Quality of Life Research (2019) 28:1265–1269
https://doi.org/10.1007/s11136-018-2089-8

BRIEF COMMUNICATION

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)

Thomas M. Atkinson[1] · Joel Palefsky[2] · Yuelin Li[1] · Andrew Webb[1] · J. Michael Berry[2] · Stephen Goldstone[3] · Rebecca Levine[4] · Timothy J. Wilkin[5] · Gary Bucher[6] · David Cella[7] · Jack E. Burkhalter[1] on behalf of the ANCHOR HRQoL Implementation Group

Accepted: 13 December 2018 / Published online: 7 January 2019

Abstract

Purpose The Anal Cancer HSIL Outcomes Research (ANCHOR) trial aims to determine whether treating precancerous anal high-grade squamous intraepithelial lesions (HSIL), versus active surveillance, is effective in reducing anal cancer incidence in HIV-infected individuals. We evaluated the reliability (i.e., internal consistency, test–retest) and between-group stability of a 25-item ANCHOR Health-Related Symptom Index (A-HRSI).

Methods ANCHOR participants at least 1-month post-randomization to treatment or active surveillance completed the A-HRSI via telephone. Participants were contacted 7–10 days later to complete the A-HRSI and a participant global impression of change (PGIC) item.

Results Participants ($n = 100$) were enrolled (mean age = 51.4, 79% cisgender-male, 73% African American, 9% Hispanic) from five ANCHOR sites. Cronbach's α was good for the physical symptoms (0.82) domain and fair for the physical impacts (0.79) and psychological symptoms (0.73) domains. Intraclass correlation coefficients were good for each of respective domains (i.e., 0.80, 0.85, and 0.82). There were no significant differences in PGIC between the treatment ($n = 56$) and active surveillance ($n = 44$) groups ($F(1,98) = 2.03$, $p = 0.16$).

Conclusions The A-HRSI is able to reliably assess participant-reported symptoms and impacts of anal HSIL across a 7–10 days of timeframe. Future work will involve the establishment of construct and discriminant validity prior to inclusion in the full ANCHOR trial.

Keywords Patient-reported outcomes · Health-related quality of life · Clinical outcome assessments · Neoplasms · ANCHOR trial

Introduction

Anal cancer incidence is on the rise in the United States, especially among those infected with HIV [1, 2]. For decades, the primary focus of anal cancer prevention has

Jack E. Burkhalter is the senior author.

The members of The ANCHOR HRQoL Implementation Group are listed in Acknowledgement section.

✉ Thomas M. Atkinson
atkinsot@mskcc.org

[1] Memorial Sloan Kettering Cancer Center, New York, NY, USA

[2] University of California-San Francisco, San Francisco, CA, USA

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

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

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

[6] Anal Dysplasia Clinic Midwest, Chicago, IL, USA

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

been screening for and treating HPV-associated precancerous high-grade squamous intraepithelial lesions (HSIL) [3]. However, the effectiveness of treating anal HSIL is not yet known. As such, in 2015 the US National Cancer Institute (NCI) funded Anal Cancer HSIL Outcomes Research (ANCHOR), a phase III clinical trial (clinicaltrials.gov identifier: NCT02135419) to be conducted within the AIDS Malignancy Consortium to determine whether treating anal HSIL, versus active surveillance, is effective in reducing incidence of anal cancer in HIV-infected individuals. Since anal HSIL and its treatment are believed to negatively impact various aspects of health-related quality of life (HRQoL) [4–6], we have initiated the development and validation of the ANCHOR Health-Related Symptom Index (A-HRSI), a self-report instrument that will be used to prospectively capture physical symptoms, physical impacts, and psychological symptoms within ANCHOR and similar anal cancer screening programs [7].

Using qualitative techniques based on best practices in measure development [8–10], we have synthesized direct participant input to establish the content validity of A-HRSI [7]. The next step in the A-HRSI psychometric development process is to evaluate its ability to provide stable, reproducible results [11]. Therefore, we sought to establish internal consistency and test–retest reliability of this newly developed measure, as well as determine whether this instrument is similarly stable for treatment and active surveillance groups.

Methods

Participants

An independent sample of English-speaking people living with HIV who were diagnosed with anal HSIL in the last 9 months and were at least 1-month post-randomization to treatment or active surveillance were recruited from five ANCHOR sites in Bronx, NY, New York, NY (two sites), Chicago, IL, and San Francisco, CA. This assessment and follow-up window was selected by an expert panel comprising investigators and clinicians from the ANCHOR trial [7], with the rationale that participants would have relatively stable experiences of symptoms and impacts during this timeframe. Eligible participants were recruited from the ANCHOR sites until a total of 100 had completed both the initial and follow-up assessment. All study staff and data collection was based at Memorial Sloan Kettering Cancer Center (MSK). The study was reviewed and approved by the NCI's Cancer Therapy Evaluation Program and the Institutional Review Boards at each study site.

