Title

Construct validity and responsiveness of a health-related symptom index for persons either treated or monitored for anal high-grade squamous intraepithelial lesions (HSIL): AMC-A01/-A03.

Journal

Quality of Life Research, 32(8)

Authors

Atkinson, Thomas
Lensing, Shelly
Lee, Jeannette
et al.

Publication Date

2023-08-01

DOI

10.1007/s11136-023-03391-4

Peer reviewed

HHS Public Access

Author manuscript

Qual Life Res. Author manuscript; available in PMC 2023 August 01.

Published in final edited form as:

Qual Life Res. 2023 August ; 32(8): 2293–2304. doi:10.1007/s11136-023-03391-4.

Construct validity and responsiveness of a health-related symptom index for persons either treated or monitored for anal high-grade squamous intraepithelial lesions (HSIL): AMC-A01/-A03

Thomas M. Atkinson[1], **Shelly Lensing**[2], **Jeannette Y. Lee**[2], **Di Chang**[2], **Soo Young Kim**[1], **Yuelin Li**[1], **Kathleen A. Lynch**[1,3], **Andrew Webb**[1], **Susan M. Holland**[1], **Erica I. Lubetkin**[4], **Stephen Goldstone**[5], **Mark H. Einstein**[6], **Elizabeth A. Stier**[7], **Dorothy J. Wiley**[8], **Ronald Mitsuyasu**[9], **Isabella Rosa-Cunha**[10], **David M. Aboulafia**[11], **Shireesha Dhanireddy**[12], **Jeffrey T. Schouten**[11,12], **Rebecca Levine**[13], **Edward Gardner**[14], **Jeffrey Logan**[14], **Hillary Dunleavy**[15], **Luis F. Barroso**[16], **Gary Bucher**[17], **Jessica Korman**[18], **Benjamin Stearn**[19], **Timothy J. Wilkin**[20], **Grant Ellsworth**[20], **Julia C. Pugliese**[21], **Abigail Arons**[21], **Jack E. Burkhalter**[1], **David Cella**[22], **J. Michael Berry-Lawhorn**[23], **Joel M. Palefsky**[23]

[1]Department of Psychiatry & Behavioral Sciences, Memorial Sloan Kettering Cancer Center, 633 Third Ave, New York, NY 10017, USA

[2]University of Arkansas for Medical Sciences, Little Rock, AR, USA

[3]School of Global Public Health, New York University, New York, NY, USA

[4]CUNY School of Medicine, New York, NY, USA

[5]Laser Surgery Care, New York, NY, USA

[6]Rutgers-New Jersey Medical School, Newark, NJ, USA

[7]Boston Medical Center, Boston, MA, USA

[8]University of California, Los Angeles School of Nursing, Los Angeles, CA, USA

[9]UCLA Center for Clinical AIDS Research and Education, Los Angeles, CA, USA

[10]Jackson Memorial Hospital, Miami, FL, USA

[11]Virginia Mason Medical Center, Seattle, WA, USA

✉Thomas M. Atkinson, atkinsot@mskcc.org.

Author contributions Per International Committee of Medical Journal Editors (ICMJE) guidelines, all authors made substantial contributions to the conception or design of this manuscript; or the acquisition, analysis, or interpretation of data for the work; and drafted and/or revised the manuscript critically for important intellectual content; and reviewed and approved the final manuscript; and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Data analysis was performed by TMA, SL, JYL, DC, SYK, and YL. The first draft of the manuscript was written by TMA and all authors commented on and/or approved each version of the manuscript.

Competing interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the NCI's Cancer Therapy Evaluation Program and Institutional Review Boards at each study site.

Consent to participate For the construct validity phase, verbal consent was obtained from all included participants. Informed consent was obtained from all participants in the responsiveness phase.

[12]Harborview Medical Center, Seattle, WA, USA

[13]Montefiore Medical Center, New York, NY, USA

[14]Public Health Institute at Denver Health, Denver, CO, USA

[15]University of Colorado Hospital, Aurora, CO, USA

[16]Atrium Health Wake Forest Baptist, Winston-Salem, NC, USA

[17]Anal Dysplasia Clinic MidWest, Chicago, IL, USA

[18]Metropolitan Gastroenterology Group, Washington, DC, USA

[19]Dupont Circle Physicians Group PC, Washington, DC, USA

[20]Weill Cornell Medicine, New York, NY, USA

[21]ANCHOR Data Management Center of The Emmes Company, LLC, Rockville, MD, USA

[22]Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[23]University of California, San Francisco, CA, USA

Abstract

Purpose—To determine whether treatment of anal high-grade squamous intraepithelial lesions (HSIL), vs active monitoring, is effective in reducing incidence of anal cancer in persons living with HIV, the US National Cancer Institute funded the Phase III ANal Cancer/HSIL Outcomes Research (ANCHOR) clinical trial. As no established patient-reported outcomes (PRO) tool exists for persons with anal HSIL, we sought to estimate the construct validity and responsiveness of the ANCHOR Health-Related Symptom Index (A-HRSI).

Methods—The construct validity phase enrolled ANCHOR participants who were within two weeks of randomization to complete A-HRSI and legacy PRO questionnaires at a single time point. The responsiveness phase enrolled a separate cohort of ANCHOR participants who were not yet randomized to complete A-HRSI at three time points: prior to randomization (T1), 14–70 (T2), and 71–112 (T3) days following randomization.

Results—Confirmatory factor analysis techniques established a three-factor model (i.e., physical symptoms, impact on physical functioning, impact on psychological functioning), with moderate evidence of convergent validity and strong evidence of discriminant validity in the construct validity phase ($n = 303$). We observed a significant moderate effect for changes in A-HRSI impact on physical functioning (standardized response mean = 0.52) and psychological symptoms (standardized response mean = 0.60) from T2 ($n = 86$) to T3 ($n = 92$), providing evidence of responsiveness.

Conclusion—A-HRSI is a brief PRO index that captures health-related symptoms and impacts related to anal HSIL. This instrument may have broad applicability in other contexts assessing individuals with anal HSIL, which may ultimately help improve clinical care and assist providers and patients with medical decision-making.

Plain English summary

A randomized clinical trial called ANCHOR is currently underway for persons who are living with HIV and are found to have precancerous anal lesions. The ANCHOR trial is testing whether treatment or regular observation is more effective in reducing anal cancer. Little is known about how treatment or observation will impact the symptoms or quality of life for participants in this study. Prior research with participants from the ANCHOR trial helped us to select the symptoms and areas of quality of life that they felt would be most important to include in the tool, which is called the ANCHOR Health-Related Symptom Index, or A-HRSI. Our goal in the present study was to make certain that the A-HRSI will allow these study participants to accurately report how treatment or observation was impacting their symptoms or quality of life. We gave the A-HRSI to 303 ANCHOR participants to make certain that the index is measuring what it is supposed to be measuring, that is, physical symptoms, impact on physical functioning, and impact on psychological functioning. We also had 103 ANCHOR participants complete the A-HRSI survey three times to determine whether this index could identify changes to participant's symptoms or quality of life over time. We found that A-HRSI is a valid and accurate index to allow people to report their symptoms and quality of life related to treatment or observation for precancerous anal lesions. These findings may help doctors to better understand the experiences of their patients.

Keywords

Neoplasms; HIV; Patient reported outcome measures; Quality of life; Anus neoplasms; Psychometrics

Introduction

Anal cancer is a growing problem in the United States (US), particularly for people living with HIV (PLWH) [1, 2]. The incidence of anal cancer in the US general population was 1.9/100,000 among men and women from 2013 to 2017 [3]. Anal cancer prevention efforts have focused on two approaches: (1) vaccination against human papillomavirus (HPV), the underlying causative agent of anal cancer [4, 5], for men and women through age 26 [6], and (2) screening for and treating precancerous HPV-associated high-grade squamous intraepithelial lesions (HSIL) prior to progression to cancer for those who have been already exposed to HPV [7–10].

To determine whether treatment (i.e., topical or ablative) of anal HSIL, vs active monitoring, is effective in reducing incidence of anal cancer in PLWH, the US National Cancer Institute (NCI) funded the Phase III ANal Cancer/HSIL Outcomes Research (ANCHOR) clinical trial (clinicaltrials.gov identifier: NCT02135419) in 2014. A secondary objective of ANCHOR is the capture of participant-reported health-related quality of life (HRQOL), defined as a "participant's general perception of the effect of illness and treatment on physical, psychological, and social aspects of life" [11]. Accurate documentation of HRQOL is feasible and necessary as part of clinical trials in oncology [12–14]. Patient-reported outcomes (PROs), defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" [11], that are experienced throughout the course of a trial cannot

be reliably graded by clinicians or their staff [15]. Therefore it is essential to capture this information verbatim from participants [16].

Since no appropriate PRO tool existed for the capture of HRQOL related to anal disease in PLWH, it was essential to develop a novel HRQOL index to capture the unique physical symptoms (e.g., anal pain, itching in or around the anus), impact on physical functioning (e.g., problems with sitting/participating in leisure activities), impact on psychological functioning (e.g., decreased desire/enjoyment of anal sexual activity) related to treatment or active monitoring for anal HSIL. We established the ANCHOR health-related symptom index (A-HRSI) to be consistent with the US regulatory recommendations for psychometric properties of PRO instruments [11].

Through the use of expert consultation (i.e., ANCHOR clinical investigators), concept elicitation ($n = 41$) and cognitive interviews ($n = 45$) with participants eligible for ANCHOR, we previously estimated the content validity of the 25-item A-HRSI [17], as well as the test–retest reliability [18]. The present study sought to estimate construct validity and responsiveness of A-HRSI in ANCHOR participants. Additionally, we aimed to estimate the acceptability of administering A-HRSI via an electronic PRO (ePRO) system. It was anticipated that the best fitting factor structure for A-HRSI would be a three-factor model (i.e., physical symptoms, impact on physical functioning, and impact on psychological functioning) and that A-HRSI would be sensitive to changes (i.e., change for the better or change for the worse) in participant-reported performance status (i.e., standardized mean change of at least 0.46) across multiple assessment time points.

Methods

Participants

Participants were recruited from 17 US ANCHOR sites. The construct validity phase utilized a single time point observational design, whereas the responsiveness phase employed the use of a prospective cohort design. For the construct validity phase, we enrolled English-speaking ANCHOR participants who were within two-weeks of randomization. For the responsiveness phase, we recruited English-speaking participants who were enrolled into ANCHOR but who had not yet been randomized. All data collection was centrally coordinated through Memorial Sloan Kettering Cancer Center (MSK), with responsiveness phase data collected through the ANCHOR Data Management Center (Emmes Company, LLC). Both phases were reviewed and approved by the NCI's Cancer Therapy Evaluation Program (CTEP) and Institutional Review Boards at each study site. For the construct validity phase, verbal consent was obtained from all included participants. Informed consent was obtained from all participants in the responsiveness phase.

Measures

ANCHOR Health-Related Symptom Index (A-HRSI)—(all assessments) [17, 18]—A-HRSI is a 25-item HRQOL index that assesses physical symptoms (nine items), impact on physical functioning (seven items), and impact on psychological functioning (nine items) over the past 7 days via a numeric rating scale (i.e., 0 = not at all; 1 = a little bit; 2 =

somewhat; 3 = quite a bit; 4 = very much). Domain scores are derived through calculating the mean of the completed items (i.e., domain ranges 0–4), provided that at least 50% of items within that domain have been completed. Higher domain scores indicate worse experience of symptom or impact burden.

Participant-Reported version of the Eastern Cooperative Oncology Group Performance Status (Participant-Reported ECOG PS)—(all assessments) [19]—The Participant-Reported ECOG PS measure was adapted from clinician-to-participant language through focus groups, interviews and comparisons of clinician and participant responses. The measure consists of a single item that asks participants to rate their current performance status from 0 to 4 (i.e., 0 = fully active, 1 = difficulty with physically strenuous activity but ambulatory, 2 = unable to work but in bed < 50% of the time, 3 = limited self-care and in bed > 50% of the time, 4 = completely disabled).

Functional Assessment of Cancer Therapy–General (FACT-G)—(construct validity phase only) [20]—The FACT-G is a well-established and psychometrically sound measure of HRQOL. This measure includes 27 items that result in an overall HRQOL (FACT-G) score, as well as separate indices for Physical Well-Being (7-items), Social/Family Well-Being (7-items), Emotional Well-Being (6-items), and Functional Well-Being (7-items). Higher total and subscale scores are indicative of better HRQOL.

M.D. Anderson Symptom Inventory (MDASI)—(construct validity phase only) [21]—MDASI is a widely used and well-validated measure of symptom severity and interference with everyday function. This tool consists of 13 items that ask participants to rate the severity of their symptoms on a 0–10 numeric rating scale (NRS; i.e., 0 = not present, 10 = as bad as you can imagine), as well as ratings of the level of interference with six areas of everyday function using a 0–10 NRS (i.e., 0 = did not interfere, 10 = interfered completely). The mean of the items represents overall symptom severity and interference with everyday function, with higher means indicative of worse severity or symptom distress.

Participant global impression of change (PGIC)—(responsiveness follow-up assessments only) [12]—PGIC is a single item that allows participants to rate whether their overall HRQOL has changed since the last time they were assessed using a seven-point scale (i.e., −, − 2, − 1, 0, 1, 2, 3) that represented HRQOL ranging from “very much worse” to “very much better.”

Additionally, we administered a brief demographic questionnaire that captured participants’ gender identity, race, and ethnicity at time of enrollment for both phases. Figure 1 includes the study schema by phase and measures completed.

Procedure

Construct validity phase—Potentially eligible participants from ANCHOR were referred to contact the MSK Clinical Research Coordinator (CRC) via telephone to confirm eligibility. During this interaction, eligible participants were asked for verbal consent to participate in the study. Once enrolled, participants were asked to complete their one-time assessment (i.e., A-HRSI, Participant-Reported ECOG PS, FACT-G, and MDASI) during

this interaction or to schedule an alternative telephone session to complete the measures within the assessment window (i.e., within two-weeks post-randomization to treatment or active monitoring arms). Study co-authors and ANCHOR investigators selected this time point, as it was thought that participants would have experienced symptoms or impacts to their HRQOL due to treatment or assignment to active monitoring during this timeframe. Upon completion, participants were mailed a $50 United States Postal Service money order as compensation for their time.

Responsiveness phase—As part of the consenting to ANCHOR, participants were provided with information about an optional HRQOL assessment where they would complete A-HRSI and Participant-Reported ECOG PS at three time points: time of enrollment up until time of trial randomization (T1), 14–70 days following randomization (T2), and 71–112 days following randomization (T3). Participants were also asked to complete PGIC at T2 and T3. These assessment times were selected to maximize study compliance and with input from study co-authors and ANCHOR investigators based on clinical experience to represent baseline HRQOL (T1), a period during which meaningful changes to HRQOL would occur (T2), and a time during which such meaningful changes to HRQOL would potentially return to baseline (T3). Individuals enrolled to this phase of the study were asked to complete the measures via a secure ePRO interface (via AdvantageEDC^{SM}, Emmes Company, LLC) that they could either access on their own mobile device privately or in clinic during a scheduled visit, or via a telephone facilitated interview with a MSK CRC. The ePRO tool allowed the participant to complete PRO tools directly into the study data entry system and is compatible with all major internet browsers and mobile phone operating systems. As an incentive, participants were provided with $25 per completed assessment (i.e., up to $75 total) via a reloadable debit card.

Statistical analysis

Construct validity phase—Construct validity was established using confirmatory factor analysis (CFA). CFA is a technique that allows for the comparison of models at a latent factor level based on commonalities within the observed variables [22]. A baseline unitary factor model (Model 1; 25-items) was compared with a two-factor (Model 2; physical (i.e., physical symptoms and impact on physical functioning combined [16-items] and psychological [nine-items]), three-factor (Model 3; physical symptoms [nine-items], impact on physical functioning [seven-items], and impact on psychological functioning [nine-items]), and four-factor (Model 4; physical symptoms [nine-items], impact on physical functioning [seven-items], impact on psychological functioning [five-items], impact on sexual desire and enjoyment [four-items] models based on several fit indices, using maximum likelihood estimation. Root-mean-squared error of approximation (RMSEA) is a measure of the average of the residual variance and covariance; good models have RMSEA ≤ 0.08 [23]. Comparative fit index (CFI) is a metric that ranges from 0 to 1, with values ≥ 0.90 considered to be good fitting models [24]. When calculating CFA models, χ^2 values are computed and then compared between alternative models, given changes in degrees of freedom. Lower χ^2 values, given an equal number of degrees of freedom, are indicative of a better fitting model [23]. Additionally, standardized factor loadings, representing the

correlations between observed items and latent factors should be ≥ 0.40 in the selected model [25].

Convergent and discriminant validity is an indication that constructs assessed by a new PRO instrument should or should not be correlated with constructs captured by legacy instruments that have already been validated and widely used in the literature. As an additional indicator of construct validity, convergent and discriminant validity was investigated via pairwise correlations between the four sets of measure scores (i.e., A-HRSI, Participant-Reported ECOG PS, FACT-G, and MDASI). We hypothesized that A-HRSI subscales (i.e., physical symptoms, impact on physical functioning, impact on psychological functioning) would have moderate (i.e., 0.30–0.70) pairwise Pearson $|r|$ correlations with similar subscales of legacy instruments (e.g., FACT-G Physical Well-Being, FACT-G psychological well-being, MDASI severity) as evidence of convergent validity. With respect to discriminant validity, we hypothesized low (i.e., < 0.30) pairwise Pearson $|r|$ correlations between A-HRSI subscales and subscales from legacy measures that were not intended to be captured by A-HRSI (e.g., FACT-G Social/Family Well-Being). Given a lack of standard cutoff values for convergent and discriminant validity, we defined evidence of convergent validity as pairwise Pearson $|r|$ values ranging between 0.30 and 0.70, whereas discriminant validity is indicated by Pearson $|r|$ values < 0.30.

For the construct validity phase, including convergent and discriminant validity, our goal was to enroll at least 10 participants per A-HRSI item (i.e., 250 participants) [26], with an oversampling target of 300 participants. The power calculation ($N = 300$) yielded an 80% power to detect a minimal Pearson r of 0.33 at a two-sided type-I error rate of 5%. Amos Version 26 [27] was used for the CFA, with R (Version 4.0.4) [28] used for the analysis of convergent and discriminant validity.

Responsiveness phase—A responsive index will reveal a difference between patients who experience changes in symptoms and patients who experience no change. Participants at follow-up time points (i.e., T2 and T3) were categorized into two sets of three groups based on Participant-Reported ECOG PS or PGIC responses: “worsened” (a response of “very much worse,” “moderately worse,” or “a little worse” on PGIC or an increase ≥ 1 in Participant-Reported ECOG PS score), “no change” (“about the same” on PGIC or no change in Participant-Reported ECOG PS score), and “improved” (“a little better,” “moderately better,” and “very much better” or a decrease ≥ 1 in Participant-Reported ECOG PS score). The primary responsiveness analysis evaluated both sets of three groups separately using one-way analysis of variance (ANOVA).

