AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Otolaryngology–Head and Neck Surgery
2020, Vol. 163(5) 979–985

DOI: 10.1177/0194599820927364
http://otojournal.org
SAGE

Validation of a Novel, Multidomain Head and Neck Cancer Appearance- and Function-Distress Patient-Reported Outcome Measure

Evan M. Graboyes, MD, MPH[1,2], Brittany N. Hand, PhD, OTR/L[3], Mark A. Ellis, MD[1], Andrew T. Huang, MD[4], Marci L. Nilsen, PhD, RN[5,6], Patrik Pipkorn, MD, MSCI[7], Courtney H. Marsh[1], Stacey Maurer, PhD[8], Terry A. Day[1], and Katherine R. Sterba, PhD, MPH[2,9]

Abstract

Objectives. Distress with self-perceived changes in appearance and function can result in body image disturbance (BID), which is common in head and neck cancer (HNC) survivors and a major source of psychosocial morbidity. To address the lack of psychometrically sound patient-reported outcome measures (PROMs) of HNC-related BID, we aim to create and validate the Inventory to Measure and Assess imaGe disturbancE–Head & Neck (IMAGE-HN).

Study Design. Survey study.

Setting. Multiple academic centers.

Subjects and Methods. Following item development, HNC survivors from 4 academic centers completed the IMAGE-HN. Item responses were psychometrically analyzed using confirmatory factor analysis (CFA) and Rasch analysis.

Results. Item development resulted in a 31-item PROM consisting of 5 individual domains and a global domain. In total, 305 HNC survivors of diverse ages, HNC subsites, and reconstructive paradigms completed the initial items. After removal of 3 items for local dependence, CFA confirmed the unidimensionality and local independence (item residual correlations $<|0.20|$) for each domain. Rasch analysis indicated acceptable fit (infit and outfit mean squares <2.0), monotonicity of all rating scale categories, and low person misfit ($<4\%$). Person separation indices and person reliability were adequate for each domain except appearance concealment, which was removed (4 items). This resulted in the IMAGE-HN, a psychometrically acceptable 24-item PROM of HNC-related BID consisting of a global scale and 4 subscales measuring unique constructs and comprised independent items.

Conclusions. IMAGE-HN is a novel, psychometrically sound, multidomain PROM of HNC-related BID for use in clinical and research settings.

Keywords

head and neck cancer, appearance, disfigurement, body image, survivorship, pyscho-oncology, patient reported outcome measure, PROMIS

Received December 10, 2019; accepted March 18, 2020.

Head and neck cancer (HNC) arises in cosmetically and functionally critical areas, including the tongue, mandible, larynx, pharynx, neck, and face. Optimal treatment of HNC incorporates various combinations and sequences of surgery, radiation, and cytotoxic chemotherapy.[1] As a result of HNC and its treatment, debilitating changes in highly visible and socially significant parts of the body are common and persist as late and long-term treatment toxicity.[2,3] Self-perceived disfigurement and appearance concerns as well as distress with functional impairments in smiling, swallowing, and speaking can result in substantial psychosocial morbidity, including depression,[4] anxiety,[5] suicide,[6] and concerns about body image.[2,3] These negative self-perceived changes in appearance and function, when severe, result in body image disturbance (BID).[7] BID is prevalent in HNC survivors, associated

[1]Department of Otolaryngology–Head and Neck Surgery, Medical University of South Carolina, Charleston, South Carolina, USA
[2]Hollings Cancer Center, Medical University of South Carolina, Charleston, South Carolina, USA
[3]School of Health and Rehabilitation Sciences, The Ohio State University, Columbus, Ohio, USA
[4]Bobby R. Alford Department of Otolaryngology–Head and Neck Surgery, Baylor College of Medicine, Houston, Texas, USA
[5]Department of Otolaryngology–Head and Neck Surgery, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
[6]Department of Acute and Tertiary Care, University of Pittsburgh School of Nursing, Pittsburgh, Pennsylvania, USA
[7]Department of Otolaryngology–Head and Neck Surgery, Washington University School of Medicine, St Louis, Missouri, USA
[8]Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina, USA
[9]Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina, USA

