CONTENT VALIDITY OF FACE-Q FOR PEDIATRIC HEAD AND NECK CANCER

ESTABLISHING CONTENT VALIDITY OF THE FACE-Q CRANIOFACIAL MODULE FOR PEDIATRIC HEAD AND NECK CANCER

By YI WANG, B.H.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University

McMaster University MASTER OF SCIENCE (2020) Hamilton, Ontario

TITLE: ESTABLISHING CONTENT VALIDITY OF THE FACE-Q CRANIOFACIAL MODULE FOR PEDIATRIC HEAD AND NECK CANCER

AUTHOR: Yi Wang, B.H.Sc. SUPERVISOR: Dr. Anne Klassen

NUMBER OF PAGES: x, 44

Lay abstract

The FACE-Q is a patient-reported outcome measure developed to assess outcomes important to patients aged 8-29 years with craniofacial conditions. The current study aimed to determine its content validity for use in patients with head and neck cancer (HNC). Cognitive interviews with patients with HNC aged 8-29 years (n=15) were conducted and feedback from experts in pediatric oncology (n=21) was obtained. A total of 1573 codes from patient comments and 234 codes from expert feedback were developed. A total of 12 items were flagged for review from qualitative interviews along with 4 items from expert feedback among the core scales for comprehensibility. Instructions and response options were found to be comprehensible and appropriate. A total of 10 missing items were identified across the core scales by patient participants while experts identified 1 missing item. The FACE-Q evidenced content validity for core scales along with limited evidence for remaining scales.

Abstract

Objective: Existing patient-reported outcome measures (PROM)s for patients with **facial differences** lack content validity, as few items address appearance and function issues. The FACE-Q is a new PROM developed to measure outcomes important to patients aged 8-29 years with craniofacial conditions. A process was needed to determine if the FACE-Q content is relevant to patients with head and neck cancer (HNC).

Methods: Cognitive interviews with patients with HNC aged 8 to 29 years (n=15) were conducted and feedback from experts in pediatric oncology (n=21) was obtained. Input was sought on all aspects of the FACE-Q content.

Results: A total of 1573 codes were developed from patient comments and 234 codes were developed from expert feedback that related to the COSMIN criteria for judging content validity. A total of 12 items were flagged for review from qualitative interviews and 4 comments were coded from expert feedback among the core scales for comprehensibility. Instructions, time frame, and response options were found to be comprehensible and appropriate by almost all patient and expert participants. Participants identified a total of 10 missing items identified across the core scales, while no additional items were identified by experts for the core scales. However, 4 experts identified swallowing/dysphagia as an important item missing from the mouth function scale.

Discussion: Content validity of the FACE-Q for patients with HNC was evaluated through cognitive interviews with patients and feedback from pediatric oncology experts. The core scales were answered by all participants and demonstrate overall content validity from feedback offered by both patients and experts.

Conclusion: The FACE-Q showed evidence of content validity for its core scales along with limited evidence that the remaining scales covered issues relevant to specific HNC patients.

Assessment of the psychometric properties of the new measure is forthcoming as part of an international FACE-Q field-test study.

Acknowledgments

I would first like to thank my thesis supervisor, Dr. Anne Klassen, for her continuous support throughout my thesis work over the years. I am truly appreciative of her wisdom as a mentor, providing me with guidance throughout a research process I had never experienced before. I thank her for believing in me when I was unsure of my next steps, for always challenging me to reach my potential, and to seize new opportunities as they arise. I am very grateful to Dr. Klassen for molding me not only in academia, but also in my personal development.

I would also like to thank my thesis committee members, Drs. Sophocles Voineskos and Joyce Obeid, for their feedback and guidance. Their expertise and critical insight helped to guide and refine the writing of this thesis. I also thank Dr. Claudia Malic for her supportive role as external examiner.

Many thanks go to the experts who were involved in the preliminary work, research process, and completion of this thesis. In particular, the support and guidance offered by Dr. Elena Tsangaris and Charlene Rae are greatly appreciated. Thank you for always sharing your knowledge and expertise, as well as guiding me through the qualitative interview and thesis-writing process.

Finally, I must express my gratitude to my parents, friends, and Forrest, for providing me with their unwavering support and continuous encouragement throughout my years of study. This accomplishment would have not been possible without them.

Table of Contents

List of Figures

List of Tables

List of Abbreviations:

PROM: Patient-reported outcome measure

HNC: Head and neck cancer

COSMIN: COnsensus-based Standards for the selection of health status Measurement INstruments

HRQOL: Health-related quality of life

QOL: Quality of life

REDCap: Research Electronic Data Capture

FACE-Q: FACE-Q Craniofacial Module

Declaration of Academic Achievement

The content of this document has been completed by Yi Wang. The thesis describes the process of establishing content validity of the FACE-Q Craniofacial Module for use in patients with head and neck cancer. The content focuses on the methodology and results obtained from qualitative research with patients with facial differences from head and neck cancer, as well as experts in the field of pediatric oncology. Yi Wang contributed to the qualitative interviews, transcription, coding, analysis, and writing of the paper. Identification and contact of patient participants were done by healthcare professionals at the hospital sites. Therefore, Dr. