ensure that patients receiving new treatment strategies are well represented in the patient sample. HCPs may be any professional involved in the treatment of relevant patients and with specialist knowledge of the treatments and condition.

The list of new or modified quality of life issues may be administered in the form of a semi-structured personal interview in which the following questions are asked: (i) *Are the issues included relevant to the new treatment(s)?* (ii) *are the proposed changes appropriate?* and (iii) *are there issues missing from this list that are considered relevant to the new treatment?*

It may be relevant to explore with patients the reasons why some items may have caused difficulty in previous studies.

Phase 2:

The new issues are converted into items (and any necessary revisions to existing items) as described for Phase 2 of development. All wording and layouts should conform with recommendations in the Guidelines for Module Development and Translation. Consultation of the EORTC Item Bank (www.eortc.be/itembank2) is recommended to prevent duplication.

Phases 3 and 4:

The pre-testing and validation of the updated module will follow standard guidance as for new module development. In addition it may be appropriate to compare compliance and acceptability of the new version with the previous version.

5 • MERGING TWO MODULES

Background

The QLQ modules provide organ specific assessment tools for QL in a wide range of tumour sites. These are specifically developed for each tumour site. Occasionally it may be appropriate to combine two existing modules, where two organs or conditions with existing modules are very close anatomically or physiologically.

Examples of combination of modules are the oesophagogastric module QLQ-OG25, (from a combination of the oesophagus and stomach modules, QLQ-OES18 and QLQ-STO 22) and the cholangio - carcinoma module derived from the pancreas and hepatocellular carcinoma modules QLQ-PAN26 and QLQ-HCC18). In each case it is important to consider whether it makes clinical sense to try to combine modules or whether to start a completely new module. Factors to consider include the degree of overlap of symptoms of the two tumour sites and the extent of similarity of progression of the diseases and the treatments offered.

Combining modules is not necessarily easier than starting from scratch but may have the advantage of using questions that have been tested and studied using psychometrics as part of a Phase 3 or Phase 4 study.

Methods

Approval of the MDC should be obtained before starting work. It is necessary to demonstrate the need for a combined module and that it is appropriate for the two modules concerned.

Phase 1:

A literature search should be performed using all relevant terms relating to the new diagnosis/organ and all issues arising should be listed as described for Phase 1 development.

The two existing modules should be reviewed and all the questions combined into a single set of logical, clinically sensible groupings (probably, but not necessarily corresponding to scales of the existing questionnaires). Some existing scales will be combined in this process. In addition, any new issues arising from the literature search should be included in these groupings of items. This may result in item groupings with a combination of “issues” and “questions” which may be difficult to work with. Some patients and health care professionals may be confused by the variation between items (questions) and issues (described features of QL) and may prefer one or other format which could bias responses. Therefore, the issues should be converted into questions/items at this stage, if possible using items from the Item Bank that have been used previously in validated modules and as such will have been translated.

This item list should be evaluated by patients and health care professionals as described for Phase 1 module development

Phase 2:

Because the majority of items are derived from existing questionnaires and additional issues are already framed as items, Phase 2 is relatively straightforward. Decision rules for inclusion and exclusion of items should be agreed and a final item list derived.

After removing unwanted items, the original item groupings may be used as “hypothesised scales” or they may be rearranged into clinically meaningful scales with additional individual items, if necessary.

At this point the Module Developers should decide whether a new module is needed at all, or whether using one of the original modules would suffice. It is recommended that this decision be discussed with the Chair of the Module Development Committee, who will require submission of a report of Phases 1 & 2 for review before the work can progress to Phase 3.

Phase 3:

The provisional item list should then be tested in a further sample of patients of all relevant stages of the disease and from different countries. Standard psychometric tests may be applied to the results to check correlation of questions and internal validity of the Questionnaire.

Phase 4:

A field study should be carried out.

6 • PUBLICATION OF MODULES

Modules that have completed Phase 2 or Phase 3 may not be published. Descriptions of the module development may be published, including a description of the issues contained in the module but these publications should not contain the text of the questionnaire.

No restrictions are made with respect to the publication of the text of Phase 4 modules. However all publication of modules should carry the EORTC logo and copyright must be asserted. Publications describing the development process of Phase 3 or 4 module development should include in the authors list “*on behalf of the EORTC QLG*” and should be approved by the Executive Committee of the QLG before submission for publication.

When researchers other than the Module Developers use Phase 3 modules, the following rules for publication of the research apply:

- 1) The module itself may not be published other than by its developers;
- 2) The Module Developers should, in principle, have the right to publish their data first - However, if this is not possible, publications should be negotiated on a case by case basis;
- 3) Collaboration between the principal investigator(s) of the module and its users are required with respect to the scoring and scale structure of the module;
- 4) At least one developer of the module should be a co-author on publications that includes information on the psychometric performance of the module;
- 5) The module developers and other researchers should agree in advance on the required access of the Module Developer to the data derived from the module and such Socio-demographic/clinical data as would be necessary for the purpose of psychometric/clinical validation.

7 • OWNERSHIP AND USE OF MODULES

Ownership

The modules developed under the auspices of the EORTC Quality of Life Group are the property of the Group. Users' agreement and copyright procedures will follow those drawn up for the core questionnaire.

Module developers retain copyright of their module until it has completed Phase 4 validation, at which time copyright reverts to the EORTC Quality of Life Group.

