



The future of cancer therapy

EORTC QUALITY OF LIFE GROUP ITEM LIBRARY USER GUIDELINES

First Edition

2022

Claire Piccinin, Dagmara Kuliś, Andrew Bottomley, Kristin Bjordal, Corneel Coens, Anne-Sophie Darlington, Colin Johnson, Galina Velikova, Mogens Groenvold

on behalf of the EORTC
Quality of Life Group

Table of Contents

Aims	4
Glossary	4
Background & Rationale	5
General Guidelines & Recommendations	6
Implementing item lists from the Item Library	7
Item Selection.....	7
Clinical Trial Use	8
Single Items vs. Multi-Item Scales	8
Combining Item Lists with Standard Questionnaires	9
Selecting Specific Versions of Wording	10
QLQ-F17.....	10
CAT Items.....	11
Item Ordering	11
Response & Time Scale Selection	12
Formatting.....	13
Scoring, Analysis & Interpretation.....	13
Psychometrics.....	16
Electronic administration	16
Translations	16
Publications/Referencing	17
Naming	17
Sharing.....	18
References.....	19

Aims

This manual has been developed by the EORTC Quality of Life Group (QLG), with the aim of providing guidance on the use of the EORTC Item Library. These guidelines should be referred to by anyone who has agreed to the Item Library's terms of use for the creation of item lists. This is a developing document and will be updated on a regular basis.

Glossary

AE: Adverse event

CAT: Computerized adaptive testing

CTT: Classical test theory

eCOA: Electronic clinical outcome assessment

EORTC: European Organisation for Research and Treatment of Cancer

EORTC Item Library: Interactive online platform displaying EORTC QLG measures and enabling the creation of customized item lists

ePRO: Electronic patient-reported outcome administration (administered via e.g. a computer or tablet)

HRQOL: Health-related quality of life

IRT: Item response theory

Item: Question from any EORTC measure

Item List: Customized selection of items

Module: EORTC questionnaire

PMDC: Project and Module Development Committee

PRO: Patient-reported outcome

QLD: Quality of Life Department

QLG: Quality of Life Group

QLQ: Quality of life questionnaire

QLQ-C30: EORTC Core Quality of Life Questionnaire

QLQ-F17: EORTC Core Function Questionnaire

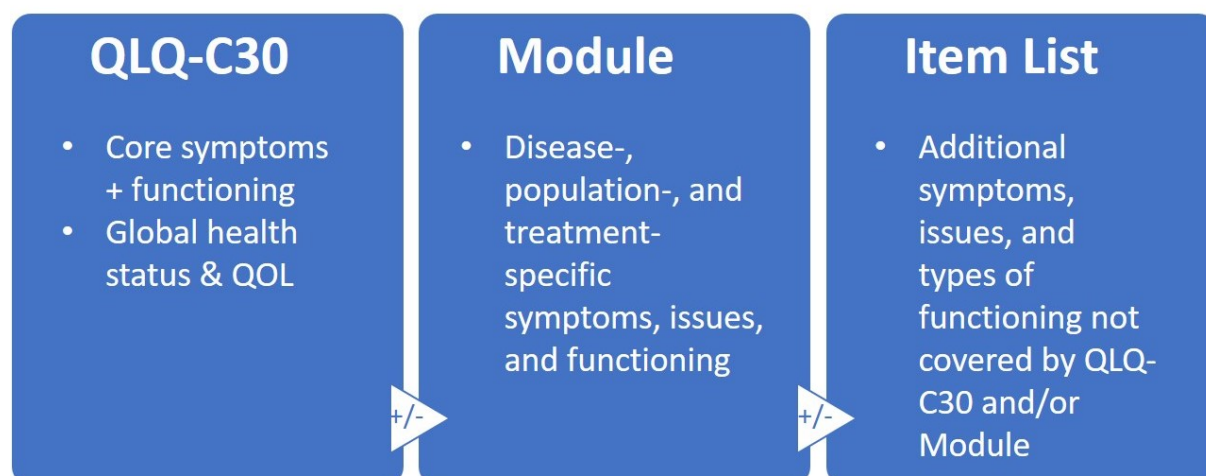
QOL: Quality of life

Standalone questionnaire: EORTC questionnaire which can be administered on its own (i.e., not in conjunction with the QLQ-C30)

Background & Rationale

The historical approach to questionnaire development has helped to ensure high validity, reliability, and cross-cultural meaningfulness of the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group (QLG) questionnaires. However, the lengthy process of validation, wherein questionnaires can take several years to finalize, is less adapted to some forms of contemporary therapeutic intervention or rare cancers, whose symptoms and side effects might not always be captured by existing questionnaires (1–3). In light of rapidly evolving treatment modalities and the need to ensure high content validity, the new strategy recently put forward by the EORTC QLG proposes increased flexibility in the form of custom-made ad hoc item lists that can be used to supplement the core QLQ-C30 and, when relevant, a module (see Figure 1). Users have the ability to choose from the wide range of items in the Item Library (4) when creating item lists, in order to better capture symptoms, functional status, and events that are relevant to the research questions under investigation.