Corresponding Author:
Evan M. Graboyes, MD, MPH, Department of Otolaryngology–Head & Neck Surgery, Medical University of South Carolina, 135 Rutledge Ave, MSC 550, Charleston, SC 29425, USA.
Email: graboyes@musc.edu

with social isolation, stigmatization, and depression, and can have a devastating negative impact on quality of life (QOL).[8]

BID, because of its subjective nature and poor correlation with observer-rated measures of disfigurement,[9,10] is optimally measured by a patient-reported outcome measure (PROM). Unfortunately, psychometrically sound PROMs to measure HNC-related BID are lacking.[11] Existing studies evaluating BID in patients with HNC have had to rely on PROMs primarily intended to measure image concerns in non-HNC patient populations that may fail to reliably capture appearance and functional concerns relevant to patients with HNC.[11] To address this gap, we aimed to develop and validate the IMAGE-HN (Inventory to Measure and Assess imaGe disturbancE–Head & Neck), a novel multidomain PROM for the assessment of HNC-related body image concerns in the clinical and research settings.

Methods

Study Design

The IMAGE-HN was developed in accordance with the Patient Reported Outcomes Measurement Information System (PROMIS) guidelines for the development and validation of PROMs.[12] Steps in the process include (1) comprehensive literature search to identify existing measures, (2) qualitative work to assess domain coverage, (3) cognitive interviewing of items for feedback on language and item clarity, (4) confirmation of domain factor structure and item analysis using Rasch analysis, and (5) validity testing of the final instrument. We have published our literature review (step 1)[11] and qualitative work (step 2).[13] The current study describes the cognitive interviewing and confirmation of factor structure and item analysis (steps 3-4).

Instrument Development

Instrument development was guided by our conceptual framework of 5 key domains of HNC-related BID: (1) other-oriented appearance concerns (OOA; verbal and nonverbal reactions by others to the appearance of patients with HNC), (2) personal dissatisfaction with appearance (PDA; HNC patients' dissatisfaction with their appearance), (3) distress with functional impairments (DFI; challenges related to speaking, swallowing, oral competence, etc), (4) appearance concealment (AC; attempts to camouflage the head and neck), and (5) social avoidance and isolation (SA; avoidance of social interaction due to image concerns).[13] We then drafted items for each of the 5 scale domains using patient quotes or key vocabulary from our qualitative work.[13] An item stem was developed (''Because of the way that HNC or its treatment has changed my body . . .'') so that the scale could be used pretreatment and longitudinally thereafter, addressing limitations of prior PROMs.[14] An item response set was selected from PROMIS preferred options to assess frequency (never, rarely, sometimes, often, always),[12] corresponding to a Likert scale of 0 to 4.

These items were reviewed for content validity and clarity by 18 content experts (a multidisciplinary group of HNC clinicians) from 12 academic medical centers. We performed cognitive interviews with HNC survivors (n = 11) using standard techniques[15] to ensure optimal item content relevance, vocabulary, complexity, clarity, and readability. The 31-item, preliminary IMAGE-HN then underwent multi-institution validation by HNC survivors. The study was approved by the Medical University of South Carolina (MUSC) Institutional Review Board. A waiver of written informed consent was granted for the validation portion of the study.

Validation Study Participants and Procedures

To enhance external validity of IMAGE-HN by enrolling a sample diverse with respect to demographic, clinical, and geographic considerations, patients were recruited from multidisciplinary HNC clinics within National Cancer Institute (NCI)–designated cancer centers at 4 tertiary care academic medical centers (MUSC, Baylor College of Medicine, University of Pittsburgh Medical Center, and Washington University School of Medicine). Eligibility criteria included (1) $\geq$18 years of age, (2) history of surgically managed HNC, and (3) no known active, untreated disease. Of the 309 eligible subjects approached for the study, 305 (98.7%) enrolled; 2 declined due to insufficient time, 1 for inability to use the technology and 1 for reasons unknown. At all 4 sites, patients with HNC were screened for eligibility and enrolled during routine posttreatment or survivorship clinic visit. Using an electronic tablet, participants completed the 31-item IMAGE-HN and a self-report demographic and clinical questionnaire. Participants received $10 for completing the IMAGE-HN. Researchers at each site were trained to ensure consistent patient selection, accrual, and questionnaire administration.

Confirmatory Factor Analysis

Confirmatory factor analyses (CFAs) were conducted following PROMIS specifications to determine the degree to which items in each domain represented a single unidimensional construct.[12] This includes input of a polychoric correlation matrix for participants with complete data (N = 305; 100%) and use of robust diagonal weighted least squares estimation.[16] A range of model fit indices was examined, including root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and Tucker-Lewis index (TLI). Standardized factor loadings were examined to evaluate the relationship between the item and latent construct; factor loadings $\geq|0.32|$ were interpreted as a statistically significant relationship.[17] Item residual correlations were examined for local dependence; items with residual correlations $\geq|0.20|$ were interpreted as locally dependent. If items were found to have local dependence, the item with the lowest factor loading on the latent construct was removed. CFA was conducted in R using package ''lavaan.''

Rasch Analysis

Once a unidimensional set of items for each IMAGE-HN domain was identified, separate Rasch analyses were

performed. Rasch analysis is a state-of-the-art technique recommended for developing novel PROMs[12] that compares observed patient responses with predicted responses at the item level.[18] Separate Rasch analyses using a rating scale model and joint maximum likelihood estimation were conducted using WINSTEPS (Winsteps.com), version 3.90.0.28. First, the appropriateness of the rating scale was evaluated using the following criteria: (1) at least 10 observations of each category, collapsed across all items; (2) monotonicity of rating scale categories (ie, 0-4), as evidenced by an increase in average category difficulty with increasing category value; and (3) outfit mean square is <2.0. Second, the fit of the items and persons to the Rasch model was evaluated by examining infit and outfit mean squares and standardized z values. Any items or persons with mean square values >1.60 and standardized z values greater than 2.0 were considered indicative of misfit to the Rasch model.[19] Third, reliability indicators were examined, including (1) person reliability, which represents the reproducibility of person ordering and was interpreted such that values >0.5 were considered adequate, >0.80 were considered good, and >0.90 were considered high,[20] and (2) the separation index, which was used to calculate the number of statistically distinct ability strata in the sample. The number of person strata is calculated according to the formula $(4 * G + 1)/3$, where G is the person separation index and is an indicator of the number of statistically distinct person measures with centers 3 calibration errors apart.[21] Ideally, a measure would be able to separate respondents at least into "high," "medium," and "low" on the latent construct.[21] Test targeting, test coverage, and item hierarchy were examined visually using person-item maps.

Global Domain

After identifying psychometrically sound sets of items in each IMAGE-HN domain, we sought to determine whether a global domain existed for the purpose of generating a single overall measure. First, we performed a single-factor CFA by combining remaining items from all domains together. We evaluated the same model fit indices described above (ie, RMSEA, SRMR, CFI, TLI) and identified locally dependent items (residual correlations >|0.20|). Instances of local dependence were addressed by eliminating items from the global domain with the highest residual correlations with the greatest number of items, allowing us to retain as many items as possible. Once we identified a unidimensional set of items without local dependence, we performed a Rasch analysis using joint maximum likelihood estimation and a rating scale model as described above (ie, evaluation of the rating scale, person and item fit, reliability indicators, item hierarchy, and test targeting and coverage).