A sample size of 90 was required to obtain 80% power if the three groups differ by a standardized change of 0.46, in a one-way ANOVA using a two-sided type-I error rate of 5%. We estimated the statistical power by running a simulation where the “worsened” group had a mean change in A-HRSI score of − 0.46 (simulated from a standard normal of mean = − 0.46 and standard deviation = 1.0), the “no change” group had a mean change of 0.0, and the “improved” group had a mean change of + 0.46. Additional statistical assumptions included a correlation of 0.35 between the assessment scores and $n = 30$ in each of the three

groups. To allow for 10% dropout, 100 participants (i.e., 50 in each study arm) were targeted for enrollment in this phase.

To examine the extent to which patients' HRQOL changes over time, the standardized response mean (SRM) was computed as the mean change in A-HRSI subscale scores divided by the standard deviation of change scores within each change category. Values greater than 0.8 were considered large and values between 0.5 and 0.8 were considered moderate [29]. SAS Version 9.4 [30] was used for the responsiveness analysis.

Acceptability of administering A-HRSI via an ePRO tool was descriptively analyzed. Acceptability was defined as > 50% of participants at each follow-up time point who preferred completing their next assessment via the ePRO tool vs facilitated interview with the MSK CRC. Additionally, time to complete A-HRSI via the ePRO tool was captured in minutes.

Results

Construct validity phase

Between February 2017 and July 2018, we screened 323 individuals for the construct validity phase, and 20 were ineligible for participation as they were no longer in the assessment window. The 303 confirmed eligible participants (Table 1) were enrolled in the construct validity phase (median age = 51.0, 75.6% cisgender male [n = 229], 21.5% cisgender female [n = 65], 2.0% transgender female [n = 6], 0.3% transgender male [n = 1], 0.7% gender non-conforming [n = 2]). Most participants were African American (n = 195, 64.4%), with 9.9% (n = 30) identifying as Hispanic or Latino. One hundred sixty participants were assigned to the treatment arm (52.8%), with the remaining 143 (47.2%) assigned to active monitoring. Most participants in this phase self-reported a Participant-Reported ECOG PS 0 or 1 (82.8%, n = 251), with 52 participants (17.2%) indicating a Participant-Reported ECOG PS score of 2–4.

According to the CFA fit indices (Table 2), Model 2 had a lower RMSEA (0.090), higher CFA (0.749), and a statistically significant change in χ^2, given the corresponding change in degrees of freedom in comparison to Model 1 ($\chi^2(1) = 42.61$, $p < 0.05$). Model 3 was statistically superior to Model 2 with respect to RMSEA (0.065), CFA (0.862), and change in χ^2, given the change in degrees of freedom ($\chi^2(2) = 146.85$, $p < 0.05$). Model 4 did not significantly improve upon these fit indices (i.e., RMSEA = 0.065; CFI = 0.870; ($\chi^2(1) = 22.75$, $p < 0.05$). As such, the parsimonious Model 3 (i.e., physical symptoms, impact on physical functioning, impact on psychological functioning) was confirmed as the best fit for the data. Figure 2 displays the path diagram and standardized factor loadings for Model 3. All standardized factor loadings were ≥ 0.40 except for the "general pain," "constipation," and "impact with work" items.

Table 3 includes the pairwise Pearson r coefficients between the A-HRSI, Participant-Reported ECOG PS, FACT-G, and MDASI. Cronbach's α ranged 0.72–0.84, 0.74–0.83, and 0.89–0.91 for the A-HRSI, FACT-G and MDASI subscales, respectively. There was moderate evidence of convergent validity, with the majority of pairwise Pearson $|r|$

coefficients falling between 0.3 and 0.7. None of the pairwise |*r*| coefficients were > 0.7. The FACT-G Social/Family Well-Being scale had low pairwise Pearson *r* coefficients between the Physical Symptoms (− 0.03, $p = 0.64$), Impact on Physical Functioning (− 0.07, $p = 0.23$), and Impact on Psychological Functioning (− 0.21, $p < 0.01$) subscales of A-HRSI, providing strong evidence of discriminant validity.

Responsiveness phase

One hundred thirty participants were enrolled into the responsiveness phase between July and October 2019. Twenty-two individuals were ineligible post-enrollment, with the primary reason being they did not have biopsy-proven anal HSIL, and five participants were enrolled but provided no A-HRSI data at T1. Therefore, analyses for this phase were based on 103 participants who completed the A-HRSI and Participant-Reported ECOG PS questionnaires at T1. Of the 103 participants in the analytic set (84.5% cisgender male [$n = 87$], 12.6% cisgender female [$n = 13$], 1.9% transgender female [$n = 2$], 1.0% gender non-conforming [$n = 1$]), participants' median age was 52.5 years at entry with a range of 35.0–74.2. Most participants were white (43.7%) or African American (38.8%); 28.1% were Hispanic or Latino. (Table 1). Fifty-five participants were assigned to active monitoring (53.4%), whereas 48 participants were assigned to the treatment arm. A total of 86 participants completed A-HRSI and Participant Reported ECOG PS within the T2 assessment window and 92 participants completed T3 assessments. Average days (SD) from randomization were 45 (29) days at T2 and 101 (23) days at T3.

Administration of A-HRSI via ePRO was acceptable. At T1, 63 (61.2%) participants indicated that they preferred their T2 assessment to be delivered via ePRO. Fifty (60.2%) participants stated at T2 that they preferred that their T3 assessment to be delivered via ePRO. Average time in minutes (SD) to complete the A-HRSI was 10.6 (6.3) minutes at T1, 8.7 (9.2) minutes at T2, and 6.4 (9.8) minutes at T3.

Most participants ($n = 92$, 89.3%) reported their Participant-Reported ECOG PS as 0 or 1 at T1, with 9 (8.7%) participants indicating their Participant-Reported ECOG PS was 2–4. From T1 to T2, 40 (47.1%) participants indicated their Participant-Reported ECOG PS did not change, with 21 (24.7%) participants expressing that their Participant-Reported ECOG PS changed for the better and 24 (28.2%) indicating that their Participant-Reported ECOG PS changed for the worse. When comparing T3 to T2, 43 (50.6%) participants reported that their Participant-Reported ECOG PS did not change, with 18 (21.2%) participants indicating that their Participant-Reported ECOG PS changed for the better and 24 (28.2%) expressing that their Participant-Reported ECOG PS had worsened (Table 4).

In reviewing PGIC scores, 30 (35.7%) participants at T2 and 35 (38.9%) participants at T3 indicated that their overall HRQOL had changed for the better. No change in HRQOL was reported by 46 (54.8%) participants at T2 and 42 (46.7%) participants at T3. Only 8 (7.9%) participants at T2 and 13 (11.9%) participants at T3 indicated that their HRQOL had changed for the worse (Table 4).

