

International validation of the EORTC QLQ-ANL27, a field study to test the anal cancer-specific health-related quality of life questionnaire

Running title: EORTC QLQ-ANL27 Validation

Abstract

Background

The European Organisation for Research and Treatment of Cancer (EORTC) health-related quality of life (HRQoL) questionnaire for anal cancer (QLQ-ANL27) supplements the EORTC cancer generic measure (QLQ-C30) to measure concerns specific to people with anal cancer treated with chemoradiotherapy (CRT). This study tests the psychometric properties and acceptability of the QLQ-ANL27.

Materials and methods

People with anal cancer were recruited from 15 countries to complete the QLQ-C30 and QLQ-ANL27 and provide feedback on the QLQ-ANL27. Item responses, scale structure (multi-trait scaling, factor analysis), reliability (internal consistency and reproducibility) and sensitivity (known group comparisons and responsiveness to change) of the QLQ-ANL27 were evaluated.

Results

Data from 382 people were included in the analyses. The EORTC QLQ-ANL27 was acceptable, comprehensive, and easy to complete, taking an average 8 minutes to complete. Psychometric analyses supported the EORTC QLQ-ANL27 items and reliability (Cronbach's alpha ranging from 0.71 to 0.93 and test-retest coefficients above 0.7) and validity of the scales (particularly non stoma bowel symptoms and pain/discomfort). Most scales distinguished people according to treatment phase and performance status. Bowel (non-stoma), pain/discomfort and vaginal symptoms were sensitive to deteriorations over time. The stoma-related scales remained untested due to low

numbers of people with a stoma. Revisions to the scoring and question ordering of the sexual items were proposed.

Conclusions

The QLQ-ANL27 has good psychometric properties and is available in 16 languages, for people treated with CRT for anal cancer. It is used in clinical trials and has a potential role in clinical practice.

Keywords

Anal cancer; EORTC QLQ-ANL27; Health-related quality of life; Patient-reported outcomes

Introduction

Background

Anal carcinoma is rare, accounting for less than 1% of all cancer diagnoses [1-3]. Chemoradiotherapy (CRT) is the mainstay of treatment for localised disease achieving high locoregional control and cure rates [1]; however, patients experience significant acute and late toxicity which impact health-related quality of life (HRQoL) [4].

Commonly reported problems related to bowel function (e.g., diarrhea, incontinence, tenesmus, increased frequency, and urgency) [5-9] impact daily and social activities [10] and cause embarrassment and distress [7]. A stoma, when fitted for significant symptoms or disease at presentation or follow up, can also impact HRQoL [11]. Sexual dysfunction following anal cancer is widely documented [4]. In women, anal cancer and its treatment can lead to painful sexual intercourse, diminished enjoyment in sex, vaginal narrowing, and dryness, while 60-71% men report erectile problems [8, 12, 13]. For both sexes, reduced libido can be severe and chronic [14]. A recent review [4] highlighted skin reactions such as radiation dermatitis and desquamation around the perianal region as common radiotherapy-related toxicities. Patients also report burning, itchy and irritable skin around the treatment region causing pain and discomfort [7]. Furthermore, urinary function problems (urinary frequency and poor bladder control) are common long-term side effects of radiotherapy [12].

In the absence of a validated, anal cancer specific HRQoL measure, previous research in anal cancer has relied upon generic cancer measures (e.g., EORTC QLQ-C30 (QLQ-C30) [15]) or measures developed for other cancers (e.g., EORTC QLQ-CR29 (QLQ-CR29) [16, 17], FACT-C [18]). Although there are overlapping HRQoL concerns with colorectal cancer, there is a clear need for a validated anal cancer specific measure for use in clinical trials and routine care [4].

The development of the EORTC QLQ-ANL27 (QLQ-ANL27) follows the EORTC Quality of Life Group (QLG) module development framework [19]; in summary, 197 issues were generated from a literature review [4] and interviews with 43 patients. The list was refined to 134 issues and reviewed by 34 health care professionals (HCPs) and 10 patients. A 65-item draft module was pilot tested by 100 patients and led to the modification and removal of questions [7]. The objective of this study is to test the hypothesised scale structure, reliability, responsiveness to change and validity of the QLQ-ANL27 in people from different countries diagnosed with anal cancer and treated with CRT with or without a stoma.