Results

A total of 305 participants were included in the study and completed the IMAGE-HN, of whom 301 provided demographic and oncologic information (**Table 1**). The median age was 65.3 years (range, 21-92); 70.4% were male (212/301), 42.9% had oral cavity cancer (129/301), 61.8% had

Table 1. Sociodemographic and Oncologic Characteristics of the Study Cohort (N = 301).[a]

Characteristic	No. (%)
Age, median (range), y	65.3 (21-92)
Sex	
Female	89 (29.6)
Male	212 (70.4)
Race	
White	251 (83.4)
African American	34 (11.3)
Asian	5 (1.7)
More than 1 race/prefer not to answer	11 (3.7)
Ethnicity	
Non-Hispanic	278 (92.4)
Hispanic	8 (2.7)
Prefer not to answer	15 (5.0)
Insurance	
Private	104 (34.6)
Medicare	162 (53.8)
Medicaid or self-pay	23 (7.7)
Other	12 (4.0)
Marital status[b]	
Married/current partner	204 (68.0)
Single/separated/divorced/widowed	96 (32.0)
Living situation[b]	
Spouse	196 (65.3)
Parents/children/friends	33 (12.4)
Self	55 (18.3)
Other	16 (5.3)
Educational attainment[c]	
Less than high school	27 (9.0)
High school graduate	85 (28.4)
Some college	77 (25.8)
College graduate	69 (23.1)
Graduate school	41 (13.7)
Employment[c]	
Part or full-time paid work (including at home)	83 (27.1)
Unemployed	11 (3.7)
Disability	54 (17.9)
Retired	153 (50.8)
Rurality[b]	
Rural	117 (39.0)
Suburban	147 (49.0)
Urban	36 (12)
Tumor location	
Oral cavity	129 (42.9)
Oropharynx	50 (16.6)
Larynx/hypopharynx	38 (12.6)
Sinonasal	6 (2.0)
Unknown primary	8 (2.6)
Major salivary gland	20 (6.6)
Facial cutaneous malignancy	50 (16.6)
Cancer treatment	
Surgery	115 (38.2)

(continued)

adjuvant therapy (186/301), and 60.1% underwent microvascular reconstruction (181/301).

CFA

CFA supported the unidimensionality of the 5 hypothesized domains of HNC-related BID: OOA, SA, PDA, AC, and DFI.[13] The final models for each of the 5 specific domains demonstrated acceptable model fit indices (**Table 2**). The initial CFA revealed that 2 domains, DFI and PDA, had local dependence concerns. After removing 2 items from the DFI domain and 1 from the PDA domain, the unidimensionality of all domains was established. All items in the 5 IMAGE-HN domains had a significant relationship with the latent construct (as determined by standardized factor loadings $\geq|0.32|$; see Suppl. Table S1 in the online version of the article) and were locally independent (item residual correlations $<|0.20|$; data not shown).

Rasch Analysis

A summary of the Rasch analysis results is provided in **Table 3**. Examination of the rating scale indicated acceptable fit (infit and outfit mean squares <2.0) and monotonicity of all rating scale categories. All items in all domains demonstrated adequate fit to the Rasch model as evidenced by infit mean squares <1.60 and standardized residuals (standardized z values) <2.0. Person fit demonstrated low levels of significant misfit to the model in any domain (2% in SA and AC, 4% in DFI). Floor effects ranged from 12% (DFI domain) to 36% (PDA domain); ceiling effects were minimal. Separation indices and person reliability were acceptable or better for each domain except AC (separation index, 0.82; person reliability, 0.40). Strata calculation revealed that the OOA, PDA, and DFI domains separated persons into 2 groups—those who had higher and those who had lower scores on each construct. Because of its inadequate separation index, person reliability, and strata calculations, the AC domain was removed from IMAGE-HN as an individual domain and component of the global domain. The Rasch item measures and fit statistics for each individual domain are shown in Supplemental Table S2 (in the online version of the article). Overall, the mean person measures (ranging from 19.41-33.44) were substantially lower than the mean item measure (anchored at a score of 50), which suggests the present sample was skewed toward having low scores (see Suppl. Figure S1 in the online version of the article).

Global Domain

A global IMAGE-HN domain consisting of items from the OOA, SA, PDA, and DFI domains was analyzed with a single-factor CFA. Local dependence was observed for 1 item from the SA domain and 2 items from the DFI domain. These 3 items were removed, leading to improved fit indices in the final model for the global domain (**Table 2**). No local dependence was observed (ie, all residual correlations were $<|0.2|$), and all items had standardized factor loadings $>|0.32|$ on the global domain (see Suppl. Table S1 in the online version of the article).

Table 1. (continued)

Characteristic	No. (%)
Surgery and adjuvant radiation	98 (32.6)
Surgery and adjuvant chemoradiation	88 (29.2)
Reconstructive surgery	
None	102 (46.2)
Local or regional flap	13 (4.3)
Microvascular free flap	181 (60.1)
Other	5 (1.7)
Osseous microvascular free flap reconstruction	
No	259 (86.0)
Yes	42 (14.0)
Time since completion of treatment	
0-6 months	119 (39.5)
6-12 months	49 (16.3)
12-24 months	52 (17.3)
>24 months	81 (26.9)
Academic medical center	
Baylor College of Medicine	39 (13)
Medical University of South Carolina	206 (68.4)
University of Pittsburgh Medical Center	28 (9.3)
Washington University School of Medicine	28 (9.3)

[a]N = 301 due to missing information.
[b]N = 300 due to missing information.
[c]N = 299 due to missing information.

Rasch analysis of the global domain revealed the rating scale met all criteria for acceptable fit, number of observations, and monotonicity of rating scale categories (**Table 3**). Approximately 6.5% of participants demonstrated significant misfit to the model. Person reliability was adequate (0.80), and the strata calculation revealed that the global domain separated respondents into 3 statistically distinct groups—those with high, medium, and low HNC-related BID. While no ceiling effects were observed, approximately 10% of participants achieved the lowest possible score. The Rasch item measures and fit statistics for the global domain are in Supplemental Table S2 and Supplemental Figure S1 (in the online version of the article).

Discussion

Herein we describe the creation and multi-institution validation of the IMAGE-HN, a novel, psychometrically sound, 24-item, multidomain PROM consisting of 4 subscales and a global scale that can be used to measure key aspects of HNC-related BID[13] in clinical and research settings (see Suppl. Table S3 for IMAGE-HN scoring information in the online version of the article). Although BID is a common in patients with HNC and a key driver of psychosocial morbidity,[2] to date, clinicians have lacked tools to identify patients with HNC-related BID in a clinical setting. In addition, significant gaps in the epidemiology, prevention, and management of BID in patients with HNC remain,[3] in part, because of the lack of a validated, psychometrically robust PROM of HNC-related BID in the research setting.[3,11] Without such a tool,

Table 2. Confirmatory Factor Analysis Model Fit Statistics for Each IMAGE-HN Domain and the Global Domain.

Fit Statistic	Other-Oriented Appearance	Social Avoidance and Isolation	Personal Dissatisfaction with Appearance	Appearance Concealment	Distress with Functional Impairments	Global
Comparative fit index	0.93	0.97	0.97	0.96	0.95	0.84
Tucker-Lewis index	0.90	0.95	0.92	0.88	0.93	0.82
Root mean square error of approximation	0.06	0.05	0.10	0.08	0.08	0.06
Standardized root mean square residual	0.03	0.03	0.01	0.03	0.04	0.05

Abbreviation: IMAGE-HN, Inventory to Measure and Assess imaGe disturbancE–Head & Neck.

Table 3. Rasch Analysis Results Summary for Each IMAGE-HN Domain and Global Domain.

Criterion	Other-Oriented Appearance	Social Avoidance and Isolation	Personal Dissatisfaction with Appearance	Appearance Concealment	Distress with Functional Impairments	Global
Item fit	7/7	6/6	4/4	4/4	7/7	20/21
Person misfit, %	3.0	2.3	4.2	2.3	3.9	6.6
Ceiling, %	0.3	0.3	1.3	0.7	0.7	0.3
Floor, %	31.8	35.4	36.4	54.4	12.1	9.8
Mean person measure	19.69	19.41	22.86	32.48	33.44	29.76
Separation index	1.34	1.09	1.70	0.82	1.57	2.01
Person reliability	0.64	0.55	0.74	0.40	0.71	0.80
No. of person strata	2.12	1.78	2.60	1.43	2.43	3.01
Cronbach's α	0.91	0.89	0.91	0.76	0.88	0.95