At T2, there were no significant differences in change from T1 in A-HRSI subscales across the six PGIC or Participant-Reported ECOG PS groups. Statistical comparisons found

no significant difference between PGIC change for the better and change for the worse categories. At T3, there were significant differences in change from T2 in A-HRSI Physical Symptoms, Impact on Physical Functioning, and Impact on Psychological Functioning across the three Participant-Reported ECOG PS groups (p's < 0.05, Table 5). There were no significant interactions between study arm and PGIC or Participant-Reported ECOG PS group for any subscale at any time point.

There was a significant moderate effect for changes in A-HRSI impact on physical functioning (SRM = 0.52) and impact on psychological functioning (SRM = 0.60) from T2 to T3, providing initial evidence of responsiveness. All other SRMs were small in magnitude.

Discussion

Accurate capture of PRO data is essential as part of all clinical trials in oncology [31, 32]. While prior work has explored an association between PRO symptoms and anal HSIL [33], such information has yet to be captured (1) as part of a randomized controlled trial, or (2) using a PRO tool that was rigorously validated for use in this specific population. The present study estimated the construct validity and responsiveness of A-HRSI.

As part of the construct validity phase, A-HRSI exhibited acceptable evidence of discriminant validity when compared with the well-established Participant-Reported ECOG PS, FACT-G, and MDASI. Consistent with the underlying development of item content for A-HRSI, the CFA demonstrated that a 3-factor representation of physical symptoms, impact on physical functioning, and impact on psychological functioning is the best fitting model for this 25-item tool.

Responsiveness was assessed in a cohort of participants who were administered A-HRSI at time of enrollment up until time of trial randomization (T1), 14–70 days following randomization (T2), and 71–112 days following randomization. While SRMs representing changes from T1 were small in magnitude, there was a significant moderate effect for changes in A-HRSI impact on physical functioning and impact on psychological functioning, providing initial evidence of responsiveness. Additionally, we estimated that administration of A-HRSI via ePRO is acceptable and is minimally burdensome to participants, with an average time of administration of less than 10 min for the follow-up assessments.

There were numerous limitations to this study. A-HRSI psychometric validation was limited to participants who indicated that English was their preferred language for healthcare delivery. A separate study established the linguistic and content validity of a Spanish version of A-HRSI [34]. Refusal information was not captured from potentially eligible individuals who declined to participate. For the construct validity phase, the "general pain," "constipation," and "impact with work" items had standardized factor loadings below 0.40. It should be noted that these items were retained in A-HRSI due to participant indication of their importance during content validation rather than based upon their psychometric characteristics [17]; removal of these items has negligible impact on model fit (i.e., RMSEA

= 0.068, CFI = 0.888). While maximum likelihood estimation was used for the CFA because of the sample size (n = 303), it is acknowledged that additional CFA estimation methods exist that may be better suited for ordinal PRO data with larger samples (e.g., diagonally weighted least squares). Additionally, the broad assessment windows (i.e., T2 = 14–70 days, T3 = 71–112 days), and small sample sizes in the "change for the worse" PGIC category for evaluating change in A-HRSI (i.e., n = 8 at T2 and 12 at T3) may have further limited the power to detect differences. Participants completed T2 45-days post T1 on average. Based on consultation with ANCHOR clinical investigators, any symptoms or impacts due to randomization to treatment or active monitoring may have resolved by 4-weeks post-randomization. As such, we recommend that subsequent studies that utilize A-HRSI make use of a baseline (i.e., pre-treatment) assessment time point, with subsequent A-HRSI assessment 2–7 days (+ 3 days) and 4-weeks (± 1 week) following treatment. Future validation studies that replicate our efforts with narrow follow-up assessment windows, larger sample sizes, and well-defined criteria for detecting change (e.g., 0.5 standard deviation) [35] will allow for more rigorous psychometric testing (e.g., invariance testing) and a better understanding of A-HRSI responsiveness. It should also be noted that the ECOG PS estimates were based on participant self-report and not traditional physician assessment; physician-assessed ECOG PS ≤ 1 is an inclusion criteria for ANCHOR. Future evaluations of A-HRSI should include both objective and subjective measures of participant performance status. Finally, it is acknowledged that A-HRSI was administered via telephone in the construct validity phase and via ePRO in the responsiveness phase. While mode of administration bias may have been introduced, recent literature has demonstrated that mixing modes of PRO administration is associated with minimal mean response differences and provides participants with greater flexibility and independence when completing a PRO tool [36–38].

A-HRSI is a PRO index validated in a diverse cohort of persons eligible for ANCHOR that captures health-related symptoms and impacts related to anal HSIL. Use of this instrument will provide valuable context to the ANCHOR trial as investigators attempt to better understand participant experiences related to randomization to the treatment or active monitoring arms. Additionally, A-HRSI may have broad applicability beyond ANCHOR for other contexts where individuals have variable degrees of immunosuppression and anal HSIL. Ultimately, providing individuals with a brief, valid, condition-specific index that will allow them to self-report their symptoms and impacts related to anal HSIL may help to improve clinical care and assist providers and patients with medical decision-making.

Acknowledgements

We gratefully acknowledge the contributions of study participants and the referral sites and their staff in meeting the study recruitment goals: We also acknowledge the contributions and support of Joseph Sparano, MD of the AMC; and Don Vena, Dustin Utt, Saron Aria, Carolina Novoa, and Maria Botello-Harbaum, Ph.D., of the ANCHOR Data Management Center of the EMMES Company, LLC. This research (AMC-A01/-A03; ClinicalTrials.gov: NCT02836522) was funded in part through 2 UM1 CA121947-09, 3U54CA137788-08S1, and a National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30 CA008748, which provides partial support for the Patient-Reported Outcomes, Community Engagement, and Language Core Facility used in this investigation. The content of this research is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Data Availablity

Deidentified data that support the findings of this study are available from the authors upon reasonable request, with the permission of the AIDS Malignancy Consortium.