Measures

ANCHOR Health-Related Symptom Index (A-HRSI)

The A-HRSI is a 25-item HRQoL measure that assesses participant physical symptoms (9 items), physical impacts (7 items), and psychological symptoms (9 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) [7]. Domain scores are derived through calculating the mean of the completed items (i.e., domain ranges 0–4). Higher domain scores indicate worse participant experience of symptom or impact burden.

Participant global impression of change (PGIC) (follow-up 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 7-point scale (i.e., −3, −2, −1, 0, 1, 2, 3) that represented HRQoL ranging from "very much worse" to "very much better."

Procedure

Eligible participants from the five ANCHOR sites were referred to contact study staff at MSK via telephone to learn more about the study and confirm eligibility. Upon enrollment, a trained clinical research coordinator (CRC) reviewed the study procedures and secured verbal consent to participate, followed by scheduling the initial assessment. The CRC then contacted the eligible participant via telephone on the scheduled date and administered the A-HRSI, followed by scheduling of the follow-up assessment 7–10 days later. This procedure was repeated at the follow-up assessment, with participants contacted on the scheduled assessment date to complete the A-HRSI and PGIC. A protocol was established to ensure that a minimum of 40 participants from each study arm (i.e., treatment and active surveillance) would be enrolled in the study.

Statistical analysis

The accrual target ($n = 100$) was established to ensure that both treatment and active surveillance arms were adequately represented (i.e., ≥ 40 participants from each group) for the between-group stability analysis. Cronbach's α was calculated for each of the three primary domains at the initial assessment time point as an indicator of internal consistency [13]. Cronbach's α values ranging from 0.70 to 0.79 are considered to be fair, whereas values ≥ 0.80 are indicative of

good internal consistency [14]. Intraclass correlation coefficients (ICC) were calculated between initial and follow-up assessments for each of the three primary domain scores, with values ≥ 0.70 representing good test–retest reliability [15]. A one-way analysis of variance (ANOVA) was used to determine whether the active monitoring and treatment groups differed with respect to their PGIC ratings at the follow-up assessment. Analyses were completed using SPSS version 25 [16].

Results

A total of 103 participants were enrolled in the study, with 100 (mean age = 51.4, 21% non-male-cisgender, 73% African American, 9% Hispanic) completing both the initial and follow-up assessments. Three participants were lost to follow-up after failing to respond within their respective test–retest windows. The mean number of days from initial to follow-up assessment was 7.51 (SD = 0.88). Fifty-six percent of the participant sample was receiving treatment, with the remaining participants randomized to active surveillance. Participant demographic and clinical characteristics are displayed in Table 1.

Means and standard deviations for the A-HRSI items and domain scores are displayed in Table 2 by assessment time point. Participant responses for all A-HRSI items ranged from 0 to 4, with the exception of "problems participating in leisure activities" during the initial assessment (0–3). Cronbach's α was good (i.e., ≥ 0.80) for the physical symptoms domain (0.82) and fair (i.e., 0.70–0.79) for the physical impacts (0.79) and psychological symptoms (0.73) domains. ICCs were good (i.e., ≥ 0.70) for each of the physical symptoms (0.80), physical impacts (0.85), and psychological symptoms (0.82) domains.

Table 3 is a display of treatment and active surveillance group responses to the PGIC item asked at the 7–10 days of follow-up. The one-way ANOVA showed no significant differences between groups when compared on their PGIC responses ($F(1,98) = 2.03$, $p = 0.16$).

Table 1 Participant demographic and clinical characteristics ($n = 100$)

Characteristics	*n* (%)
Age (years)	
Mean (SD)	51.36 (8.0)
Range	35–75
Gender	
Cis-male	79 (79.0)
Cis-female	19 (19.0)
Trans-male	0 (0.0)
Trans-female	2 (2.0)
Race	
Non-White	73 (73.0)
Ethnicity	
Hispanic	9 (9.0)
Referring site	
Chicago, IL	45 (45.0)
New York, NY (Site 1)	17 (17.0)
New York, NY (Site 2)	15 (15.0)
San Francisco, CA	23 (23.0)
Treatment modality	
Ablation	49 (49.0)
Topical	6 (6.0)
Surgical	0 (0.0)
Active surveillance	45 (45.0)