Abbreviation: IMAGE-HN, Inventory to Measure and Assess imaGe disturbancE–Head & Neck.

scientists have been limited in their ability to advance our understanding of the epidemiology (eg, diagnostic characteristics, risk factors, temporal trajectory) of this disorder, accurately identify patients with HNC with clinically meaningful image-related distress, recruit patients of comparable BID severity into trials to test novel interventions to manage BID, or assess the effectiveness of interventions in this population. The IMAGE-HN thus addresses a critical unmet need for measuring image concerns in patients with HNC in the clinical and research realms.

PROMs enhance the delivery of patient-centered HNC care[22] and result in improved symptom control, higher levels of patient satisfaction, and decreased mortality when incorporated into clinical oncology practice.[23,24] The paradigm shift toward measuring BID using PROMs, which harness the patient's words and experiences, has been reinforced by the recognition that BID in patients with cancer is not driven primarily by objective measures of disfigurement.[9,10] The IMAGE-HN can facilitate patient-centered care by providing clinicians with a comprehensive and accurate method of quantifying information about the multidimensional nature of HNC-related BID.

The IMAGE-HN addresses a number of methodologic and content deficiencies in existing PROMs used to assess image concerns in patients with HNC. To date, researchers have relied on PROMs developed for and validated in non-HNC patient populations to assess BID among HNC survivors.[11] For example, the Body Image Scale, the most widely used PROM to assess BID in patients with HNC,[11] was developed based on qualitative work with patients with breast cancer and providers and validated in patients with breast cancer.[14] The lack of content validity of the Body Image Scale for HNC-related BID manifests in its failure to address key HNC image concerns (eg, eating in public, challenges speaking) and its inclusion of items unrelated to HNC-related BID (eg, ''Did you find it difficult to look at yourself naked?''). The IMAGE-HN, with its 4 subscales and a global domain confirmed by CFA that map to the conceptual domains of HNC-related BID identified in our qualitative work,[13] provides for a robust and comprehensive assessment of the diverse and heterogeneous facets of BID experienced by patients with HNC.

Recently, the McGill Body Image Concerns Scale–Head and Neck Cancer (MBIS-HNC), a novel PROM to measure image concerns in patients with HNC, was published.[25] This validated 2-domain instrument was based on extensive qualitative work,[26] underwent rigorous psychometric analysis,[25] and represents a significant improvement of prior PROMs to measure HNC-related BID. However, a few limitations that may hinder its utility. Patients were recruited from a single academic medical center, limiting external validity of study findings. In addition, the most common subsite/histology among subjects was facial skin cancer, which is not representative of the target HNC patient population. CFA model fit indices reflected inadequate model fit with the 3-factor solution, and the final MBIS-HNC consists of only 2 of the 4

originally hypothesized domains, suggesting that it may not be measuring key aspects of HNC-related BID.[13,26,27] The method by which the authors progressed from the 3-factor CFA solution to 2 domains in the Rasch analysis is also not readily apparent. In contrast, the IMAGE-HN was validated using a multi-institutional cohort in a representative sample of patients with HNC. The IMAGE-HN comprehensively measures multiple domains that are of high priority to patients with HNC in a fashion that is psychometrically robust.