Figure 1. Updated EORTC QLG Measurement Strategy



The Item Library currently contains over 1,000 questions, referred to as ‘items’, which have been systematically developed, tested, validated, and translated since 1993. Since their original development, many items have undergone modifications to reflect important cultural and linguistic changes. These items are derived from the full portfolio of validated EORTC QLG instruments, including the QLQ-C30, modules, standalone questionnaires, and computerized adaptive testing (CAT) item banks. These measures have been developed and validated cross-culturally following the EORTC Module Development Guidelines (5) and published extensively in peer-reviewed studies. Some items have been used in several questionnaires, to capture different underlying issues. The portfolio of EORTC QLQ questionnaires is widely used in oncology to measure patient-reported

Quality of Life (QOL) across a variety of domains and cancer sites (6–10). Together, the items cover multiple symptomatic and functional domains, including physical symptoms (e.g., nausea), social functioning (e.g., socializing with friends), cognitive functioning (e.g., memory and concentration), emotional functioning, physical functioning, sexual functioning, activities of daily living (ADL), communication, and satisfaction with provision of care (11). Although the Item Library provides registered users with considerable flexibility, the QLG has compiled general guidelines to help guide usage in order to promote good practice and scientific rigour. Over time, as more data are collected on the use and implementation of the Item Library, this document will evolve to reflect new evidence and findings. Existing guidelines on the design and use of PROs, provided by the Consolidated Standards of Reporting Trials (CONSORT) PRO Extension (12), United States Food and Drug Administration (FDA) (13,14), European Medicines Agency (EMA) (15), and Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-PRO Extension (16) were consulted in order to extract relevant recommendations.

General Guidelines & Recommendations

It is crucial that the ad hoc output created through use of the Item Library be referred to as an ‘item list’. Customized lists of items cannot be considered fully validated static questionnaires since they will not have undergone the psychometric testing required of EORTC QLG questionnaires and instruments (see EORTC Project and Module Development Committee (PMD) Module Development Guidelines (5)). Indeed, item lists should generally be considered ad hoc lists aimed at supplementing the EORTC QLQ-C30 and, when relevant, other EORTC modules. Although users may choose to undertake such psychometric analyses on their own, this work should be clearly distinguished from that which is carried out by the QLG (see [Psychometrics](#) section below). Modified questionnaires, such as those that have been shortened (through removal of items), should be distinguished from fully validated measures as item lists, and their use should be approached with careful consideration.

Although users may wish to remove specific items from fully validated questionnaires and modules, the EORTC QLG recommends that this method be approached with considerable caution. Users should only remove items and sections of validated questionnaires if there is ample evidence to suggest that they are not relevant and/or may be overly burdensome to patients. Such a modified measure must also be referred to as an item list, to distinguish it from a fully validated static measure. Users choosing to take this approach are advised that recommended interpretation of resulting scales and data will no longer be fully applicable and may need to be changed from that in existing published EORTC guidelines. As such, modifications should only be undertaken if the

advantages outweigh the risks. Due to the underlying validated multidimensional scoring structure, the EORTC QLG advises against removing parts of a scale. In general, items should be added or removed in conjunction with the other items that make up the underlying domain scale, unless there is a strong rationale to omit only certain items.