Methods and Materials

Study design and participants

The study protocol is available from the authors upon request and was peer-reviewed by executive members of the EORTC as part of the grant review process. Over a 3-year period (December 2017 to November 2020), participants were recruited from 19 centers across 15 countries: Australia, Brazil, Canada, Cyprus, Denmark, France, Greece, Italy, Norway, Portugal, Spain, Sweden, Turkey, The Netherlands, and UK, to ensure a culturally and linguistically diverse sample. The study was approved by the Sponsor and the national research ethics committee and health research authority. Local approvals were granted at each participating site.

Participant eligibility included (a) histologically confirmed diagnosis of newly diagnosed or recurrent locoregional anal (squamous or cloacogenic) cancer; (b) current (at any point during their treatment) or previous treatment (within the last 5 years) with definitive radiotherapy or CRT with or without a stoma placed; (c) adults (lowest age of inclusion defined by local institution), and (d) informed written consent given in accordance with national/local regulations and procedures. Participants involved in previous phases of the QLQ-ANL27 development were not eligible.

To ensure good representation of participants covering the full range of the intended population (patients treated for anal cancer at different points of time), we conducted purposive sampling to recruit comparable numbers of participants in the following groups: Acute: treatment started within 3 months; Early: between 3 months and 12 months since treatment began; and Late: between 1 and 5 years since the treatment start date.

Sub-groups of participants were invited to complete the QLQ-ANL27 alongside the QLQ-C30 on two occasions to test: 1) Responsiveness to change in two scenarios: “Deterioration” including participants at the start of treatment and again at least 4 weeks later during the last week of treatment and “Improvement” to include patients during the last week of treatment and 6 -12 weeks later to coincide with their routine follow-up clinic appointment; 2) Consistency in responses at two time points (1-2 weeks apart) where clinical stability is assumed (no changes in health status reported by the patient or researcher): towards the end of treatment (high symptom presentation and burden) and, in the Late treatment group (low levels of symptoms and high functioning).

Using Fayers and Machin’s (2016) rule of thumb of 10-15 responses required per question to adequately test a measure’s properties [20], the overall target sample size was set as 375. For the double administration groups, a target of at least 40 patients in each of the responsiveness to

change groups was set to allow detection of a change of 0.5 standard deviations (SDs), power 80% and alpha level 5%, assuming a correlation of 0.4 between the repeated measurements. To test consistency in responses, a minimum of 40 patients was required overall (combining the responses of the acute and late group participants) to provide an estimated intraclass correlation (ICC) of 0.7 with a 95% confidence interval of width 0.3 [21].

The QLQ-ANL27 includes 27 questions with six multi-item scales hypothesized prior to the study. Symptom scales (higher scores representing a higher level of symptoms) include four multi-item scales: bowel symptoms (non-stoma), bowel symptoms (stoma), stoma care, pain/discomfort and 5 single items (frequent urination, keeping clean, proximity to toilet, lower limb edema, planning activities). Functional scales (with higher scores representing better functioning) include two multi-item scales: sexual functioning (male) and sexual functioning (female). There is also one sexual activity screening question which does not form part of the scoring and is also included for descriptive purposes. The sexual functioning items are rated according to functioning during the past 4 weeks; the timeframe for all other questions is the past week. Except for the QLQ-C30 overall health and HRQoL questions, a 4-point Likert scale ranging from “not at all” (1) to “very much” (4) is used. A linear transformation is applied to produce scale scores with a possible range from 0 to 100, with high scores indicating better functioning on the functional and Global health/HRQoL scales but worse symptoms on the symptom scales [22]. In terms of handling missing data, scales scores were calculated if at least half of its items were completed with the scoring algorithm applied to the number of items completed (i.e., if a scale is made up of 5 items and at least 3 are answered, the scores for the 3 items are summed and divided by 3 to calculate a raw score which is then standardised), otherwise the scale score was regarded as missing.

Questionnaire translations were prepared according to the EORTC QLQ guidelines [19] using an iterative forward-backward procedure.

Socio-demographic (Supplementary Table 1) and clinical data (TNM staging, diagnosis date, recurrence, treatment type and dates, and performance status using the Eastern Cooperative Oncology Group (ECOG) [23] see Supplementary Table 2) were retrieved from medical notes or collected directly from participants either before they completed the questionnaires or when they were contacted (in person or over the telephone) to complete the debriefing questions. For those completing the QLQ-C30 and QLQ-ANL27 on a second occasion, ECOG performance status, hospital admissions, post-treatment complications, newly diagnosed medical conditions and new treatments were recorded by the researcher in consultation with the patient.