The major strengths of this study include its extensive qualitative background, rigorous approach using PROMIS methodology, multi-institutional validation, and robust Rasch psychometric analysis. Nevertheless, the study has limitations. First, Rasch analysis confirmed an important floor effect for the IMAGE-HN. Whether the floor effect is a result of an intrinsically psychometrically imperfect instrument or the result of the validation sample being skewed toward less severe HNC-related BID cannot be known. Future work identifying clinically relevant values for the IMAGE-HN domains and global score will help address this potential limitation. Although the IMAGE-HN is a psychometrically sound PROM capable of separating patients into person strata based on the latent construct, further research is necessary to determine clinically relevant minimal important differences between individuals at a given point in time as well as the association of demographic and oncologic characteristics with BID severity as measured by the IMAGE-HN. In addition, longitudinal studies will be necessary to determine clinically meaningful changes in IMAGE-HN scores for an individual over time. Although our study was multi-institutional in nature and included patients diverse by demographic and clinical characteristics, a single institution (MUSC) accounted for two-thirds of the patient accrual, which could limit external validity. Further studies validating the IMAGE-HN in other geographic regions and different cultures will therefore be critically important. While the tablet-based completion of the IMAGE-HN was feasible, data about time required to complete the PROM, optimal method of delivery, and strategies for integration into the clinical setting are all lacking and should be addressed in the future.

Conclusions

The IMAGE-HN is a novel, psychometrically sound, 24-item, multidomain PROM for assessing BID in patients with HNC that has been validated in a multi-institutional cohort. The IMAGE-HN, consisting of 4 subscales and a global scale, addresses a critical unmet need for addressing multidimensional image concerns in patients with HNC. It can be successfully and consistently used across multiple institutions and clinical settings in the pre- and posttreatment periods to facilitate the delivery of patient-centered HNC care through the identification of patients with HNC-related BID requiring supportive intervention. The IMAGE-HN provides researchers with a more precise and accurate tool that will be essential to future studies addressing gaps in our knowledge about the epidemiology, prevention, and management of HNC-related BID in an attempt to minimize psychosocial morbidity for HNC survivors and improve QOL.

Acknowledgments

We thank W. Greer Albergotti, MD, Vasu Divi, MD, Jeff Houlton, MD, Sobia Khaja, MD, Brian Nussenbaum, MD, MHCM, Nitin Pagedar, MD, MPH, Megan Scallion, MSW, MPH, Tiffany K. Garris, NP-C , and Joseph Zenga, MD, for their critical feedback on the question-writing portion of the scale development.

Author Contributions

Evan M. Graboyes, substantial contributions to the conception or design of the work; acquisition, analysis, and interpretation of data for the work; drafting the work and revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Brittany N. Hand**, substantial contributions to the conception or design of the work; acquisition, analysis, and interpretation of data for the work; drafting the work; final approval of the version to be published; agreement 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; **Mark A. Ellis**, substantial contributions to the conception or design of the work; acquisition, analysis, and interpretation of data for the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Andrew T. Huang**, substantial contributions to the conception or design of the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Marci L. Nilsen**, substantial contributions to the conception or design of the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Patrik Pipkorn**, substantial contributions to the conception or design of the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Courtney H. Marsh**, substantial contributions to the conception or design of the work; interpretation of data for the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Stacey Maurer**, substantial contributions to the conception or design of the work; interpretation of data for the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Terry A. Day**, substantial contributions to the conception or design of the work; interpretation of data for the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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; **Katherine R. Sterba**, substantial contributions to the conception or design of the work; acquisition, analysis, and interpretation of data for the work; revising it critically for important intellectual content; final approval of the version to be published; agreement 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.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: ACS IRG-16-185-17 from the American Cancer Society (ACS) to Evan M. Graboyes; AAO-HNS/F CORE 577381 from the American Academy of Otolaryngology–Head and Neck Surgery Foundation to Mark A. Ellis; database support from the South Carolina Clinical and Translational Research Institute (UL1TR000062). No role in the design and conduct; collection, analysis, and interpretation of the data; or writing or approval of the manuscript.

Supplemental Material

Additional supporting information is available in the online version of the article.