Discussion

In the context of increasing anal cancer incidence in the United States and efforts underway to evaluate preventive treatment of precancerous lesions versus active surveillance, it is important to provide those who are diagnosed with anal HSIL with a tool that enables them to accurately and reliably report symptoms and impacts. The 2009 United States Food and Drug Administration Guidance for Industry for Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims included regulatory recommendations that instruments must demonstrate evidence of internal consistency and test–retest reliability in order to be used for clinical trial purposes [11]. In the present study, we found that the internal consistency of the three primary domains (i.e., physical symptoms, psychological symptoms, physical impacts) of the A-HRSI was fair to good, and test–retest reliability surpassed the targeted threshold for the three domains when the measure was administered twice over a 7–10 days of timeframe. These findings provide initial evidence in support of the psychometric properties of this HRQoL measure.

In addition, participants were asked to complete a subjective significance question (PGIC) at the 7–10 days of follow-up in order to rate how much their HRQoL had changed since the initial assessment. While the assessment window of 1-month post-treatment or randomization to active surveillance was selected by an expert panel of ANCHOR trialists because it was expected that this would be a time when symptoms or impacts should be stable, there was variability in the levels of change that were reported by participants (Table 3). Despite this, there were no statistically significant differences in patterns of variability between the treatment and active surveillance groups. This finding is an initial indication that the A-HRSI exhibits between-group stability in the capture of participant-reported change for these

Table 2 Means (SD) for A-HRSI items by domain and assessment time point ($n = 100$)

Item	Domain	Mean (SD)	
		Baseline	7–10 day follow-up
Anal pain	Physical Sx	0.57 (1.03)	0.39 (0.94)
Bleeding from the anus	Physical Sx	0.45 (0.88)	0.34 (0.79)
Burning sensations in anal area	Physical Sx	0.33 (0.74)	0.37 (0.87)
Constipation	Physical Sx	0.79 (1.15)	0.66 (1.16)
Discharge in anal area	Physical Sx	0.44 (0.90)	0.41 (0.93)
Itching in or around the anus	Physical Sx	0.76 (1.05)	0.61 (1.10)
Pain during bowel movements	Physical Sx	0.59 (1.04)	0.41 (0.85)
Pain other than anal pain	Physical Sx	1.18 (1.42)	1.45 (1.49)
Urgency for bowel movements	Physical Sx	0.83 (1.26)	0.64 (1.11)
Anxiety	Psych Sx	1.52 (1.45)	1.34 (1.49)
Decreased desire for anal sexual activity	Psych Sx	1.45 (1.27)	1.27 (1.12)
Decreased desire for other sexual activity	Psych Sx	1.62 (1.22)	1.55 (1.18)
Decreased enjoyment of anal sexual activity	Psych Sx	1.03 (1.15)	0.98 (1.13)
Decreased enjoyment of other sexual activity	Psych Sx	1.47 (1.28)	1.40 (1.20)
Depression	Psych Sx	1.57 (1.46)	1.36 (1.47)
Difficulty concentrating	Psych Sx	1.13 (1.25)	1.15 (1.31)
Problems with intimate relationships	Psych Sx	1.41 (1.13)	1.59 (1.51)
Worried about condition getting worse	Psych Sx	1.46 (1.47)	1.29 (1.40)
Problems completing daily household chores	Physical impacts	0.76 (1.18)	0.66 (1.16)
Problems participating in leisure activities	Physical impacts	0.33 (0.83)	0.36 (0.77)
Problems participating in social activities	Physical impacts	0.65 (1.18)	0.71 (1.23)
Problems taking care of myself	Physical impacts	0.37 (0.85)	0.26 (0.72)
Problems with physical ability to move around	Physical impacts	0.66 (1.12)	0.73 (1.20)
Problems with sitting	Physical impacts	0.42 (1.03)	0.53 (1.09)
Problems with work productivity	Physical impacts	0.86 (1.08)	0.74 (1.00)
Physical symptom domain score		0.66 (0.68)	0.59 (0.68)
Psychological symptom domain score		1.43 (0.81)	1.32 (0.87)
Physical impacts domain score		0.58 (0.71)	0.57 (0.72)

A-HRSI ANCHOR Health-Related Symptom Index; *Sx* symptoms

Table 3 Participant global impression of change responses by treatment and active surveillance groups

Response	Active surveillance		Treatment	
	n	% Within group	*n*	% Within group
Very much worse	2	4.5	9	16.1
Moderately worse	0	0.0	3	5.4
A little worse	2	4.5	3	5.4
About the same	24	54.5	20	35.7
A little better	4	9.1	8	14.3
Moderately better	7	15.9	6	10.7
Very much better	5	11.4	7	12.5

distinct approaches to treatment. We will further explore this in future validation efforts. Additionally, the unexpected variability in levels of change that were reported across participants during the study assessment window underlines the need for a validated instrument that will allow individuals to report these changes to their clinical team.