References

1. Goedert JJ, Hosgood HD, Biggar RJ, Strickler HD, & Rabkin CS (2016). Screening for cancer in persons living with HIV infection. Trends in Cancer, 2(8), 416–428. 10.1016/j.trecan.2016.06.007. [PubMed: 27891533]
2. Shiels MS, Islam JY, Rosenberg PS, Hall HI, Jacobson E, & Engels EA (2018). Projected cancer incidence rates and burden of incident cancer cases in HIV-infected adults in the United States through 2030. Annals of Internal Medicine, 168(12), 866–73. 10.7326/M17-2499. [PubMed: 29801099]
3. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. SEER cancer statistics review, 1975–2019, National Cancer Institute. Bethesda, MD based on Nov 2021 SEER data submission, posted to the SEER web site, April 2022. Retrieved July 27, 2022, from https://seer.cancer.gov/statistics-network/explorer/application.html
4. Wang CJ, & Palefsky JM (2019). HPV-associated anal cancer in the HIV/AIDS patient. Cancer Treatment Research, 177, 183–209. 10.1007/978-3-030-03502-0_7. [PubMed: 30523625]
5. St Laurent J, Luckett R, & Feldman S (2018). HPV vaccination and the effects on rates of HPV-related cancers. Current Problems in Cancer, 42(5), 493–506. 10.1016/j.currproblcancer.2018.06.004. [PubMed: 30041818]
6. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, & Markowitz LE (2019). Human papillomavirus vaccination for adults: Updated recommendations of the advisory committee on immunization practices. Morbidity and Mortality Weekly Report, 68(32), 698–702. 10.15585/mmwr.mm6832a3. [PubMed: 31415491]
7. Palefsky JM (2015). Screening to prevent anal cancer: Current thinking and future directions. Cancer Cytopathology, 123(9), 509–510. [PubMed: 26237741]
8. Palefsky JM (2017). Human papillomavirus-associated anal and cervical cancers in HIV-infected individuals: Incidence and prevention in the antiretroviral therapy era. Current Opinion in HIV AIDS, 12(1), 26–30. 10.1097/COH.0000000000000336. [PubMed: 27828801]
9. Koskan AM, Brennhofer SA, & Helitzer DL (2019). Screening for anal cancer precursors among patients living with HIV in the absence of national guidelines: Practitioners' perspectives. Cancer Causes Control, 30(9), 989–96. 10.1007/s10552-019-01209-8. [PubMed: 31302838]
10. Palefsky JM, & Cranston RD (2017). Anal squamous intraepithelial lesions: Diagnosis, screening, prevention, and treatment. UpToDate Inc.
11. US Department of Health and Human Services. (2009, Dec). Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. Retrieved July 27, 2022, from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
12. Haslam A, Herrera-Perez D, Gill J, & Prasad V (2020). Patient experience captured by quality-of-life measurement in oncology clinical trials. JAMA Network Open, 3(3), e200363. 10.1001/jamanetworkopen.2020.0363. [PubMed: 32129865]
13. Cella D, Wagner L, Cashy J, Hensing TA, Yount S, & Lilenbaum RC (2007). Should health-related quality of life be measured in cancer symptom management clinical trials? Lessons learned using the functional assessment of cancer therapy. Journal of the National Cancer Institute Monographs., 37, 53–60. 10.1093/jncimonographs/lgm009.
14. Ganz PA (1994). Long-range effect of clinical trial interventions on quality of life. Cancer, 74(S9), 2620–2624. 10.1002/1097-0142(19941101)74:9 [PubMed: 7954275]
15. Basch E (2016). Missing patients' symptoms in cancer care delivery-the importance of patient-reported outcomes. JAMA Oncology, 2(4), 433–4. 10.1001/jamaoncol.2015.4719. [PubMed: 26720842]

16. Basch E (2017). Patient-reported outcomes—Harnessing patients' voices to improve clinical care. N Engl J Med, 376(2), 105–8. 10.1056/NEJMp1611252. [PubMed: 28076708]
17. Burkhalter JE, Atkinson TM, Berry-Lawhorn J, Goldstone S, Einstein MH, Wilkin TJ, et al. (2018). Initial development and content validation of a health-related symptom index for persons either treated or monitored for anal high-grade squamous intraepithelial lesions. Value Health, 21(8), 984–992. 10.1016/j.jval.2018.01.018. [PubMed: 30098677]
18. Atkinson TM, Palefsky J, Li Y, Webb A, Berry JM, Goldstone S, et al. (2019). Reliability and between-group stability of a health-related quality of life symptom index for persons with anal high-grade squamous intraepithelial lesions: An AIDS Malignancy Consortium Study (AMC-A03). Quality of Life Research, 28(5), 1265–9. 10.1007/s11136-018-2089-8. [PubMed: 30617704]
19. Basch E, Iasonos A, McDonough T, Barz A, Culkin A, Kris MG, et al. (2006). Patient vs clinician symptom reporting using the National Cancer Institute common terminology criteria for adverse events: Results of a questionnaire-based study. The Lancet Oncology, 7, 903–909. 10.1016/S1470-2045(06)70910-X [PubMed: 17081915]
20. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. (1993). The functional assessment of cancer therapy scale: Development and validation of the general measure. Journal of Clinical Oncology, 11(3), 570–579. [PubMed: 8445433]
21. Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle MT, Morrissey M, et al. (2000). Assessing symptom distress in cancer patients: The M.D. Anderson symptom inventory. Cancer, 89(7), 1634–46. [PubMed: 11013380]
22. McArdle JJ (1996). Current directions in structural factor analysis. Current Directions in Psychological Science, 5, 11–18.
23. Hu L, & Bentler PM (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria vs new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6, 1–55.
24. Bentler PM (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107, 238–246. [PubMed: 2320703]
25. Hair JF, Black WC, Babin BJ, Anderson RE, & Tatham RL (2009). Multivariate data analysis (7th ed.). Pearson Prentice Hall.
26. Nunnally JC, & Bernstein IH (1994). Psychometric theory (3rd ed.). McGraw-Hill.
27. Arbuckle JL (2019). Amos (version 26.0) [computer program]. IBM SPSS.
28. R Core Team. (2022). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org
29. Middel B, & van Sonderen E (2002). Statistical significant change vs relevant or important change in (quasi) experimental design: Some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. International Journal of Integrated Care, 2, e15. 10.5334/ijic.65. [PubMed: 16896390]
30. SAS Institute Inc. (2013). SAS/ACCESS® 9.4 interface to ADA-BAS: Reference. SAS Institute Inc.
31. Atkinson TM, Dueck AC, Satele DV, Thanarajasingam G, Lafky JM, Sloan JA, et al. (2020). Clinician vs patient reporting of baseline and postbaseline symptoms for adverse event assessment in cancer clinical trials. JAMA Oncology, 6(3), 437–9. 10.1001/jamaoncol.2019.5566. [PubMed: 31876902]
32. Basch E, & Yap C (2021). Patient-reported outcomes for tolerability assessment in phase I cancer clinical trials. Journal of the National Cancer Institute, 113(8), 943–4. 10.1093/jnci/djab017. [PubMed: 33616647]
33. Goddard SL, Templeton DJ, Petoumenos K, Jin F, Hill-man RJ, Law C, et al. (2019). Association of anal symptoms with anal high grade squamous intraepithelial lesions (HSIL) among men who have sex with men: Baseline data from the study of the prevention of anal cancer (SPANC). Cancer Epidemiology, 58, 12–6. 10.1016/j.canep.2018.10.009. [PubMed: 30439602]
34. Atkinson TM, Lynch KA, Vera J, Olivares NM, Webb A, Diamond LC, et al. (2022). Linguistic validation of the Spanish version of the anal cancer high-grade squamous intraepithelial lesions outcomes Research Health-Related Symptom Index (A-HRSI): AMC-A04. Journal of Patient-Reported Outcomes. 10.1186/s41687-022-00515-1.