A debriefing questionnaire collected participant feedback including general comments, the time required to complete the measure, help received, and items identified as upsetting, confusing or difficult to answer. Reasons for missing data were documented. Participants had the option of completing the questionnaire on paper or online using the Computer-based Health Evaluation System (CHES) [24].

Descriptive analyses were carried out on the time taken to complete the QLQ-ANL27, assistance required and item feedback. Missing data by timepoint and response distribution for each item were summarised. Items were flagged if the combined proportion of patients in the two extreme response categories did not exceed 10%.

Statistical analysis

The main psychometric analyses were performed on the baseline responses of participants regardless of their subgroup classification using SAS version 9.4 [25].

Confirmatory factor analysis (CFA) was planned to test the hypothesised six multi-item scale using robust weighted least squares estimation. CFA was carried out for males and females separately, therefore each testing a 5-factor model (bowel symptoms – non stoma and stoma separately, pain/discomfort, stoma care, and either sexual functioning male or female). The model fit was assessed using local fit testing (residuals < 0.10) [26] and global fit statistics (chi-square > 0.05 [27], comparative fit index (CFI) ≥ 0.95 [28], root mean square error approximation (RMSEA) [26] and standardised root mean square residual (SRMR) < 0.10) [26] as a guide.

To test the scale structure and investigate whether the questions are measuring the same construct, inter-item correlations were computed using polychoric correlation coefficients between each pair of items at baseline. Items with high correlation ($r \geq 0.90$; indicating over 80% shared variance) were highlighted for examination.

Consistency in responses to the different items within the scales (internal reliability) was examined using Cronbach's alpha coefficient, with a value of ≥ 0.70 regarded as acceptable [29]. Test-retest reliability assessed consistency in responses on two repeated assessments using intra-class coefficients (ICCs) on the responses of participants in the clinical stable sub-group at assessment 1 and 2 with an ICC ≥ 0.70 regarded as adequate [20].

Item-level convergent validity was defined as an item correlating highly with its own hypothesised scale, defined as a correlation of $r \geq 0.40$ (corrected for overlap) [20]. For discriminant validity, scaling successes were defined as correlations between an item and its hypothesized scale (corrected for overlap) higher than its correlation with other scales.

Pearson's product moment correlations between the QLQ-ANL27 and the QLQ-C30 were used to assess the extent to which their scales are measuring similar constructs. Scales that are conceptually related should correlate substantially with one another ($r \geq 0.40$) [20]. The QLQ-ANL27 bowel symptoms scale was hypothesised to be conceptually related to the QLQ-C30 constipation, diarrhoea, physical functioning, and social functioning scales. QLQ-ANL27 pain/discomfort was hypothesised to conceptually relate to the QLQ-C30 pain scale.

Known-group comparisons assessed the extent to which the EORTC QLQ-ANL27 differentiated between the following groups: (1) Treatment group (acute vs. early vs. late); (2) Treatment status (on vs. off); (3) Disease recurrence (yes vs. no); (4) Site (localised vs. locoregional); (5) Radiotherapy schedule (3-D conformal vs. intensity modulated radiotherapy (IMRT)); (6) Stoma (yes vs. no). The following cut-offs were used to interpret the magnitude of each effect size: small (0.20), moderate (0.50), and large (0.80) [29]. An analysis of variance was used to determine the statistical significance ($p \leq 0.05$) between groups. Analysis was performed only for classifications that included at least 20 participants in each category.

To assess sensitivity to change over time, the mean change scores from initial baseline to follow up timepoint were summarised and compared within groups using effects sizes [29] using the criteria for interpretation outlined above.

Results

Participants

A total of 387 participants were enrolled in the study over 35 months; 5 from Spain only completed one questionnaire at baseline (QLQ-C30 or QLQ-ANL27) and were not included in the analyses.

The mean (standard deviation) age of participants was 63.9 (10.4) years with a range between 34-94 years. The sample included 273 (72%) females (reflecting population incidence [30]) and 109 (28%) males. There was a good distribution of participants recruited according to geographical region with 101 (26%) participants from Northern Europe, 126 (33%) from Western Europe, 86 (23%) from Southern Europe and 69 (18%) from outside of Europe (Supplementary Table 1).