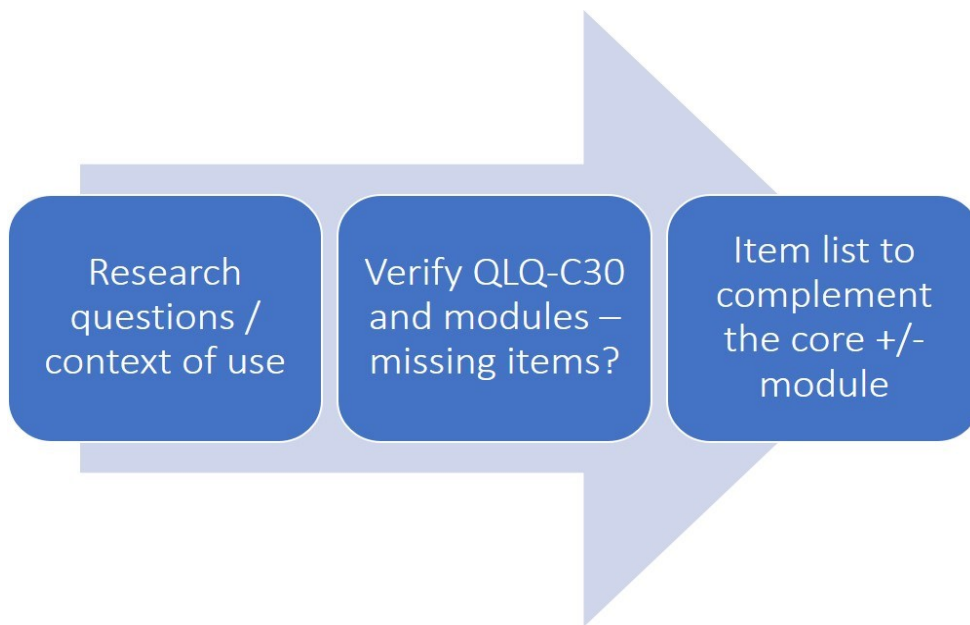
Given its extensive use and validated multidimensional model of health-related quality of life (HRQOL), the EORTC QLG advises against modifying the Core Quality of Life Questionnaire (QLQ-C30), unless there is strong evidence to support such an action. Modifications of other validated questionnaires (e.g., modules and standalone questionnaires) should also be approached with caution and follow the rationale presented above. In all cases, the resulting measures must be referred to as item lists. Similarly, if a user wishes to add items to an existing validated questionnaire or module, this must also be carried out within the context of an item list, such that the additional items remain clearly differentiated from the questionnaire/module. In such a case, the item list may be used to measure relevant clinical domains and symptoms that are not included (in whole or partially) in the original questionnaires.

Implementing item lists from the Item Library

Item Selection

Item selection must be driven by theoretical and empirical evidence. Users are encouraged to conduct thorough reviews of the literature in order to identify possible symptoms (e.g., toxicities and adverse events [AEs]) and QOL issues (e.g., functional impairment, etc.) which are relevant to the research questions under investigation. Where available, preclinical data may also be used to inform this decision. In an effort to promote content validity, users are encouraged to carry out qualitative debriefing interviews with patients and healthcare professionals, to help identify items which may be most relevant. In cases where items are not covered by existing questionnaires (i.e., the QLQ-C30 and modules), the selected items can be used to construct an item list (see Figure 2).

Figure 2. Implementing item lists from the Item Library



Clinical Trial Use

During early-phase non-randomized trials, when less is known about the treatment under investigation, item lists may represent a more preliminary collection of anticipated AEs and symptoms, based on existing empirical (e.g., preclinical) and theoretical data (17). As a study progresses, item selection may be more refined as new symptoms and AEs are identified. During later-phase randomized trials, when more is known about the treatment in question, as well as possible associated AEs and symptomatic toxicities, the issues and symptoms selected for item lists will likely be more focused. It is important to include the same items in all treatment arms, regardless of expectations, in order to minimize the potential for bias, avoid underreporting of AEs and symptoms, and ensure comparability across arms. It should be noted, though, that use of a highly customized item list, in contrast to a fully validated static questionnaire, may complicate comparability across studies, given that there may be fewer overlapping outcomes for comparison. As such, measurement of core outcomes (e.g., using the QLQ-C30) is recommended, where relevant.

Single Items vs. Multi-Item Scales

The choice of single items versus multi-item scales will depend on the symptom or domain under investigation. While some concepts might be captured by one item only (e.g., the presence of a specific symptom, like nausea or appetite loss) others may be more nuanced and as such require more items. Higher-level concepts, such as physical functioning, which involve a number of different activities, generally require several items and are often captured (as in the case of the QLQ-C30 (18) Physical Functioning scale) by a multi-item scale. Given the wide scope of activities that constitute physical functioning (e.g., walking, lifting, carrying out daily activities), one item broadly assessing

this domain would not be sufficient and would be difficult to report on. Instead, several items which provide concrete examples of different types of physical functioning are more meaningful to patients and paint a more comprehensive picture of functioning overall. Following this rationale, when items form part of a scale, addition or removal of items should occur on the scale level, whenever possible (e.g., if the scale is relevant to the research questions and inclusion is feasible). For many concepts, this helps to ensure a more detailed assessment while preserving the underlying structure, which, in turn, facilitates scoring. Users must take special care when using items that are part of a conditional outcome, as this will influence dependent items (e.g., some items are only applicable to patients who have undergone specific medical procedures or experienced relevant symptoms). The specific choice of items must ultimately reflect the research questions under investigation. For example, if a user is interested in assessing the presence of a symptom, one item may suffice. However, if symptom severity, impact on functioning, and interference with daily activities are also intended to be captured, this will likely necessitate at least a few items and perhaps the use of a multi-item scale, depending on the availability in the Item Library.

Combining Item Lists with Standard Questionnaires

When an item list is used in conjunction with the QLQ-C30 and (when relevant) a module, users must be careful to ensure that the selected items are not already accounted for in one of the other questionnaires being used.