Using a module in Research

Modules that have completed Phase 1 and 2 are not suitable for primary research and should not be used. Information about the development process may be published; unpublished material can be obtained directly from the Module Developers.

Modules that have completed Phase 3 are not freely available, but may be obtained from the principal investigators. Copyright of these modules remains with the Module Developers.

If researchers want to use Phase 3 modules, they may do so if:

- 1) They have received the explicit permission of the Module Developer;
2. They leave the module's integrity intact and do not revise items. However, if they want to add items at the end of the module they may do so after consulting the Module Developer;
2. They must provide the Module Developer with a copy of the module as used in the study and the study protocol. When the study is finished they should report any comments on performance of the module to the Module Developer;
3. They agree, if requested, to contribute data for purposes of the psychometric/clinical validation of the module;
- 4) They use the hypothesized scale structure as agreed by the Module Developer;
- 5) They respect the publication rights, rules and regulations;

Validated modules that have completed Phase 4 are the property of the EORTC QLG and can be downloaded from the EORTC QL Department website www.eortc.be/qol after a user's agreement has been signed.

8 • APPENDICES

8.1: MATRIX FOR ASSISTING IN TARGETING PATIENT RECRUITMENT

The groups used to construct a sample matrix for different modules may vary and should be decided by the Module Developers in advance of each Phase. Usually, the matrix will include selected groups from two of three categories: disease stage, treatment type, or stage of treatment. Module Developers may decide to combine cells and to avoid recruitment in some cells as appropriate, relevant to the tumour type or condition being assessed. Two examples are shown below.

In Phase 1, recruitment should be spread evenly across the cells chosen for inclusion of patients, to ensure a representative sample of patients. In Phase 3, each designated cell should contain 15 patients.

Example 1

	Surgery	Chemo-therapy	Radio-therapy	Palliative Care
Localised Disease	X	X	X	-
Advanced Disease	X	X	X	-
Palliative Care	-	-	-	X

Example 2

	Pre Treatment	Mid Treatment	Post Treatment	Palliative Care
Localised Disease	-	X	X	-
Advanced Disease	X	X	X	-
Palliative Care	-	-	-	X

8.2: GENERATION OF RELEVANT QL ISSUES IN PHASE 1: - EXAMPLE OF A PATIENT INTERVIEW

Introduction

There are two main approaches to gathering information about Health Related Quality of Life issues from patients with a particular condition, or who are undergoing a particular treatment. The researcher may ask the patient to describe their experience and allow the patient to provide information freely, or in response to predetermined questions in a semi-structured interview. Alternatively, the researcher may show the patient existing relevant material to begin the discussion and to prompt the patient's description of significant issues. The two approaches may be used sequentially.

All information voiced by the patient should be recorded, preferably on audiotape or a digital audio file and then transcribed for later analysis. This method ensures accuracy of wording as used by patients and reduces any bias that may result from the selective noting of patient comments by the researcher

Interview

The researcher should begin the interview with some introductory remarks to explain its nature and purpose. For example:

We are asking for your help in devising a questionnaire which will be used to monitor the experiences of patients who have (specific disease or treatment.) I would like to ask you a few things about your health. Can you tell me about the experiences you may have had as a result of your disease (or treatment).

Neutral probes should be used to elicit more information. e.g. 'Can you tell me more about that?' or 'Can you think of any additional experiences?' The semi-structured interview designed by the Module Developers may explore specific areas of concern with each patient.

The EORTC QLQ-C30 or any other relevant list of items or issues may be shown to the patient after the patient has provided those issues which arise spontaneously. The EORTC QLQ-C30 and other material may serve as a prompt to stimulate further suggestions. (Place the EORTC QLQ-C30 (and any relevant list of items or issues) before the patient) and continue as follows:

Here you see a list of experiences related to (condition, treatment or additional QoL dimension) which a patient who is (relevant characteristics) may have. Please could you indicate for each experience separately the extent to which you have had it during your illness.

This is an example which could be used to determine the relevance of an issue in a more complex setting e.g. during the development of a module for spiritual wellbeing:

This is a list of thoughts and/or feelings which patients with cancer may experience. Could you please go through the list and, for each one, tell me how much it has been something which you have felt or thought about.

Identify new issues

In order to identify new issues, the interviewer should explain to the patient what is required. The patient should read the EORTC QLQ -C30 and then suggest any additional issues relevant to their disease and their QoL.

This is an existing questionnaire that asks about you and your quality of life. These questions may be of value for all patients who have cancer. Could you please read these questions? You may have had some other experiences that are not included in this questionnaire.

I would like to ask you a few things about your health:

Can you tell me about the experiences you may have had as a result of your disease (or treatment)? (Interviewer may use additional neutral probes, e.g.: Can you tell me more about that? Can you think of additional experiences?)

- *Can you think of anything else that you have had (experienced/had to cope with) during your illness that is not included in this questionnaire?*
- *If yes: please name each of these experiences so I can write them down*
- *For each additional issue: could you tell me about this?*

The issues raised by the patient interviews will be transcribed and tabulated and combined with those generated by the literature review and health professional interviews.

Relative importance of issues

When using an open structured interview, there will not be a readily available list of issues to review. The interviewer should summarise the issues raised during the interview and ask the patient which issues are most important:

- *We would like to ask you which of these issues troubled you the most*
- *For each chosen issue separately: Can you tell me about that?*

Before asking patients to discuss relative importance it may be necessary to categorise issues into two lists, one for 'problems' e.g. cough, shortness of breath and one for functioning issues where the concern is capacity or ability or sometimes even positive experiences. e.g. ability to do work or other daily activities.