One hundred and thirty (34%) participants were in the acute treatment group, 65 of whom were currently on treatment, 70 (18%) were in the early, and 182 (48%) in the late treatment subgroup. Half of the participants ($n=192$) presented with localised disease, 135 (35%) had locoregional disease. At study entry, 15 participants presented with disease recurrence (8 localised and 7 locoregional). Only 34 (9%) participants had a stoma. Most participants, 297 (78%), received IMRT (2 received Volumetric Arc Therapy), 80 (21%) 3-D conformal radiotherapy. Nearly three quarters of

the participants received either concomitant mitomycin c (MMC) and capecitabine (n=153, 40.0%) or MMC and fluorouracil (5-FU) (n=133, 35%). Over half of the participants (n=249, 65%) described their performance status as “Fully active”. At least one other health condition was reported by 258 (68%) participants. For a full overview of the clinical characteristics of the sample, see Supplementary Table 2.

Questionnaires were completed on two occasions by 131 (34%) participants; 84 were assigned to the test-retest population (n=16 on treatment and n=68 post-treatment) and 47 to the responsiveness to change sub-groups (n=34 deterioration, n=13 improvement). Given the small number of participants in the improvement group, analyses were not performed.

Completion of the QLQ-ANL27

It took participants a mean 8 (SD 5.8) minutes to complete the QLQ-ANL27; ranging between 1 and 40 minutes. Assistance to complete the QLQ-ANL27 was received by 52 (14%) participants; for 28 participants, the questions were read out by the researcher or family member, 20 asked for help with interpreting some of the questions, 3 asked for a family member’s opinion (e.g., to facilitate memory), and for one the nature of help was not specified. The debriefing questionnaire was completed by 355 (93%) participants.

Comprehensibility

Clarification on the meaning of questions was requested by 14 participants; 6 queried the meaning of stoma. A total of 74 (19%) participants nominated at least one confusing question; for the majority of these (43), sex-related questions were regarded as most problematic, largely due to the perceived irrelevance of such questions. For participants who had just started treatment, a “not applicable” response option was recommended for questions relating to soreness or skin problems around the treated area. Two participants with a stoma queried the relevance of some of the bowel symptom questions. Four participants struggled with the restrictions of the past week time frame due to their variability of symptoms. Five participants identified problematic questions which relied upon interpretations of words such as “often” and “frequently”. Finally, difficulties in assigning causality to anal cancer / treatment or another condition was described by seven participants.

Acceptability

Seventeen (4%) participants (from nine countries) reported feeling upset by at least one of the questions with the sex questions largely responsible, mentioned by 12 (3%) participants as intrusive and a reminder of the effects of anal cancer and its treatment. Twenty-eight participants wrote about their positive experience of completing the QLQ-ANL27 and described it as “helpful” and

“relevant”; one participant wrote that it prompted her to talk about issues which she might not have been able to bring up herself.

Comprehensiveness

Sixty-six (17%) participants proposed the following areas of concern which were not covered by the questionnaires: psychological aspects of anal cancer, including body image and coping (18, 5% participants), information needs and provision (12, 3%), urinary function issues such as incontinence, and painful and night-time urination (7, 2%), problems around the pelvic region (6, 1.6%), and neuropathy (4, 1%). Seven (2%) participants recommended asking more detailed questions about bowel problems and sexual functioning given their significant impact on their lives.

Missing data

Missing responses to individual QLQ-ANL27 items ranged from 1% to 20%. Items with 5% or higher missing data (calculated from the total number of participants eligible to answer the questions) at both timepoints included all sexual functioning (except for the sexual activity screening question) and stoma questions. Pain during intercourse had the highest percentage of missing responses across both timepoints (19% and 20%) followed by the female-specific sexual questions (ranging from 11% and 13% across the timepoints).

Changes to hypothesized scale structure

CFA tested a 3-factor model: bowel symptoms (non-stoma), pain/discomfort and either sexual functioning female or male. Two factors (bowel symptoms (stoma) and stoma care) were excluded from the originally planned model due to low sample size for stoma patients (n=34). The sample size for the male subgroup was small for CFA resulting in low power for this model (n=109). Both models did not fit the data well (chi-square <0.0001, CFI <0.95) (see Supplementary Table 3). Other analyses (not reported) also suggested these sexual functioning multi-item scales were not performing well psychometrically and resulted in a revised scale structure with three single items item for interest in sex, sex life and pain during intercourse, one single item for males (erectile problems) and a 3-item scale for females (vaginal symptoms). From here on, the analyses reported relate to the psychometric properties for this revised scale structure.