This study is not without limitations. Our study was open to English-speaking participants only. Preliminary efforts are underway to translate the A-HRSI into Spanish to allow us to capture the symptoms and impacts of anal HSIL in Spanish-speaking participants who have limited English proficiency.

The present results provide us with confidence that the A-HRSI has good internal consistency and test–retest reliability. The next steps in the validation process are to establish construct validity, as well as the ability of the A-HRSI to be clinically responsive to prospective score changes across participants. Ultimately this process will yield a validated measure to be deployed in the ANCHOR trial and future anal cancer screening initiatives. It is anticipated that the A-HRSI outcomes from ANCHOR will help clinicians and those who are impacted understand and enhance decision-making regarding treatment or surveillance of anal HSIL.

Acknowledgements The ANCHOR HRQoL Implementation Group members are Susan M. Holland, Jeannette Lee, Erika I. Lubetkin, Kathleen A. Lynch, Jeff Taylor, and Amanda Watsula-Morley.

Funding This research (ClinicalTrials.gov: NCT02836522) was funded in part through 2 UM1 CA121947-09, 3U54CA137788-08S1, and the NIH/NCI Cancer Center Support Grant P30 CA008748-50, which provides partial support for the Patient-Reported Outcomes, Community-Engagement and Language Core Facility used in this investigation.

References

1. Nelson, R. A., Levine, A. M., Bernstein, L., Smith, D. D., & Lai, L. L. (2013). Changing patterns of anal canal carcinoma in the United States. *Journal of Clinical Oncology, 31*(12), 1569–1575.
2. Shiels, M. S., Kreimer, A. R., Coghill, A. E., Darragh, T. M., & Devesa, S. S. (2015). Anal cancer incidence in the United States, 1977–2011: Distinct patterns by histology and behavior. *Cancer Epidemiology, Biomarkers & Prevention, 24*(10), 1548–1556.
3. Palefsky, J. M. (2015). Screening to prevent anal cancer: Current thinking and future directions. *Cancer Cytopathology, 123*(9), 509–510.
4. Chang, G. J., Berry, M. J., Jay, N., Palefsky, J. M., & Welton, M. L. (2002). Surgical treatment of high-grade anal squamous intraepithelial lesions. *Diseases of the Colon & Rectum, 45*(4), 453–458.
5. Siegenbeek van Heukelom, M. L., Richel, O., Nieuwkerk, P. T., De Vries, H. J., & Prins, J. M. (2016). Health-related quality of life and sexual functioning of HIV-positive men who have sex with men who are treated for anal intraepithelial neoplasia. *Disease Colon Rectum, 59*(1), 42–47.
6. Tinmouth, J., Raboud, J., Ali, M., Malloch, L., Su, D., Sano, M., et al. (2011). The psychological impact of being screened for anal cancer in HIV-infected men who have sex with men. *Disease Colon Rectum, 54*(3), 352–359.
7. Burkhalter, J. E., Atkinson, T. M., Berry-Lawhorn, J., Goldstone, S., Einstein, M. H., Wilkin, T. J., Lee, J., 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.
8. Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity-establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) Instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting concepts for a new PRO instrument. *Value in Health, 14*(8), 967–977.
9. Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity-establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2—Assessing respondent Understanding. *Value in Health, 14*(8), 978–988.
10. Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks: Sage.
11. US Department of Health and Human Services. (2009). *Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims*. Accessed September 19, 2018, from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
12. Osaba, D., Rodrigues, G., Myles, J., Zee, B., & Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology, 16*(1), 139–144.
13. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.
14. Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55.
15. Ng, S. S. M., Leung, W. W., Wong, C. Y. N., Hon, S. S. F., Mak, T. W. C., Ngo, D. K. Y., & Lee, J. F. Y. (2013). Quality of life after laparoscopic vs open sphincterpreserving resection for rectal cancer. *World Journal of Gastroenterology, 19*(29), 4764–4773.
16. IBM Corp. Released 2017. IBM SPSS statistics for Windows, version 25.0. Armonk, NY: IBM Corp.