- *I would like to ask you which of these problems, including the problems you have mentioned yourself troubled you most? Please look again at these lists and pick out five (to 15) problems that caused you the greatest trouble (nuisance, distress)*
OR
- *We would like to ask you which of these abilities including any you have mentioned yourself, you value most highly? Please look again at these lists and pick out five (to 15) issues which are particularly important to you.*

If the total number of issues is small (15-20); it may be sufficient to ask patients to identify five key issues; however for larger number of issues it may be necessary to ask patients to identify 10 - 15 issues.

Follow-up questions

The use of follow-up questions or "probes" will be required in the majority of interviews. The appropriate wording is dependent on the topic at hand, but should always be in an open, non-judgemental way. For example:

- If the answer is too general and indefinite, the follow-up may be
 - *"In what way?" "Just how do you mean?" "Can you give me an example?"*
- If the answer is incomplete, the questions may be:
 - *"Any other reasons?" "Would you tell me a little more about that?"*
- Other follow-ups could ask: *"What makes you think this?" "What was there in the question that made you feel that way?"*

It may be useful to prompt the patient to consider specific domains, especially if the literature review has suggested that these may be relevant to the patient group. Some examples are shown:

- *Do you have any other symptoms not mentioned in the questionnaire?*
- *Do you have other problems with your physical functioning/health/changes in sleep patterns?*
- *What are you not able to do that you would formerly do before your illness, any why?*
- *Are you limited in normal daily activities (e.g.-shopping) or self care (e.g.-washing/bathing) compared to before your illness? What is it that limits you?*
- *Are you undertaking fewer social activities (e.g..-hobbies, meeting up with friends) and why?*
- *Have changes in relationships with family/friends occurred?*
- *Do you have financial problems or worries due to your illness?*
- *Have your personal feelings changed (e.g.-satisfaction with life, spirituality)?*
- *Has your emotional wellbeing changed (e.g.-feelings of anxiety or worrying)?*
- *Are there any other issues or comments you would like to make regarding your illness and treatment and your quality of life?*

Generation of list of issues

The list of QoL issues raised by patients in Phase 1 will be reviewed, together with the responses of the health care professionals. In principle, all issues should be considered for inclusion in the provisional item list, because the generation of issues is based on responses of a relatively small number of patients and all information should be used at this stage. There will be an opportunity in Phase 3 to exclude items that have low importance or relevance, for example.

Exclusion of issues

Decisions to exclude issues raised during patient interviews should be based on the following features:

- *Redundancy*, either because of overlap with the core questionnaire or because of the generation of multiple closely related issues
- *Upsetting*, issues which are potentially *distressing* (e.g. "anxiety about approaching death") may be excluded, if no acceptable alternative wording can be found (e.g. "concerns about approaching the end of life")
- *Lack of relevance*, if an issue is raised by only one patient and is scored very low for relevance by the health care professionals it may be excluded



Inclusion of new issues

If at least two patients mention an additional issue, it should be included at this stage in the list of issues to be considered in Phase 2, provided that the motivation is plausible. In some cases, an issue mentioned by only one patient may warrant inclusion.

8.3: GENERATION OF RELEVANT QL ISSUES IN PHASE 1: - EXAMPLE OF A HEALTH CARE PROFESSIONAL INTERVIEW

Introduction

The interviews with health care professionals complement the patient interviews. They may be able to identify important but uncommon issues that may not be found in the relatively small number of Phase 1 patient interviews. It is most useful to conduct the health professional interviews after generation of relatively complete lists of issues, that is after the literature review and after completion of (most of) the patient interviews.

Relevance

The following is an example which could be used to determine the relevance of an issue in a simple situation – e.g. Health professionals' views of lung cancer patients' experiences of a cough, chest pains or tingling fingers.

We already have a questionnaire assessing quality of life aspects of cancer patients in general. Quality of life aspects relevant to specific diagnostic patient groups are not included in this questionnaire. We are asking your help in devising a questionnaire which will be used to assess the quality of life of patients who have (specific disease or treatment).

Place list with issues before the health professional:-

*Here you can see a list with issues relevant to cancer patients with (specific disease or treatment).
Could you please indicate for each issue separately the extent to which you find it relevant for this patient group.*

Response categories could range from (1) not relevant to (4) very relevant.

“Relevance” refers to the frequency with which a specific complaint occurs and if it “occurs”, the trouble it may cause. Thus the more frequently a complaint occurs and the more trouble it causes, the more relevant it will be for this patient group.

After completion the interviewer asks:-

Could you please tell me for each issue for which you circled 1 (not relevant) or 2 (a little relevant) why you consider it not or only a little relevant?

Interviewer notes down the reasons.

Here is an example which could be used to determine the relevance of an issue in a more complex setting.

We would like your help in developing a questionnaire which will be used to monitor the experiences of patients with cancer. This is a list of thoughts and/or feelings which patients with cancer may experience. Could you please go through the list and, for each one, tell me whether you think it is something your patients have ever considered. (see also the data collection form on pX)

Relative Importance

It may be necessary to select some issues and omit others from the provisional item list, especially if the number of issues raised is large. To assist selection, the health care professionals should be asked to rank the issues in order of importance, or to pick out the most impor-

tant issues that should definitely be included. The Module Developers should ask the subject to identify a number of issues between 5 and 10.