Item distributions

All item response options were used (Supplementary material 4). For all QLQ-ANL27 items, >10% participants endorsed the “not at all” or “very much” response options at baseline. In terms of the sexual activity screening question, 71% of the population reported not being sexually active at study entry (73% at follow-up) and, at baseline, 41.9% participants reported that anal cancer and its treatment had affected their sex life “quite a bit” or “very much” while 39% of males reported “quite

a bit” / “very much” difficulty achieving or maintaining an erection. Bowel function issues relating to bowel urgency (45.4%) and frequency (36%) were endorsed as “quite a bit” or “very much” at the first assessment.

At follow-up, skewed responses towards the “not at all” / “a little” options were identified for swelling in the legs or ankles (91.9%) and two of the stoma questions (92.3% for sore skin around stoma and leakage of stools).

Scale structure

Item pair correlations (Supplementary Table 5) reached the threshold for concern ($r \geq 0.90$) for only one pair (difficulty getting or maintaining an erection and unintentional release of gas/flatulence from stoma) which was only answered by 12 participants due to the conditional nature of the questions and missing responses.

Reliability

Consistency in responses to items within scales - Internal reliability

The Cronbach’s alpha for all multi-item scales, except for stoma care (0.65 at baseline), was acceptable (ranging from 0.71 to 0.93) at both time points (Table 1).

Consistency in responses to items over time - Test-retest reliability

For most items and scales, there was acceptable consistency in responses over time with the ICCs above the threshold of 0.7 for all scales except for stoma care (0.59) and all single items except for keeping clean (0.68) and proximity to toilet (0.69) (Table 2). Paired mean differences were generally small, ranging from 0 to 4.76 points (stoma care).

Validity

Convergent and divergent validity

All items, except for unintentional release of gas/flatulence from the stoma bag, correlated higher with its own scale items compared with items from other scales (Table 3).

The anticipated scale correlations between the QLQ-C30 and QLQ-ANL27 were broadly met for the bowel symptoms (non-stoma) and pain/discomfort scales (see Table 4). The correlation between bowel symptoms (non-stoma) and QLQ-C30 constipation was lower than anticipated (0.21) and borderline for QLQ-C30 physical functioning (0.36). Pain/discomfort scales correlated (0.70) with the QLQ-C30 pain scale. The bowel symptoms (stoma) scale did not correlate with any of its hypothesized scales.

Known group comparisons

Most scales could distinguish across treatment phases well, particularly pain/discomfort and vaginal symptoms with large and moderate effect sizes respectively (see Table 5). All scales, except vaginal

symptoms had small to moderate effect sizes for performance status. Vaginal symptoms distinguished between radiotherapy type with a small effect size (0.20). The pain/discomfort scale also distinguished between patients according to whether they were on treatment with a small effect size (0.20). None of the scales could distinguish participants according to cancer stage (localised vs locoregional).

Responsiveness to change

A large deterioration over time (from the beginning to the end of treatment) was observed for pain/discomfort (mean difference -34.9, $p < 0.0001$, effect size -1.3) and vaginal symptom scales (mean difference -31.7, $p = 0.001$, effect size 0.9). For the bowel (non-stoma) scale, a moderate effect size (-0.7) was observed (mean difference -17.9, $p = 0.001$) (Table 6).

Discussion

This international field study has tested the psychometric properties and acceptability of different language versions of the first ever anal cancer specific HRQoL questionnaire. The QLQ-ANL27 was feasible; most participants completed the questionnaire independently.

The results from the psychometric testing were favorable in terms of supporting the items included in the QLQ-ANL27 and provided evidence for acceptable reliability and validity for most scales. The full range of response options was used for the questions and the HRQoL concerns assessed were relevant across the different disease and treatment phases covered in this study. Consistent with reviews on HRQoL issues of people with anal cancer [4, 31], bowel (in particular, urgency) and sex-related problems were the most prominent across all treatment groups. Due to the low number of participants with a stoma, reflective of the treatment of this patient group [1], two scales applicable only to stoma patients (bowel symptoms and stoma care) remain largely untested. However, these items were taken from the EORTC QLG colorectal questionnaire (QLQ-CR29) [17] and have been fully validated in the colorectal cancer context. Psychometric testing supported the structure of the bowel symptoms (non-stoma) scale as well as a single pain/discomfort scale. However, problems with the sexual functioning items were highlighted. A modified scale structure was proposed with the sexual functioning scales split into three single items applicable to all patients (sex life, interest in sex and pain during intercourse), a single item (erectile problems) for men and a three-item vaginal symptoms scale for females which demonstrated favorable psychometric properties. Participant feedback revealed that where upsetting questions were identified, the sex questions were largely responsible; however given the widespread and potentially long-lasting impact of anal cancer and its treatment on sexual functioning it is considered important to retain these items [4, 8, 12, 14, 32]. It also follows that problems in this area will not be addressed if they are not asked.