35. Norman GR, Sloan JA, & Wyrwich KW (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. Med Care, 41(5), 582–92. 10.1097/01.MLR.0000062554.74615.4C. [PubMed: 12719681]

36. Bennett AV, Dueck AC, Mitchell SA, Mendoza TR, Reeve BB, Atkinson TM, et al. (2016). Mode equivalence and acceptability of tablet computer-, interactive voice response system-, and paper-based administration of the U.S. National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Health and Quality of Life Outcomes., 14, 24. 10.1186/s12955-016-0426-6. [PubMed: 26892667]

37. Gwaltney CJ, Shields AL, & Shiffman S (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. Value Health, 11(2), 322–333. 10.1111/j.1524-4733.2007.00231.x. [PubMed: 18380645]

38. Muehlhausen W, Doll H, Quadri N, Fordham B, O'Donohoe P, Dogar N, et al. (2015). Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. Health and Quality of Life Outcomes, 13, 167. 10.1186/s12955-015-0362-x. [PubMed: 26446159]

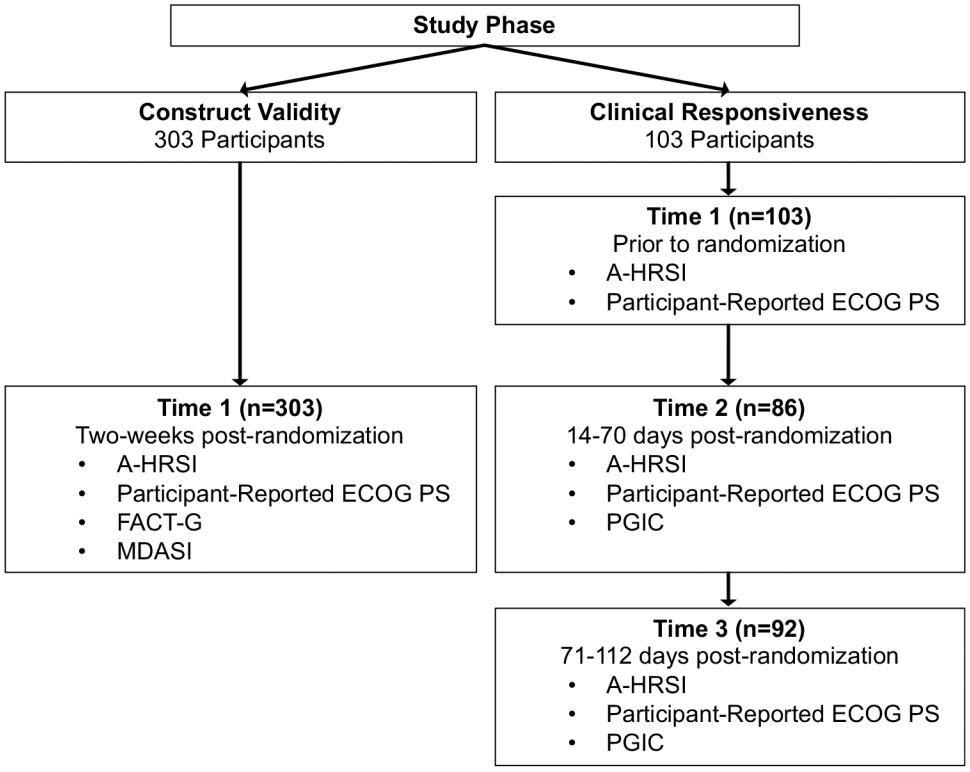

Fig. 1.
Study schema by phase and measures completed

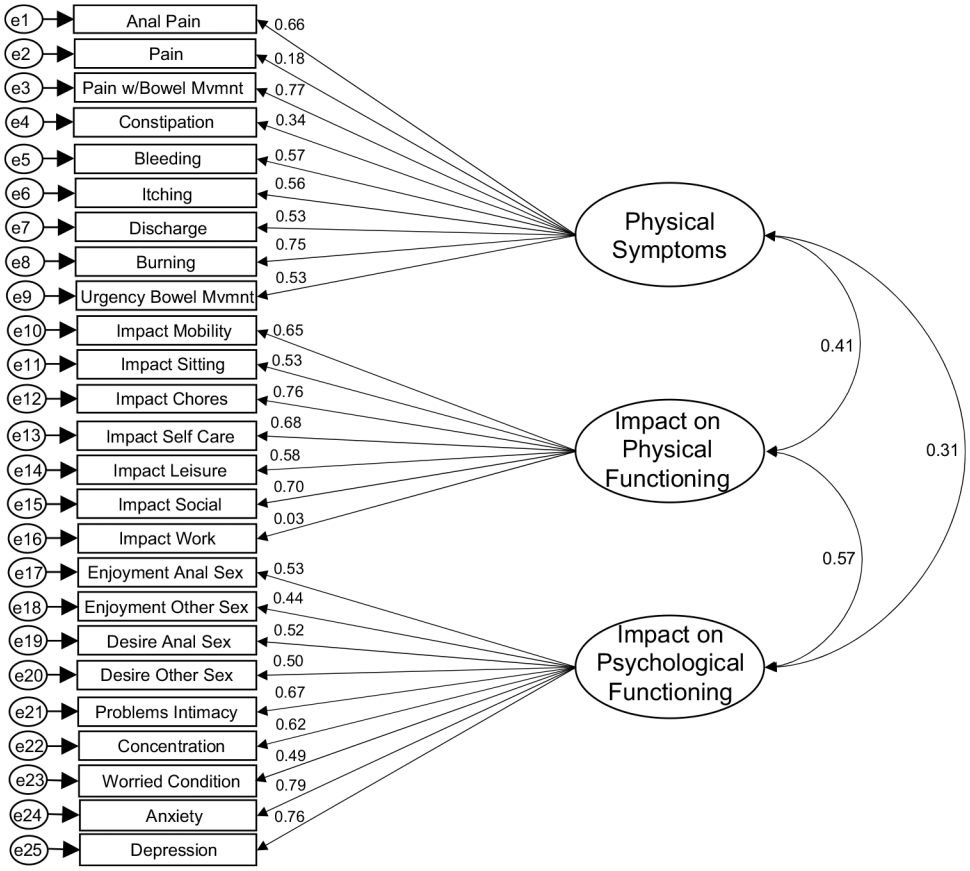

Fig. 2.
Confirmatory factor model for the three-factor solution

Table 1

Participant demographics and clinical characteristics by study phase

	Study phase	
	Construct validity (*N* = 303)	**Responsiveness (*N* = 103)**
Characteristics	*n* (%)	*n* (%)
Age (years)		
Mean (SD)	51.5 (8.1)	51.0 (9.4)
Median	51	52.5
Range	35–73	35–74
Gender identity		
Cisgender female	65 (21.5%)	13 (12.6%)
Transgender female	6 (2.0%)	2 (1.9%)
Cisgender male	229 (75.6%)	87 (84.5%)
Transgender male	1 (0.3%)	0 (0.0%)
Gender non-conforming	2 (0.7%)	1 (1.0%)
Race		
White	88 (29.0%)	45 (43.7%)
African American	195 (64.4%)	40 (38.8%)
Asian or Pacific Islander	3 (1.0%)	3 (2.9%)
American Indian or Native Alaskan	0 (0%)	3 (2.3%)
Other or multiple races	17 (5.6%)	12 (11.7%)
Ethnicity		
Hispanic or Latino	30 (9.9%)	29 (28.1%)
Non-Hispanic or Latino	273 (90.1%)	72 (69.9%)
Not reported	0 (0.0%)	2 (2.0%)
Study Arm		
Active monitoring	143 (47.2%)	55 (53.4%)
Treatment	160 (52.8%)	48 (46.6%)
Participant-reported ECOG performance status		
0–1	251 (82.8%)	92 (89.3%)
2–4	52 (17.2%)	9 (8.7%)

ECOG indicates Eastern Cooperative Oncology Group

Table 2

Fit indices for confirmatory factor models ($N = 303$)

Model	RMSEA	95% CI	CFI	χ^2	*df*	$\Delta\chi^2/df$	*p*
Model 1	0.093	0.087; 0.099	0.733	956.23	265		
Model 2	0.090	0.084; 0.097	0.749	913.62	264	42.61	< 0.05
Model 3	0.065	0.058; 0.072	0.862	619.93	262	146.85	< 0.05
Model 4	0.065	0.058; 0.072	0.870	597.18	261	22.75	< 0.05

RMSEA indicates Root-Mean-Squared Error of Approximation; *CI* confidence interval, *CFI* confirmatory fit index; *df* degrees of freedom, $\Delta\chi^2/df$, change in χ^2 given corresponding change in degrees of freedom

Table 3