The list of issues (including any new issues which you have identified) is too long to be administered to patients. Therefore a subset of issues must be chosen. Please could you mark those items that, in your opinion, affect the quality of life of these patients most profoundly and that we should definitely include in the final questionnaire. You may choose a limited number of issues (specify the exact number) that you consider to be most relevant and that you think should definitely be included. If there are items that you think should definitely be excluded please mark these also and say why you think they are not a priority.

Breadth of coverage

To assess whether the list of issues covers all aspects of QL in the target patient group (including all possible subgroups of disease or treatment), the researcher should explore the breadth of the list of issues.

(Place the EORTC QLQ-C30 before the health care professional).

This is the existing questionnaire that assesses the quality of life of cancer patients in general. Could you please read these questions? You may have thought of other things that are not included in this questionnaire nor in the previous list of issues you have just rated.

*Please consider patients at all stages of disease and patients undergoing any type of treatment for this condition. Can you think of anything else that may be of relevance to this patient group and is not included in these two questionnaires?
If yes: Please name each of these issues so I can write them down.
For each additional issue: Could you tell me about this?*

Review of list of issues

After the interviews with patients and health care professionals have been analysed there may be issues in the provisional list that were not considered in the interviews with health care professionals. The full list of issues may be shown to the health professionals at this stage; with a response scale for each issue to record the professional's rating of relevance (from 1 to 4).

8.4: DECISION RULES FOR SELECTION OF QL ISSUES IN PHASE 1

In principle, if one or more patient or health care provider mentions an issue, it should be included, provided that the rationale is plausible.

At this stage, one should feel reluctant to exclude issues. However, if the number of patients interviewed is large (>30) and the list of issues has been scored by patients or by health care providers, issues that have a low (e.g. mean < 2) mean score for relevance or importance may be considered for exclusion.

The Module Developers should review each issue in the context of the proposed scale structure (i.e. each scale considered in turn as a group of issues). It is necessary to consider the meaning of each issue, whether there is overlap or redundancy within the proposed new issues and whether the issue is already assessed by the QLQ-C30. Some issues must be handled sensitively when creating a questionnaire. For example, issues about approaching death are clearly important to some patients, but may cause distress to others. Alternative phrasing (refer to “approaching the end of life”) may be more acceptable.

If a larger numbers of patients (>30) have contributed to the list of issues, the threshold for inclusion of an issue into the new module should be that it was mentioned by more than 5%.of patients.

In some circumstances a comparative approach is needed. For example, in selecting issues for inclusion in the QLQ-ELD15 (module for elderly patients), the percentage prevalence of each issue was determined in both the >70 years and 50-69 years control group to determine if it was a general concern of all cancer patients, or if it specifically applied to older cancer patients. Issues that were cited by at least 1.5x older patients than younger (a ratio of 3:2) were considered for inclusion in the new questionnaire.

All decisions about inclusion should be reviewed by all the Module Developers to ensure consensus in the inclusion or exclusion of issues. The Module Developers may agree to vary the criteria for particular issues, if there is a strong argument for doing so. This should be recorded in the Phase 1&2 report.

8.5: EXAMPLE OF A PATIENT INTERVIEW IN PHASE 3

Introduction

Pretesting is designed to collect response data, to record evaluation of relevance and importance and to record the subjective impression of the patients after they have completed the EORTC QLQ-C30 and the new provisional module/item list.

For short provisional modules/item lists, the interview should examine each item individually. For longer lists, the interview should ask the patient to identify particular aspects of the whole questionnaire and discuss these in detail.

Administration of EORTC QLQ-C30 and the module

The patient is asked to complete the EORTC QLQ-C30 and the new provisional module/item list.

We have two questionnaires that ask about you and your health and quality of life. I will ask you first to complete these questionnaires. After you have completed them, I will interview you to make sure we asked the right questions in the right way. We want to be sure that we cover the most important aspects of patients' experience of (disease /treatment /characteristics).

(Place EORTC QLQ-C30 before the patient who then completes it)

As a result of your (illness/treatment) you may have experiences in common with other patients who have the same problem. These particular experiences are not covered by this more general questionnaire. We would like to add some extra questions to take account of those things which may be important to you and other patients who have (disease /treatment /characteristics). We are now asking your help in devising these additional questions.

We think that this questionnaire may be more useful for patients who have (specific disease or treatment, or additional QL dimension).

(Place the provisional module before the patient who completes it)

Interview directed to each item separately

The wording of the interview questions will be dependent on whether the module item refers to a problem or ability and how the respondent has completed the particular item (i.e., no problem at all versus a problem to some degree).

For items referring to problems the patient has experienced, ask the following questions:

I see that you have (particular problem) to some degree.

- *Is this correct?*
- *Can you tell me about this problem?*
- *Do you think that this problem is related to (disease or treatment)?*
- *Did you have difficulty in replying to this question?*
- *Did you find this question annoying?*
- *Did you find this question confusing?*
- *Did you find this question upsetting?*
- *How would you have asked this question?*

For problems the respondent did not endorse, ask the following questions:

I see that you did not have this problem during the previous week.

a. Is this correct?

b. Have you ever experienced this problem before last week?

If not, go to question (e)

c. If yes, do you think that had something to do with your disease (or treatment)?

If not, go to question (e)

d. If yes, can you tell me about this problem?

e. Did you have difficulty in replying to this question?

f. Did you find this question annoying?