References

1. National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Head and Neck Cancers*. National Comprehensive Cancer Network; 2019.
2. Fingeret MC, Teo I, Goettsch K. Body image: a critical psychosocial issue for patients with head and neck cancer. *Curr Oncol Rep*. 2015;17:422.
3. Rhoten BA, Murphy B, Ridner SH. Body image in patients with head and neck cancer: a review of the literature. *Oral Oncol*. 2013;49:753-760.
4. Rhoten BA, Deng J, Dietrich MS, Murphy B, Ridner SH. Body image and depressive symptoms in patients with head and neck cancer: an important relationship. *Support Care Cancer*. 2014; 22:3053-3060.
5. Wu YS, Lin PY, Chien CY, et al. Anxiety and depression in patients with head and neck cancer: 6-month follow-up study. *Neuropsychiatr Dis Treat*. 2016;12:1029-1036.
6. Osazuwa-Peters N, Simpson MC, Zhao L, et al. Suicide risk among cancer survivors: head and neck versus other cancers. *Cancer*. 2018;124:4072-4079.
7. Fingeret MC, Vidrine DJ, Reece GP, Gillenwater AM, Gritz ER. Multidimensional analysis of body image concerns among newly diagnosed patients with oral cavity cancer. *Head Neck*. 2010;32:301-309.
8. Fingeret MC, Yuan Y, Urbauer D, Weston J, Nipomnick S, Weber R. The nature and extent of body image concerns among surgically treated patients with head and neck cancer. *Psychooncology*. 2012;21:836-844.
9. Pezner RD, Lipsett JA, Vora NL, Desai KR. Limited usefulness of observer-based cosmesis scales employed to evaluate patients treated conservatively for breast cancer. *Int J Radiat Oncol Biol Phys*. 1985;11:1117-1119.
10. Sneeuw KC, Aaronson NK, Yarnold JR, et al. Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer: 1. Comparison of patients' ratings, observers' ratings, and objective assessments. *Radiother Oncol*. 1992; 25:153-159.
11. Ellis MA, Sterba KR, Brennan EA, et al. A systematic review of patient-reported outcome measures assessing body image disturbance in patients with head and neck cancer. *Otolaryngol Head Neck Surg*. 2019;160:941-954.
12. PROMIS® Instrument Development and Validation Scientific Standards Version 2.0. Periodical [serial online]. 2013. http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf. Accessed December 3, 2019.
13. Ellis MA, Sterba KR, Day TA, et al. Body image disturbance in surgically treated head and neck cancer patients: a patient-centered approach. *Otolaryngol Head Neck Surg*. 2019;161:278-287.
14. Hopwood P, Fletcher I, Lee A, Al Ghazal S. A body image scale for use with cancer patients. *Eur J Cancer*. 2001;37:189-197.
15. Mullin PA, Lohr KN, Bresnahan BW, McNulty P. Applying cognitive design principles to formatting HRQOL instruments. *Qual Life Res*. 2000;9:13-27.
16. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45:S22-S31.
17. Tabachnick BG, Ridell LS. *Using Multivariate Statistics*. Pearson Education; 2013.
18. Christensen KB, Kreiner S, Mesbah M. *Rasch Models in Health*. John Wiley; 2013.
19. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measure Trans*. 1994;8:370-371.
20. Linacre JM. Winsteps Rasch measurement computer program: user's guide 2016. Accessed September 12, 2019 .https://www.winsteps.com/winman/reliability.htm
21. Wright B, Masters G. Number of person or item strata. *Rasch Measure Trans*. 2002;16:888.
22. Ringash J, Bernstein LJ, Cella D, et al. Outcomes toolbox for head and neck cancer research. *Head Neck*. 2015;37:425-439.
23. Basch E. Patient-reported outcomes—harnessing patients' voices to improve clinical care. *N Engl J Med*. 2017;376:105-108.
24. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol*. 2016;34:557-565.
25. Rodriguez AM, Frenkiel S, Desroches J, et al. Development and validation of the McGill Body Image Concerns Scale for use in head and neck oncology (MBIS-HNC): a mixed-methods approach. *Psychooncology*. 2019;28:116-121.
26. Henry M, Ho A, Lambert SD, et al. Looking beyond disfigurement: the experience of patients with head and neck cancer. *J Palliat Care*. 2014;30:5-15.
27. Albornoz CR, Pusic AL, Reavey P, et al. Measuring health-related quality of life outcomes in head and neck reconstruction. *Clin Plast Surg*. 2013;40:341-349.