Patient Burden Considerations

Although the specific number of items selected by users for any given item list will vary considerably as a function of the research questions and context of use, users are advised to avoid duplication of concepts in an effort to minimize patient burden. Specifically, if a symptom, issue, or type of functioning is already captured with one questionnaire, it should not generally be included in the item list, unless there is a specific rationale for its inclusion. Such potential for duplication is especially relevant in studies in which other instruments and PRO measurement systems are used as items may quickly accumulate, which could lead to increased burden. Although lengthier measures (or batteries of measures) may be more burdensome for patients to complete, timing of administration is also important. For example, patients may be more willing to complete a long measure (or battery of measures) if administration of such measures does not occur regularly (19,20); whereas longer measures (or batteries of measures) may be more burdensome for patients to complete when administration occurs regularly (e.g., weekly).

Perceived burden of questionnaire completion may also vary as a function of several factors and differ considerably from one individual to another. When patients feel that the concepts being

measured are highly relevant, they may be less burdened by the number of items and length of measures (21,22). Other factors such as administration mode (e.g., paper vs. electronic), patients' literacy levels, the presence of sensitive items, and the formatting of the measure have also been linked to burden (19). Patients' disease stage and severity, along with the perceived difficulty of questionnaire completion, may also be linked to burden.

Whenever possible, users are strongly encouraged to carry out a pilot test to more closely assess duration of administration and burden for patients. Adopting a robust approach to item selection by involving both patients and healthcare professionals can also help manage issues of potential burden by helping to ensure that the items included are relevant and understandable and presented in such a way that completion of measures is feasible and not overly burdensome.

Selecting Specific Versions of Wording

For items for which there are multiple versions, users are encouraged to opt for the 'recommended wording' version as indicated in the Item Library. These versions represent items for which (after years of development and evolution) the wording has been deemed the most comprehensible, grammatically sound, and easily translatable. However, in cases where an item list is comprised of a simple subset of items from a standard questionnaire, the original wording (as in the standard questionnaire) should be retained, to ensure comparability with the standard (source) questionnaire.

Although users can choose between different versions of wording for some items, beyond this they cannot modify wording of the Item Library items and other text (e.g., response scales and instructions). This is to ensure that items remain in their original validated form, and to facilitate analyses across item lists and other EORTC QLG instruments. Moreover, users cannot use items originating from another source outside of the Item Library. Users who perceive an error in any text or translation can report this to the administrators of the Item Library: admin.itemlibrary@eortc.org.

QLQ-F17

The FDA has recommended assessing well-defined functioning scales (e.g., physical and role function), symptomatic adverse events, overall side effect impact, and disease-related symptoms relevant to a given trial, while ensuring minimal patient burden (1,3,23). In response to this communication, the EORTC has taken a pragmatic approach in creating the EORTC QLQ-F17 Core Function Questionnaire (24). Derived from the QLQ-C30, it contains only the functional scales (i.e., physical, role, emotional, cognitive, and social functioning, and global health status/QOL) from the questionnaire, without the symptom scales or single items, in their original order and wording. By offering a shorter alternative to the QLQ-C30, the QLQ-F17 helps to ensure flexible PRO

measurement by allowing users to assess well-defined functioning scales, along with other modules and item lists as relevant. Although the QLQ-F17 is comprised of a subset of items from the QLQ-C30, it should be noted that it is not an ad hoc item list. While it is already available for use, the QLQ-F17 will eventually undergo equivalence testing, to ensure that its psychometric properties are equivalent to those of the QLQ-C30.

CAT Items

Users who wish to use items which have only been developed for use in the EORTC's CAT item banks (25,26), should liaise with the Quality of Life Department (QLD).

Given that these items have been developed using item response theory (IRT), and not classical test theory (CTT) as was used to develop the QLG's static questionnaires, they should generally be scored following the appropriate IRT algorithms and not the standard scoring instructions. The QLG's CAT team can provide further assistance for users wishing to use CAT items in a computerized form as well as those wishing to construct short forms. Users can read more about the CAT project here:

<https://qol.eortc.org/projectqol/eortc-cat/>

Item Ordering

When an item list is formed by removing items from a validated instrument, the order of the remaining items should generally remain unchanged, to preserve the intended structure as best as possible. This also helps to ensure that items with the same instructions, conditions (where relevant), and response and time scales remain grouped together. When items are added to a validated instrument, a few different options are available.

If an item list is designed to be used in conjunction with a fully validated questionnaire (e.g., the QLQ-C30), the item list should be formatted as an additional instrument (not inserted into the format of the QLQ-C30). This helps to preserve the structure of the validated measure and differentiates the item list in a more ad hoc manner. However, if the validated instrument contains items that are potentially sensitive in nature (e.g., assessing sexuality and intimacy), it is best to integrate the items from the item list so that the potentially sensitive items from the validated instrument remain at the end of the measure, as they might affect patients in such a way that responses to subsequent items could be influenced (20).

When formatting item lists, new items should be integrated so that similar formats remain grouped together. As with the standard validated instruments, items with different response and/or time scales should be clearly distinguished. With regard to item content, efforts should be made to ensure that similar items are grouped together, provided they meet the aforementioned criteria

(i.e., matching response and time scales). For example, items assessing symptom presence may be grouped together (with specific localizations/sites also listed together, when relevant) while those assessing functional impairment can also be listed together, keeping in mind the other criteria discussed earlier. As above, potentially sensitive items capturing sexuality and intimacy should generally be placed at the end of the item list, to ensure that subsequent responses are not influenced.

Ultimately, given the many different factors which need to be taken into consideration, along with the specific research questions and needs of any given study, the issue of ordering needs to be examined on a case-by-case basis. To date, only limited research has looked at possible ordering effects in EORTC QLG instruments. As such, future work should also be carried out and consulted to inform such decisions.

Response & Time Scale Selection

The choice of specific response and time scales will also vary as a function of specific research questions. It is generally recommended that items be administered using the same response and time scales with which they were originally validated. However, some items (e.g., those used in multiple questionnaires) may have been tested with more than one possible time and response scale. Beyond that, there may also be cases in which more flexibility is sought. As such, the following general guidelines should be taken into consideration.

When users require a precise response for dependent items, a dichotomous scale (e.g., 'yes' vs. 'no') may be most appropriate. When more nuance is required, which is generally the case, an ordinal scale may be used (typically, responses range from 1-4). In general, it is recommended that users employ the response scale associated with a specific item, since this represents the format in which the item will have been tested and validated. However, depending on the specific research question(s) and context, there may be a rationale for using an alternative time or response scale, and such flexibility is possible using the Item Library. For time scales, it is generally also recommended that users select the time scale (from the list of validated Item Library options) with which the item was originally validated. In most cases, this will capture events and symptoms occurring within the last week (i.e., a 1-week time scale). This is based on evidence which suggests that longer recall periods (e.g., of 2, 3, and 4 weeks) are associated with small but increasing rates of measurement error (27). However, users must consider the symptoms, events, and research questions being investigated when making such a decision, since some items (e.g., issues concerning sexuality) may be best represented and measured by a longer recall period. Study design, context of

use, and timing of instrument/item list administration may also have an impact on the choice of specific recall period(s).

Formatting

Once they have been published in the online platform, users may request to download item lists directly in .csv format, which lends itself especially well to implementation in electronic patient reported outcome (ePRO) platforms. For users who wish to administer their item lists and other measures on paper, the QLD formats these creations following the templates used for the standard measures, incorporating the EORTC logo and formatting and citing the relevant copyright. As discussed in the Item Ordering section, items with the same time and response scales are grouped together (as well as the other ordering conventions highlighted). In addition to the automatically exported .csv files, at least one Word template is provided for individuals using ePRO, to highlight the relevant graphics, format, layout, and copyright.


Scoring, Analysis & Interpretation

With increasing flexibility comes an increased need to exercise caution and rigour when analyzing and interpreting item list data. Users have the possibility of combining several different items, and response and time scales, which adds a considerable level of complexity to scoring and interpretation of data. The standard approach to scoring should be followed, with items being scored according to the scale structure they originate from, following the guidelines as stipulated in the relevant EORTC Scoring Manual(s). These guidelines include instructions that single items and multi-item scales be scored as such and transformed to a 0-100 scale with appropriate direction. If multi-item scales are used, they may be summarized and transformed following the procedure outlined in the relevant scoring manual. However, scales can only be scored as such if they include all of the intended items. The example below (see Figure 3) highlights an item list (IL19) comprised of the physical functioning scale from the QLQ-C30. Since this item lists contains all of the items from the QLQ-C30 physical functioning scale, it can be scored as a multi-item scale, following the instructions outlined in the relevant scoring manual.

For descriptive purposes, users may consider using the raw, non-transformed scores of single items and single item scales and presenting these as simple proportions for each of the relevant response categories. Such an approach may be easier to present and interpret, compared to the use of mean scores. Indeed, the evidence suggests that there is little difference between analyzing the scale as continuous or categorical (27). However, such an approach should be regarded as novel, and distinguished appropriately, given the divergence from the standard scoring instructions.

Figure 3. Item List Example (Multi-Item Scale)

ENGLISH



EORTC IL19

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.


	Not at All	A Little	Quite a Bit	Very Much
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2. Do you have any trouble taking a <u>long</u> walk?	1	2	3	4
3. Do you have any trouble taking a <u>short</u> walk outside of the house?	1	2	3	4
4. Do you need to stay in bed or a chair during the day?	1	2	3	4
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

© Copyright 2018 EORTC Quality of Life Group. All rights reserved.

The following example (Figure 4), comprised of two single-symptom items, would be scored as such (i.e., single items) following the standard scoring algorithms.

Figure 4. Item List Example (Single Items)

ENGLISH



EORTC IL154

Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week:

	Not at all	A little	Quite a bit	Very much
1. Have you had diarrhea?	1	2	3	4
2. Have you had a rash?	1	2	3	4

© Copyright 2011 EORTC Quality of Life Group. All rights reserved.

If only a subset of items is used from any multi-item scale, these items must be scored as single items, but it is recommended that any resulting publication/dissemination specifies that the result is not an official single-item scale but rather an ad-hoc selected item. Since the psychometric properties would not be established at the single-item scale level for such an item, it is crucial to distinguish any other scoring/analysis method as ad hoc and not official.

Users should always exercise a high level of caution when interpreting scores from item lists, given that the single- and multi-item scales have been validated for use in existing source questionnaires and modules (with specific populations) but not in custom-made item lists, where the addition and ordering of other items could potentially influence patients' responses. It is currently not possible to combine multiple single items into a single score – any user wishing to undertake such analyses may only do so in a strictly exploratory manner. Single items should therefore be scored as such.

As indicated previously, items developed only for use in CAT item banks should generally be scored following the appropriate IRT algorithms and not the standard scoring instructions. Users can liaise with the EORTC QLD for more information.

Although there are currently no official recommendations for the analysis of QLQ-C30 outcomes specifically, whether pertaining to single items or otherwise, the general recommendation is to first develop a research hypothesis with clear objectives to be answered. A statistical analysis plan can then be developed, and, if necessary, tweaked based upon observed data availability and/or distribution. For outcomes based on single items, the use of statistical methods that assume normal continuous data is not advised, as they may lead to a loss in power; although, if the sample size is sufficiently large, the continuous assumption may still provide sufficient approximation (18). However, for single item scales, ordinal methods are preferable as they are easier to interpret. Equidistance between the response categories is a commonly made assumption to simplify the ordinal methods to, for example, a proportional odds model. For more detailed recommendations on the analysis of PRO data, tailored to the specific context of use and research questions, we recommend referring to published guidelines from the Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) Consortium (28).

There are currently only published minimal important difference (MID) scores provided for full instruments and modules. There are presently no guidelines for interpretation of users' item list scores. Users may consult published papers that will give some insight into clinical significance of items and scales, but caution is required when applying such information to item list data. Although we are not currently able to attribute clinical significance to item list scores, this will be investigated over time as part of the ongoing research on the Item Library being carried out by the EORTC QLQ.

Psychometrics

In contrast to the QLQ-C30, standalone, and module questionnaires, item lists created using the Item Library will not have undergone additional psychometric testing. Although items have already undergone extensive validation in their instrument(s) of origin, when items are removed from their source questionnaires and administered in populations in which they have not been tested (e.g., in a new disease site), the psychometric properties may not be retained. Thus, users wishing to cite the psychometric properties of item lists will have to do so at their own discretion, and the EORTC QLG takes no responsibility for the results and interpretation of such forms of testing. Indeed, such testing and validation should be clearly described as more experimental and be distinguished from that which is carried out by the QLG. In the absence of strict psychometric testing, we caution users against attempts to report on reliability and validity given insufficient data/analyses to support such claims.

In the future, frequently used item lists (e.g., treatment- or symptom-specific lists) may start to be systematically validated, following QLG guidelines, but these guidelines will need to be developed in parallel with the PMDC and with the consensus of the QLG.

Electronic administration

Electronic patient-reported outcome (ePRO) administration of item lists should be carried out following the guidance as stipulated in the EORTC ePRO Guidelines (29) for implementing QOL instruments in electronic platforms/applications. Users should also liaise with the EORTC QLD to ensure that the appropriate electronic clinical outcome assessment (eCOA) vendor agreement is in place, where relevant.

Translations

The Item Library currently contains over 1,000 unique items, some of which are translated in over 100 languages. However, users must verify whether or not selected items are available in the desired language, since the available translations vary across items. When an item list is comprised of items from different source questionnaires (i.e., the questionnaires in which the items were originally developed), some translations may not be available. If an item or items are not available in the language(s) of your choice, a translation request can be sent to the EORTC QLD Translation Team. Under no circumstances are users allowed to translate items on their own – this must be done following the EORTC QLG Translation Procedure (30,31), under the supervision of the EORTC QLG Translation Team Leader.

Publications/Referencing

Users may not reproduce the custom-made item list in any publications or elsewhere without prior written authorization from the EORTC QLQ. The specific terms of use are outlined in the relevant user agreement template. As with the standard questionnaires, specimen versions of the item list may be shared and made available, even before a user agreement is fully executed. When referencing item lists created using the Item Library, users must refer to the website: www.eortc.be/itemlibrary/

Users may refer to the EORTC questionnaires (whether core, standalone, or module) from which specific items originate, but only in such a way that the item list is clearly differentiated from the validated questionnaires. When an item list is used in conjunction with a validated questionnaire (e.g., the QLQ-C30 and/or a module), the validated questionnaire must also be cited as per the usual procedure by referring to the appropriate publication.