- g. Did you find this question confusing?*
- h. Did you find this question upsetting?*
- i. How would you have asked this question?*

For abilities and functioning which the respondent indicated to be limited to some extent, ask the following questions:

I see that you were able to (fill in ability) to some degree during the previous week.

- *Is this correct?*
- *Can you tell me about this (ability)?*
- *Do you think that your disease (or treatment) has affected in any sense your ability to (ability)?*
- *Did you have difficulty in replying to this question?*
- *Did you find this question annoying?*
- *Did you find this question confusing?*
- *Did you find this question upsetting?*
- *How would you have asked this question?*

For abilities and functions the respondent is completely able to perform, ask the following questions:

I see that you were able to (fill in ability) during the previous week.

- a. Is this correct?*
- b. Can you tell me about this ability?*
- c. Were you limited in your capacity to do this before last week?*
If not, go to question (e)
- d. If yes, do you think it had something to do with your disease (or treatment)?*
- e. Did you have difficulty in replying to this question?*
- f. Did you find this question annoying?*
- g. Did you find this question confusing?*
- h. Did you find this question upsetting?*
- i. How would you have asked this question?*

Interview directed to the entire module

If modules contain a large number of items (e.g., over 20), the time involved in questioning about each individual item would be prohibitive. In those cases the questions may be directed towards the entire module. For example:

- *Were there questions that you found difficult to answer?*
- *Were there questions that you found annoying?*
- *Were there questions that you found confusing?*
- *Were there questions that you found upsetting?*
- *Were there questions that you found intrusive?*
- *Do you have other comments about these questions?*

These general questions may then be supplemented by the further probing of selected module items, for example, questions that are expected to cause some difficulty and items that appear to be troublesome during the interview.

Completion of the interview

The pre-testing interview should be completed with two questions directed to the entire questionnaire that is the combination of the core questionnaire and the module:

- *Were there questions that you found irrelevant?*
- *Can you think of additional issues that are relevant for you but are not included in this questionnaire?*

Thank the patient for their contribution to the research.

8.6: DECISION RULES FOR INCLUSION OR EXCLUSION OF ITEMS IN PHASE 3

In Phase 3 it is necessary to reduce the (usually) long provisional list of items to a shorter (preferably no more than 20) list of items for the new module. In development of a module, the viewpoint of the patient should be given the greatest weight in the selection of items. At this stage in Phase 3, some selection must be applied to remove unnecessary items, balanced against the need to produce a module that adequately covers all the QoL concerns of the target patient group. Module Developers should agree decision rules for this selection before beginning the analysis, although the rules may be modified if preliminary inspection shows that they would lead to exclusion of too many or too few items.

Comments provided by patients are also important and should be taken into consideration.

Decision rules

The relevance and importance ratings provided by patients should be considered before review of the other responses and items which fail to score adequately should be excluded. A suitable cut-off should be agreed, for example if relevance is scored yes /no, an item could be retained if at least 60% respond yes; if importance was scored on a 4 point scale, the item could be retained if >60% scored 3 or 4 (quite a bit-very much). Problems (e.g., symptoms) that relatively few patients describe and abilities that relatively few patients were limited in, may be of little relevance for inclusion in the final module. These are candidate items for deletion.

Parameters of each item to be considered include the mean score and the number of patients reporting the item (score 2, 3 or 4) divided by the total number that completed the item (prevalence ratio). A full range of responses is important: Items that have limited variance should be excluded. In particular, "floor" and "ceiling" effects should be looked for in the distribution of responses to each item.

Negative items (e.g. symptoms) score more highly (3 or 4) if the symptom is greater, whereas positive items (e.g. functions) score highly if disability is less. For the purposes of these decision rules, responses can be standardised by inverting responses to the positive items to correspond with response categories ranging from 1 "no disability" to 4 "very much disability". The following cut-off points are suggested for selection of items for retention in the final module (after consideration of importance and relevance as noted above):

1. Mean score > 1.5
2. Prevalence ratio >30% **or** prevalence of scores 3 or 4 >50%
3. Range > 2 points
4. No floor or ceiling effect: responses in categories 3&4 or 1&2 >10%
5. No significant concerns expressed by patients (e.g. item is upsetting, ambiguous)
6. Consistency across languages/cultures.
7. Compliance: at least 95% response to the item

Module Developers may vary these criteria on a case by case basis. The cut-off points may be adapted depending on the number of items pre-tested, the number of items identified as having a high priority and the sample size. Any variation should be explained in the Module Development Report, or in an Appendix to the publication.

Items that meet at least five of these seven criteria may be retained in the list, unless the answers to the open interview questions suggest that this is inappropriate (e.g., for the majority of subjects, the issues are not related to the disease, or the question meant something different).

Items that meet four or fewer criteria should be excluded, unless the interviews provided strong arguments for retaining them in the list (e.g., when the importance was stressed in a considerable number of interviews). Items that meet four of these seven criteria may be retained in the list, if discussion by the Module Developers concludes that inclusion in the final module is appropriate.

Addition of new items

Additional issues (not included in the provisional item list) may arise during Phase 3. However, some uncommon issues may arise and may be felt sufficiently important to warrant consideration for inclusion. Investigators should retain a high threshold for the addition of new items. Such items may introduce new problems at a later stage in development, they have not been validated in the Phase 3 testing and they may represent rare or idiosyncratic concerns (see also Converse and Presser, 1986).