To address the issue of missing responses for the sex questions and avoid confusion as to why certain questions are asked even when sexual activity is absent, we have revised the question order of the general (not gender-specific) sex questions to start by asking respondents to rate the degree to which sexual activity has been affected by anal cancer and its treatment, followed by the extent to which they are interested in sex. The sexual activity screening question is then asked and serves as a conditional question for the painful intercourse item which now includes a not applicable option.

There was also confusion around the meaning of the word stoma and navigation to the appropriate sub-section of the questionnaire according to stoma status. These conditional questions also had a high number of missing responses. To address these problems, a screening question is now included which uses the validated wording of the QLQ-CR29.

In terms of omissions, participants identified areas of concern such as the psychological impact of anal cancer which, although widely reported in previous research [33, 34], are not specific to anal cancer. Three other areas of concern were also identified as not addressed by 1-2% of the sample. First, pain in the pelvic region which could be argued as non-specific in terms of its origin as well as being already covered by the pain or discomfort questions. Second, neuropathy which although is recognised as a chemotherapy-related toxicity [35] and can be radiation induced [36], is not commonly observed in this patient group. Finally, urinary issues such as incontinence, painful, or night-time urination were mentioned as missing. These were previously reported in the literature [12, 37] and recognised in the context of earlier questionnaire development work but excluded following pilot testing due to poor performance and low incidence [7].

The high rates of toxicity and morbidity both during and after CRT [5-9] call for the development and use of a validated HRQoL instrument that is specific for anal cancer patients. Sunesen et al., in consultation with clinicians and existing grading systems, developed an anal cancer-specific questionnaire [8] for the purposes of their own research but this has not been validated. While some of the issues measured by the QLQ-ANL27 feature in existing measures, they are limited in terms of their coverage. For example, while skin toxicity is assessed by the colorectal cancer questionnaire, the QLQ-CR29, with the inclusion of a question about sore skin, it does not capture the widespread impact of skin reactions on activities of daily living, such as ability to sit down and comfort when lying down. In addition, while painful sexual intercourse is evaluated in existing measures (i.e., the QLQ-CR29), the impact of treatment on the vaginal area has been overlooked. The QLQ-ANL27 is the first anal cancer specific HRQoL questionnaire to provide a comprehensive assessment of the impact of the full range of toxicities related to anal cancer and its treatment with CRT. This tool is

already being used in key clinical trials and gives an insight to the patient's HRQoL issues both in the acute and late effect settings and will help in their overall management and support.

The QLQ-ANL27 has five multi-item scales: bowel symptoms (non-stoma), bowel symptoms (stoma), stoma care, pain/discomfort, vaginal symptoms and nine single items (frequent urination, keeping clean, proximity to toilet, lower limb edema, planning activities, sexual interest, sex life, painful sexual intercourse, and erectile problems). Responses to the sexual interest question are reversed scored so that, for all questions, a higher score reflects poorer functioning or a high incidence of symptoms. There are also two screening questions: presence of a stoma and sexual activity which do not contribute to the scale scoring.

Limitations

Although the overall target sample size was exceeded, the number of participants completing a second administration for the purposes of testing sensitivity to change (improvement) over time was limited and did not allow for an adequately powered analysis. Although we provided evidence of consistency in responses of clinically stable participants, we relied mostly on data from participants in the late treatment group rather than those with a high incidence of symptoms (in the acute treatment group). In addition, the number of participants with a stoma was insufficient for a rigorous evaluation of the stoma-related items and scales. Stoma status and disease recurrence were both excluded from the known group analyses due to small numbers. For these reasons, and due to the revised scale structure, the QLQ-ANL27 would benefit from further validation studies. The current study represents an important first step towards demonstrating the psychometric properties of the QLQ-ANL27.

Over half of the participants had been diagnosed with a comorbidity and 16 had a previous cancer diagnosis. Although the QLQ-ANL27 instructs respondents to answer each question with reference to anal cancer, it is acknowledged that attribution of causality can be challenging. Furthermore, additional sub-group analyses according to HIV positive status and sexual orientation were beyond the scope of the current study but could further support the validity of the QLQ-ANL27.