Naming

Custom-made item lists and Item Library creations are assigned unique alphanumeric identifiers following the ILXX format (IL = item list). This convention helps to ensure that each creation is uniquely identifiable, which in turn helps to facilitate tracking and use. If the same item list is used in a new study or context, the original name is retained (provided all of the required items and elements are present). This helps to ensure consistency across studies and within documentation and may also facilitate future comparative analyses. Although users may refer to the original source questionnaire in a description or publication, an item list constructed through removal of items from a validated questionnaire (i.e., an item list comprised of a subset of items from a questionnaire) must be clearly differentiated as such, and assigned an item list number. Additional descriptive information can be added to the 'Description' field in the item list view in the Item Library (e.g., pertaining to key scales, types of functioning, symptoms, or population) to highlight the content. As in the example below (Figure 5), an item list assessing lymphoedema symptoms, designed to be administered in conjunction with the QLQ-C30, can be described as such. Given that such descriptive classifications will likely be study-specific, the alphanumeric identifier (e.g., IL100) remains the official name for an item list.

Figure 5. Item List Example in Item Library Online Platform

The screenshot displays the 'Item List' interface for 'IL100'. At the top, there is a header with the item ID 'IL100', a folder icon, and three filters: 'GENDER: Neutral', 'TYPE: Item list', and 'LAST MODIFIED: Nov 2, 2021'. Below the header, there are two tabs: 'Questions' (selected) and 'Questionnaire info'. The 'Questions' tab shows a list of 8 questions, each with a three-dot menu icon to its right. The 'Questionnaire info' tab is also visible, showing a 'Description' and 'Instructions' section. The 'Description' states: 'This item list has been designed to be used in conjunction with the EORTC QLQ-C30 as a screening tool to investigate possible lymphedema symptoms.' The 'Instructions' section contains three bullet points, each with a 'filter' button and a 'Q, 1-9' label.

Questions English

Questionnaire info Question details

Description

This item list has been designed to be used in conjunction with the EORTC QLQ-C30 as a screening tool to investigate possible lymphedema symptoms.

Instructions

- Patients sometimes report that they have the following symptoms or problems. Q, 1-9 filter
- Some of these relate to the genital area. Q, 1-9 filter
- Whether or not you have had surgery in your genital area, please indicate the extent to which you have experienced these symptoms or problems. Q, 1-9 filter

Sharing

The item list can only be reproduced for the strict purpose of generating sufficient copies for use in the intended study. Any other use is subject to prior written authorization from the EORTC QLQ.

Under no circumstances may it be distributed to third parties for sale, rental, lease, lending, or any other services. For specific guidelines regarding usage, sharing, and publishing, please refer to the user licencing agreement. Users should be aware that once it is published in the Item Library, their item list becomes visible and available to other registered users. This is to increase transparency and to foster collaboration across the scientific community. All identifying information (e.g., the name of the user who created the item list) is removed to protect the identity of all users, in accordance with EU data protection laws.

References

1. Kluetz PG, Slagle A, Papadopoulos EJ, Johnson LL, Donoghue M, Kwitkowski VE, et al. Focusing on core patient-reported outcomes in cancer clinical trials: Symptomatic adverse events, physical function, and disease-related symptoms. *Clin Cancer Res*. 2016.
2. Groenvold M, Aaronson NK, Darlington ASE, Fitzsimmons D, Greimel E, Holzner B, et al. Focusing on core patient-reported outcomes in cancer clinical trials - Letter. *Clinical Cancer Research*. 2016.
3. Kluetz PG, Papadopoulos EJ, Johnson LL, Donoghue M, Kwitkowski VE, Chen WH, et al. Focusing on core patient-reported outcomes in cancer clinical trials - Response. *Clinical Cancer Research*. 2016.
4. EORTC Quality of Life Group. EORTC Item Library [Internet]. [cited 2021 Nov 3]. Available from: <https://www.eortc.be/itemlibrary/>
5. Bjordal K, Bottomley A, Gilbert A, Martinelli F, Pe M, Sztankay M, et al. EORTC Quality of Life Group Module Development Guidelines [Internet]. 5th ed. Brussels, Belgium: EORTC; 2021. Available from: <https://qol.eortc.org/manuals/>
6. Efficace F, Fayers P, Pusic A, Cemal Y, Yanagawa J, Jacobs M, et al. Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: A pooled analysis of 557 trials. *Cancer*. 2015.
7. Bottomley A, Efficace F, Thomas R, Vanvoorden V, Ahmedzai SH. Health-related quality of life in non-small-cell lung cancer: Methodologic issues in randomized controlled trials. *Journal of Clinical Oncology*. 2003.
8. Efficace F, Bottomley A, Van Andel G. Health related quality of life in prostate carcinoma patients: A systematic review of randomized controlled trials. *Cancer*. 2003.
9. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *Journal of Clinical Oncology*. 2008.
10. Mierzynska J, Piccinin C, Pe M, Martinelli F, Gotay C, Coens C, et al. Prognostic value of patient-reported outcomes from international randomised clinical trials on cancer: a systematic review. *Lancet Oncol* [Internet]. 2019;20(12):e685–98. Available from: [http://dx.doi.org/10.1016/S1470-2045\(19\)30656-4](http://dx.doi.org/10.1016/S1470-2045(19)30656-4)
11. Piccinin, Claire; Kuliś, Dagmara; Bottomley, Andrew; Martinelli, Francesca; Mierzynska, Justyna; Pe, Madeline; Groenvold M. Organizing the EORTC item library: a mixed methods exploratory content analysis. *Qual Life Res*. 2018;27:S106.
12. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA - Journal of the American Medical Association*. 2013.
13. FDA. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance. *Health Qual Life Outcomes*. 2006.
14. FDA. Patient-Focused Drug Development: Methods to Identify What Is Important to Patients Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders, Draft