Additional issues that are mentioned by a considerable number of and that are related to the disease or treatment should be developed into questions and added to the list. Investigators may wish to agree a defined proportion of patients that report a missing issue before it could be added to the module at this stage. However, any quantitative number, such as at least a third of the patients, is arbitrary. Researchers must apply their judgement to balance potential loss of information versus inclusion of untested items. A justification for either choice should be documented in the report.

Rephrasing items

On the basis of the interviews, questions may be identified that troubled (some of) the patients. This information should be taken seriously. Even when a small number of patients had difficulty answering the questions, these should be rewritten, as others may have had some reservations or difficulties but chosen not to discuss them, such items should be rephrased, subdivided or substantially changed as appropriate.

8.7: DEBRIEFING QUESTIONNAIRE IN PHASE 4

Patient Study ID

--	--	--	--	--	--

Date of Interview:

--	--	--	--	--	--	--	--

ABCXX DEBRIEFING QUESTIONNAIRE

1. How long did it take you to complete the questionnaire?

--

 minutes

2. Did anyone help you to complete the questionnaire?

No Yes If so:

a) *What kind of help?*

b) *How much help was provided?*

3. Were there questions that you found confusing or difficult to answer?

No Yes

If so, which ones?

4. Were there questions that you found upsetting?

No Yes

If so, which ones?

5. Please use the space below if you have other comments about the questionnaire.

--

Thank you!

8.8: ITEM RESPONSE THEORY (IRT) FOR SCALE STRUCTURE & SELECTION OF ITEMS IN PHASE 4

For QLQ module development, IRT can be used as a psychometric development tool. If an early version of a questionnaire contains a lot of candidate items that are all believed to be measuring much the same thing, IRT provides an excellent means for identifying the most informative one or two items and for quantifying how much extra information or precision would be gained from increasing the scale length by including additional items. The publications describing the QLQ-C15-PAL provide examples of shortening some scales of the QLQ-C30 (Petersen et al., 2006; Groenvold et al., 2006; Bjorner et al., 2004).

Whereas traditional psychometrics explores averages (means), standard deviations and correlations of the responses to questions, IRT is concerned with the *probability* that any particular patient will select one or another response option. Factor analysis and IRT are concerned with “latent variables” that are not directly measurable, but which it is assumed that the scale-score represents.

IRT has become increasingly widely used in questionnaire development and – when applicable – possesses some major advantages over traditional methods. IRT is primarily useful when there are a number of items that all address a single homogeneous dimension. Like factor analysis, it is of no value for single items (although it may aid in selecting a single item from among a group of similar items). Unlike factor analysis, it is not suitable for multi-item scales that lack homogeneity – as might be the case if several items are deliberately chosen to extend the breadth of coverage of a concept (i.e., multi-item scales characterised by a low Cronbach’s α).

IRT can be used (a) solely as an aid to developing a scale that is then to be scored using traditional methods (such as summation, as commonly used for most HRQL scales), or (b) to develop an IRT-based scale that is also scored using IRT computer software. Examples of (b) would be primarily but not necessarily scales such as physical functioning, where a number of items might be chosen to target patients with varying levels of ability (*Can you get out of bed? ... Can you run a marathon?*) For such scales, IRT can provide a consistent scoring system.

To apply IRT, we would typically require data on at least 400 patients and the sample should contain a fairly even spread of patients across the continuum of interest (i.e., it is unhelpful and uninformative to have a lot of patients responding “no problems”). A simple introduction to IRT may be found in Fayers & Machin (2007), while a more detailed exposition is provided by Embretson and Reise (2000).

8.9: TEMPLATE FOR REPORT ON MODULE CONSTRUCTION (PHASE 1 TO 3)

The Reports, written in English, should be organised according to the following sub-headings and should include the information listed under each heading:

In the report, lists of issues, items, interview data and instructions or interview structures may conveniently be presented as 'Appendices'.

Module Developers may choose to present material in Appendices in some, all or none of the places indicated below.

Phase 1 & 2 Report

No Report is required after Phase 1 alone.

A comprehensive **Report of Phases 1 & 2** describing procedures, findings, reasoning and the provisional item list must be approved before progress to Phase 3.

1. Research objective

Justification of need for module, description of the unmet need for assessment.

Purpose or patient population for which the module was developed.

Clear statement of the Research objective.

2. Phase 1: Generation of quality of life issues

Literature search

Search headings and databases used.

List of references included in the literature search; tabulation of main findings (*Appendix*).

List of available questionnaires consulted (*Appendix* to reproduce these).

List of QoL issues from literature (*Appendix*); if this is relatively short, it may be presented as a Table in the report.

Interviews with patients

Patients

Number of patients and relevant background characteristics

Interviews:

The interview instructions (*Appendix*).

Results of patient interviews:

Quantitative results - Individual ratings, average ratings, priority ratings (*Appendix*).

Qualitative results - Comments leading to adaptations (e.g., irrelevance of issues, rewording, combining or splitting up of issues, omissions).

List of issues arising from patient interviews (*Appendix*).

Interviews with health care professionals

Health-care professionals

Number of health care professionals and their specialities

Organisation of the issues

A brief description of the categories of issues shown to or discussed with the health care professionals for example, disease symptoms, treatment-related side effects etc.

Interviews

The interview instructions (*Appendix*)

Results of health care professionals interviews

Quantitative results - Individual ratings, average ratings, priority ratings (*Appendix*)

Qualitative results - Comments leading to adaptations (e.g., irrelevance of issues, Re-wording, combining or splitting up of issues, omissions).

List of QL issues from the professionals interviews (*Appendix*).