Although this manuscript presents the QLQ-ANL27 as a comprehensive measure, the extent to which it is sensitive to the HRQoL issues experienced by people with anal cancer who follow different treatment pathways to include surgery as well as those with metastatic disease is unclear; this will be a focus of future research.

Conclusion

The QLQ-ANL27 is a HRQoL instrument for assessing the acute and long-term sequelae of anal cancer treated with CRT. The QLQ-ANL27 is usable, reliable, valid, acceptable across different geographical regions and has been translated into 16 languages. The measure has already been widely adopted within the scientific community with its inclusion in clinical trials [e.g., 38, 39] and several of the issues it covers are recommended in a core outcome set for anal cancer [32]. The QLQ-ANL27 is available upon request from the EORTC QLG from <https://qol.eortc.org>.

References

1. National Cancer Institute Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Anal Cancer, 2020. <https://seer.cancer.gov/statfacts/html/anus.html>. [accessed 18 March 2021].
2. Islami F, Ferlay J, Lortet-Tieulent J, Bray F, Jemal A. International trends in anal cancer incidence rates. *Int J Epidemiol* 2016;46(3):924-38.
3. Rao S, Guren MG, Khan K, Brown G, Renehan AG, Steigen SE et al. Anal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2021;32(9):1087-100.
4. Sodergren SC, Vassiliou V, Dennis K, Tomaszewski KA, Gilbert A, Glynne-Jones R et al. Systematic review of the quality of life issues associated with anal cancer and its treatment with radiochemotherapy. *Support Care Cancer* 2015;23(12):3613-23.
5. Allal AS, Sprangers MA, Laurencet F, Reymond MA, Kurtz JM. Assessment of long-term quality of life in patients with anal carcinomas treated by radiotherapy with or without chemotherapy. *Br J Cancer* 1999;80(10):1588-94.
6. Das P, Cantor SB, Parker CL, Zampieri JB, Baschnagel A, Eng C et al. Long-term quality of life after radiotherapy for the treatment of anal cancer. *Cancer* 2010;116(4):822-9.
7. Sodergren SC, Johnson CD, Gilbert A, Tomaszewski KA, Chu W, Chung HT et al. Phase I-III development of the EORTC QLQ-ANL27, a health-related quality of life questionnaire for anal cancer. *Radiother Oncol* 2018;126(2):222-8.
8. Sunesen KG, Nørgaard M, Lundby L, Havsteen H, Buntzen S, Thorlacius-Ussing O, Laurberg S. Long-term anorectal, urinary and sexual dysfunction causing distress after radiotherapy for anal cancer: a Danish multicentre cross-sectional questionnaire study. *Colorectal Dis* 2015;17(11):O230-9.
9. Knowles G, Haigh R, McLean C, Phillips H. Late effects and quality of life after chemo-radiation for the treatment of anal cancer. *Eur J Oncol Nurs* 2015;19(5):479-85.
10. Nigro ND, Vaitkevicius VK, Considine B Jr. Combined therapy for cancer of the anal canal: a preliminary report. *Dis Colon Rectum* 1974;17(3):354-6.
11. Näsval P, Dahlstrand U, Löwenmark T, Rutegård J, Gunnarsson U, Strigård K. Quality of life in patients with a permanent stoma after rectal cancer surgery. *Qual Life Res* 2017;26(1):55-64.
12. Bentzen AG, Balteskard L, Wanderås EH, Frykholm G, Wilsgaard T, Dahl O, Guren MG. Impaired health-related quality of life after chemoradiotherapy for anal cancer: late effects in a national cohort of 128 survivors. *Acta Oncologica* 2013;52(4):736-744.