Guidance [Internet]. 2019 [cited 2021 Nov 4]. Available from:
<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-methods-identify-what-important-patients-guidance-industry-food-and>

15. European Medicines Agency Committee for Medicinal Products for Human Use (CHMP). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man: The use of patient-reported outcome (PRO) measures in oncology studies [Internet]. 2016 [cited 2021 Nov 4]. Available from: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf
16. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols the spirit-pro extension. *JAMA - J Am Med Assoc*. 2018.
17. Trask PC, Dueck AC, Pault E, Campbell A. Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events: Methods for item selection in industry-sponsored oncology clinical trials. *Clin Trials*. 2018;15(6):616–23.
18. Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain*. 2008.
19. FDA. Guidance for Industry; Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. *Clin Fed Regist*. 2009.
20. Piccinin C, Kulis D, Bottomley A, Bjordal K, Coens C, Darlington AS, et al. PCN296 Development of scientific guidelines for use of the EORTC Item Library in cancer clinical trials. *Value Heal*. 2019.
21. Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: Is shorter better? A review and meta-analysis. *Value Heal*. 2011.
22. Atkinson TM, Schwartz CE, Goldstein L, Garcia I, Storfer DF, Li Y, et al. Perceptions of Response Burden Associated with Completion of Patient-Reported Outcome Assessments in Oncology. *Value Heal*. 2019.
23. FDA. Core Patient-Reported Outcomes in Cancer Clinical Trials Guidance for Industry DRAFT GUIDANCE. 2021;(June 2021). Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials>
24. Groenvold, Mogens; Aaronson, Neil K.; Bjordal, Kristin; Darlington, Anne-Sophie; Kuliś, Dagmara; Piccinin, Claire; Bottomley A. Measuring function: Introducing the EORTC QLQ-F17 Core Function Questionnaire. *Qual Life Res*. 2019;28:S36.
25. Petersen MA, Aaronson NK, Arraras JI, Chie WC, Conroy T, Costantini A, et al. The EORTC CAT Core—The computer adaptive version of the EORTC QLQ-C30 questionnaire. *Eur J Cancer* [Internet]. 2018;100:8–16. Available from: <https://doi.org/10.1016/j.ejca.2018.04.016>
26. Petersen MA, Aaronson NK, Conroy T, Costantini A, Giesinger JM, Hammerlid E, et al. International validation of the EORTC CAT Core: a new adaptive instrument for measuring core quality of life domains in cancer. *Qual Life Res*. 2020.
27. Donneau AF, Mauer M, Coens C, Bottomley A, Albert A. Longitudinal quality of life data: a comparison of continuous and ordinal approaches. *Qual Life Res*. 2014.

28. Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. *The Lancet Oncology*. 2020.
29. Kuliś D, Holzner B, Koller M, Ruyskart P, Itani A, Williams P, et al. Guidance on the implementation and management of EORTC quality of life instruments in electronic applications [Internet]. Brussels, Belgium: EORTC; 2018. Available from: <https://qol.eortc.org/>
30. Kuliś, Dagmara; Bottomley, Andrew; Velikova, Galina; Greimel, Eva; Koller M. EORTC Quality of Life Group Translation Procedure. 4th ed. Brussels, Belgium: EORTC; 2017.
31. Kuliś, Dagmara; Piccinin, Claire; Bottomley, Andrew; Groenvold M. Translations of item lists created with the EORTC Item Library: updated translation and review procedure. *Qual Life Res*. 2018;27:S107.

© Copyright EORTC, Brussels, 2022

D/2022/6136/01

ISBN 978-2-930064-49-9