If this is relatively short, it may be more convenient to present as a Table in the report.

Phase 2: Creation of a provisional item list

Description of all steps in the conversion of the list of issues into a provisional module/item list.
Review of issues; removal of duplications within the item list and with the EORTC QLQ- C30 (and other modules if appropriate).

Inclusion of existing items from the Item Bank.

Construction of new items (not in the Item Bank).

List of new items not included from the Item Bank.

The resulting questionnaire or provisional item list

Description of the provisional item list and conceptual groups of items (*Appendix*)

Phase 3 Report

Module Developers may submit for approval to the Chair of the MDC either a Phase 3 report or a paper intended for publication that reports the Phase 1 to 3 developments. If a paper is submitted, it may not contain all the data on which the conclusions were reached, as set out in the Template below. In that case, the Module Developers should submit to the Chair of the MDC their paper for publication and additional files containing all other material required below, with a covering letter to list the additional documents.

It is unnecessary to repeat information already contained in the Phase 1 and 2 reports.

The **Phase 3 report** should describe:-

Patients

Number of patients and their relevant characteristics

Interviews

The interview instructions (*Appendix*)

Procedure for item selection

This should include a clear description of the agreed decision rules applied to selection of items for the final module.

Results

Quantitative (may be tabulated in the report, or presented as an *Appendix*)

Qualitative – usually described in the report

Resulting module to be field-tested

General Description

Hypothesised scales

Single items

The module should be presented in final form as an *Appendix* (not for publication).

Translation

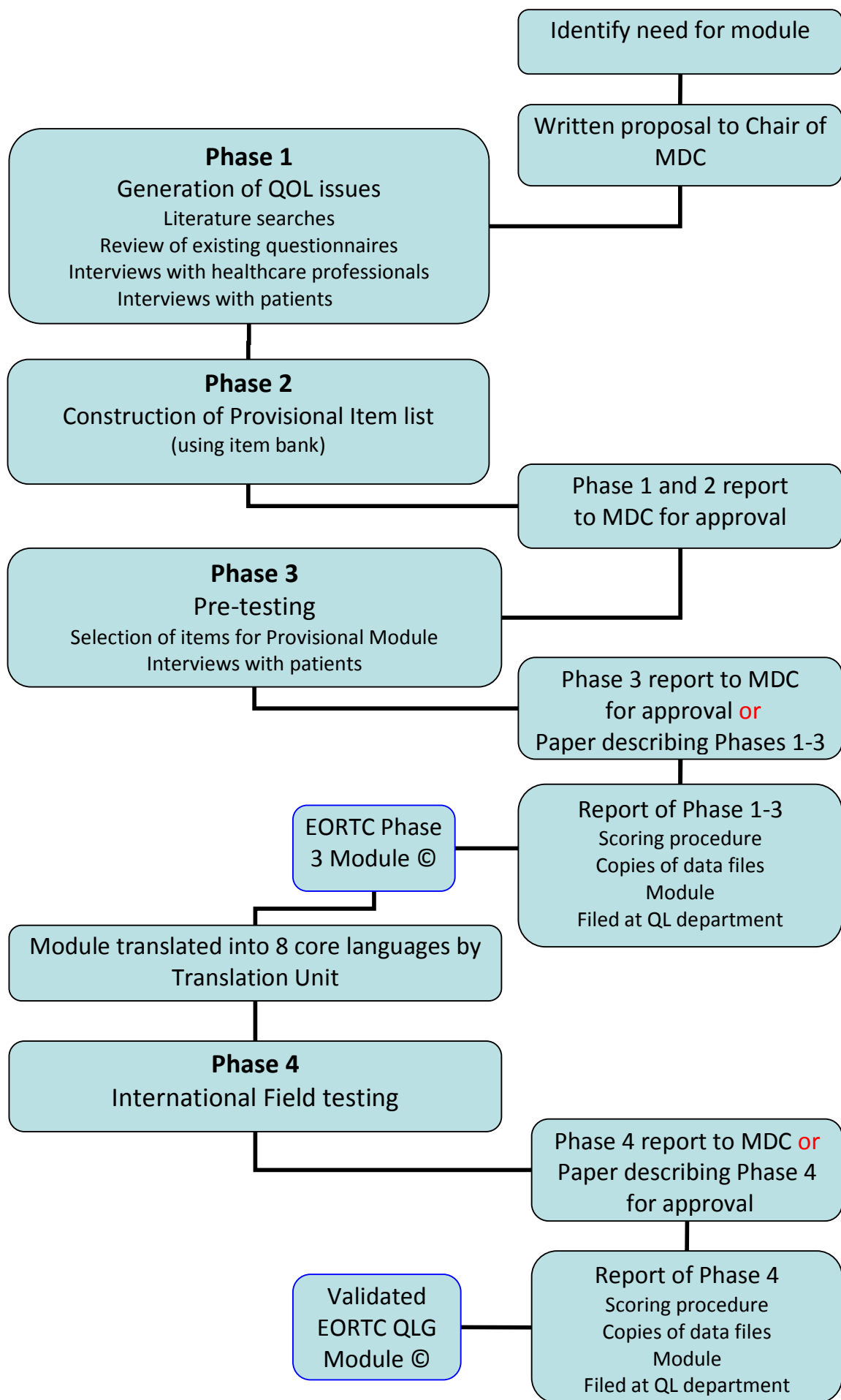
Brief description of the translation of issue and item lists used during Phase 1 and Phase 3.