13. Fakhrian K, Sauer T, Dinkel A, Klemm S, Schuster T, Molls M, Geinitz H. Chronic adverse events and quality of life after radiochemotherapy in anal cancer patients. A single institution experience and review of the literature. *Strahlenther Onkol* 2013; 189(6):486-94.
14. Gilbert A, Drinkwater K, McParland L, Adams R, Glynne-Jones R, Harrison M et al. UK national cohort of anal cancer treated with intensity-modulated radiotherapy: One-year oncological and patient-reported outcomes. *Eur J Cancer* 2020; 128:7-16.
15. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993; 85(5):365-76.
16. Sprangers MA, te Velde A, Aaronson NK. The construction and testing of the EORTC colorectal cancer-specific quality of life questionnaire module (QLQ-CR38). *European Organization for Research and Treatment of Cancer Study Group on Quality of Life. Eur J Cancer* 1999;35(2):238-47.
17. Gujral S, Conroy T, Fleissner C, Sezer O, King PM, Avery KN, Sylvester P, Koller M, Sprangers MA, Blazeby JM; European Organisation for Research and Treatment of Cancer Quality of Life Group. Assessing quality of life in patients with colorectal cancer: an update of the EORTC quality of life questionnaire. *Eur J Cancer* 2007;43(10): 1564-73.
18. Ward WL, Hahn EA, Mo F, Hernandez L, Tulskey DS, Cella D. Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. *Qual Life Res* 1999; 8(3):181-95.
19. Johnson C, Aaronson N, Blazeby JM, Bottomley A, Fayers P, Koller M et al. *Guidelines for developing questionnaire modules*. On Behalf of EORTC Quality of Life Group, 2011. <https://qol.eortc.org/manual/developing-questionnaire-modules/> [accessed 24 April 2021].
20. Fayers PM, Machin D. *Quality of Life: the Assessment, Analysis and Interpretation of Patient-reported Outcomes*. 3rd ed. Chichester: John Wiley & Sons, Ltd; 2016.
21. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res* 2004;13(4):251-71.
22. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. *The EORTC QLQ-C30 Scoring Manual* 3rd ed. European Organisation for Research and Treatment of Cancer: Brussels; 2001.
23. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982; 5(6):649-55.
24. Holzner B, Giesinger JM, Pinggera J, Zugal S, Schöpf F, Oberguggenberger AS et al. The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. *BMC Med Inform Decis Mak* 2012;12:126.
25. SAS Institute Inc. 2013. *SAS® 9.4 Statements: Reference*. 2013, SAS Institute Inc.: Cary, NC.
26. Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York: Guilford Press; 2015.
27. Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 2001; 66(4):507-14.
28. Hu L-T, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6(1):1-55.
29. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27(3 Suppl):S178-89.
30. Cancer Research UK *Anal cancer incidence statistics*. 2021 <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/anal-cancer/incidence>. [accessed 18 March 2021].
31. Sterner A, Derwinger K, Staff C, Nilsson H, Angenete E. Quality of life in patients treated for anal carcinoma-a systematic literature review. *Int J Colorectal Dis*. 2019; 34(9):1517-28.

32. Fish R, Sanders C, Adams R, Brewer J, Brookes ST, DeNardo J et al. A core outcome set for clinical trials of chemoradiotherapy interventions for anal cancer (CORMAC): a patient and health-care professional consensus. *Lancet Gastroenterol Hepatol* 2018;3(12): p. 865-73.
33. Benedict C, Philip EJ, Baser RE, Carter J, Schuler TA, Jandorf L et al. Body image and sexual function in women after treatment for anal and rectal cancer. *Psycho-Oncology* 2016;25(3):316-23.
34. Li M, Fitzgerald P, Rodin G. Evidence-based treatment of depression in patients with cancer. *J Clin Oncol.* 2012;30(11):1187-96.
35. Mols F, Beijers T, Vreugdenhil G, van de Poll-Franse L. Chemotherapy-induced peripheral neuropathy and its association with quality of life: a systematic review. *Support Care Cancer* 2014;22(8):2261-9.
36. Edwards HL, Mulvey MR, Bennett MI. Cancer-Related Neuropathic Pain. *Cancers (Basel)* 2019;11(3):373.
37. Joseph K, Vos LJ, Warkentin H, Paulson K, Polkosnik LA, Usmani N et al. Patient reported quality of life after helical IMRT based concurrent chemoradiation of locally advanced anal cancer. *Radiother Oncol* 2016;120(2):228-33.
38. ClinicalTrials.gov. A Study of INCMGA00012 in Squamous Carcinoma of the Anal Canal Following Platinum-Based Chemotherapy (POD1UM-202). <https://clinicaltrials.gov/ct2/show/NCT03597295>. [accessed 21 May 2021].
39. ClinicalTrials.gov. Carboplatin-paclitaxel With Retifanlimab or Placebo in Participants With Locally Advanced or Metastatic Squamous Cell Anal Carcinoma (POD1UM-303/InterAACT 2). <https://clinicaltrials.gov/ct2/show/NCT04472429>. [accessed 21 May 2021].