Pairwise Pearson *r* coefficients between A-HRSI subscales and legacy measures

Legacy Measure	A-HRSI subscale		
	Physical symptoms	**Impact on physical functioning**	**Impact on psychological functioning**
Participant-reported ECOG PS	0.27 (< 0.01)	0.49 (< 0.01)	0.43 (< 0.01)
FACT-G			
Physical Well-Being	− 0.52 (< 0.01)	− 0.66 (< 0.01)	− 0.51 (< 0.01)
Social/Family Well-Being	− 0.03 (0.64)	− 0.07 (0.23)	− 0.21 (< 0.01)
Emotional Well-Being	− 0.32 (< 0.01)	− 0.35 (< 0.01)	− 0.58 (< 0.01)
Functional Well-Being	− 0.25 (< 0.01)	− 0.35 (< 0.01)	− 0.45 (< 0.01)
Overall FACT-G	− 0.35 (< 0.01)	− 0.45 (< 0.01)	− 0.56 (< 0.01)
MDASI			
Severity	0.46 (< 0.01)	0.56 (< 0.01)	0.59 (< 0.01)
Interference	0.47 (< 0.01)	0.55 (< 0.01)	0.47 (< 0.01)

A-HRSI indicates ANCHOR Health-Related Symptom Index; Participant-Reported ECOG PS, Eastern Cooperative Oncology Group Performance Status; FACT-G, Functional Assessment of Cancer Therapy—General; MDASI, M.D. Anderson Symptom Inventory

Table 4

Descriptive Statistics for Changes in A-HRSI Subscales (T3 minus T2) According to Change in PGIC and Participant-Reported ECOG PS, mean (SD)

	PGIC			Participant-reported ECOG PS		
	Change for the better	**No change**	**Change for the worse**	**Change for the better**	**No change**	**Change for the worse**
A-HRSI subscale	*n* = 34	*n* = 38	*n* = 12	*n* = 18	*n* = 43	*n* = 24
Physical Symptoms	− 0.12 (0.86)	0.05 (0.74)	0.27 (0.65)	− 0.23 (0.61)	− 0.06 (0.80)	0.37 (0.78)
Impact on Physical Functioning	0.01 (1.14)	− 0.06 (0.99)	0.65 (1.25)	− 0.76 (1.18)	0.10 (0.74)	0.63 (1.22)
Impact on Psychological Functioning	− 0.22 (1.39)	0.05 (1.19)	0.90 (1.50)	− 0.34 (1.42)	− 0.08 (1.11)	0.64 (1.55)

A-HRSI indicates ANCHOR Health-Related Symptom Index; Participant-Reported ECOG PS, Participant-Reported Eastern Cooperative Oncology Group Performance Status; PGIC, Participant Global Impression of Change

Table 5

Responsiveness comparison of changes (T3 minus T2), mean difference (95% confidence interval)

A-HRSI subscale	PGIC				Participant-reported ECOG PS			
	Change for the worse vs no change	*p*	Change for the worse vs change for the better	*p*	Change for the worse vs no change	*p*	Change for the worse vs change for the better	*p*
Physical symptoms	0.22 (− 0.30 to 0.73)	0.40	0.39 (− 0.13 to 0.90)	0.14	0.43 (0.05–0.82)	0.03	0.60 (0.13–1.07)	0.01
Impact on physical functioning	0.71 (− 0.01 to 1.43)	0.06	0.65 (− 0.08 to 1.38)	0.08	0.53 (0.02–1.03)	0.04	1.39 (0.77–2.01)	0.01
Impact on psychological functioning	0.86 (− 0.02 to 1.73)	0.06	1.12 (0.23–2.01)	0.01	0.72 (0.04–1.40)	0.04	0.98 (0.14–1.82)	0.02

A-HRSI indicates ANCHOR Health-Related Symptom Index; Participant-Reported ECOG PS, Participant-Reported Eastern Cooperative Oncology Group Performance Status; PGIC, Participant Global Impression of Change