A full translation report must also be filed with the Quality of Life Department.

Note on researchers involved

The Module Developers may wish to describe the contributions of the various members of the Module Development team, for example the recruitment of patients, the analysis of qualitative

APPENDIX 9 – FLOW CHART OF MODULE DEVELOPMENT PROCESS



9 • REFERENCES

- Bjorner JB, Petersen MA, Groenvold M, Aaronson N, Ahlner-Elmqvist M, Arraras JI, Brédart A, Fayers P, Jordhøy M, Sprangers M, Watson M, Young T; European Organisation for Research and Treatment of Cancer Quality of Life Group - **“Use of item response theory to develop a shortened version of the EORTC QLQ–C30 emotional functioning scale”** - *Quality of Life Research*, 13(10): 1683–1697. (Dec 2004)
- Brod, M, Tesler LE, Christensen TL – **“Qualitative research and content validity: developing best practices based on science and experience.”** *Quality of Life Research*, 18(9): 1263-1278. (Sept 2009)
- Converse JM, Presser S - **“Survey Questions. Handcrafting the standard questionnaire”**. *Sage University Paper (63): Quantitative Applications in the Social Sciences*. Beverly Hills, CA: Sage. (1986)
- Dewolf L, Koller M, Velikova G, Johnson CD, Scott N, Bottomley A - on behalf of the EORTC Quality of Life Group - **“EORTC Quality Of Life Group Translation Procedure”**. (3rd edition). *EORTC, Brussels*. (2009)
- Embretson SE, Reise SP. **“Item Response Theory for Psychologists”**. *Laurence Erlbaum Associates, Mahwah*, (2000) ISBN-13: 978-0-805-82819-1
- Fayers PM, Hand DJ - **Factor analysis, causal indicators and quality of life.** *Quality of Life Research*, 6(2):139-150. (March 1997)
- Fayers PM, Machin D. **“Quality of Life: the Assessment, Analysis and Interpretation of Patient-reported Outcomes”** (2nd edition). *John Wiley & Sons, Chichester*. ISBN-13: 978-0-470-02452-2 (Apr 2007)
- Fayers PM, Hand DJ, Bjordal K, Groenvold M - **“Causal indicators in quality of life research”**. *Quality of Life Research*, 6(5):393-406. (Jul 1997)
- Fortune-Greeley AK, Flynn KE, Jeffery DD, Williams MS, Keefe FJ, Reeve BB, Willis GB, Weinfurt KP; PROMIS Sexual Function Domain Committee - **“Using cognitive interviews to evaluate items for measuring sexual functioning across cancer populations: improvements and remaining challenges.”** *Quality of Life Research*, 18(8): 1085-1093. (Oct 2009)
- Groenvold M, Bjorner JB, Klee MC, Kreiner S - **“Test for item bias in a quality of life questionnaire”**. *Journal of Clinical Epidemiology*, 48(6): 805-816. (Jun 1995)
- Groenvold M – **“Quality of life in breast cancer adjuvant therapy: validation and pilot testing of a combination of questionnaires”**. *The Breast*, 6(2): 97-107 (Apr 1997)
- Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, Fayers PM, de Graeff A, Hammerlid E, Kaasa S, Sprangers MA, Bjorner JB; EORTC Quality of Life Group – **“The development of the EORTC QLQ–C15–PAL: a shortened questionnaire for cancer patients in palliative care”**. *European Journal of Cancer*, 42(1): 55–64 (Jan 2006)

- Kerr, C, Nixon A, Wild D - "**Assessing and demonstrating data saturation in qualitative inquiry supporting patient reported outcomes research**". *Expert Review of Pharmacoeconomics & Outcomes Research*, 10(3): 269-281 (Jun 2010)
- Koller M, Aaronson NK, Blazeby JM, Bottomley A, Dewolf L, Fayers P, Johnson C, Ramage J, Scott N, West K; EORTC Quality of Life Group – "**Translation procedures for standardised quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach**". *European Journal of Cancer*, 43(12): 1810-1820 (Aug 2007)
- Krueger RA & Casey MA – "**Focus Groups A practical guide for applied research**". (3rd edition). *Thousand Oaks, Sage Publications, Inc* (Apr 2000)
- Nunnally JC and Bernstein IH - "**Psychometric Theory**" (3rd edition). *New York: McGraw Hill*. (1994).
- Petersen MA, Groenvold M, Aaronson N, Blazeby JM, Brandberg Y, de Graeff A, Fayers P, Hammerlid E, Sprangers M, Velikova G, Bjorner JB; European Organisation for Research and Treatment of Cancer Quality of Life Group - "**Item response theory was used to shorten EORTC QLQ–C30 scales for use in palliative care**". *Journal of Clinical Epidemiology*, 59: 36–44 (Jan 2006)
- Petersen MA, Groenvold M, Aaronson N, Brenne E, Fayers P, Nielsen JD, Sprangers M, Bjorner JB; for the European Organisation for Research and Treatment of Cancer Quality of Life Group. "**Scoring based on item response theory did not alter the measurement ability of the EORTC QLQ–C30 scales**". *Journal of Clinical Epidemiology*, 58: 902–908 (Sept 2005)
- Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD - "**Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report**". *Value in Health*. 12(8):1075-1083. (Dec 2009)
- Streiner DL, Norman GR – "**Health Measurement Scales: A practical guide to their development and use**". (4th Edition). *Oxford: Oxford University Press, USA* (Dec 2003)

ISBN 978-2-930064-413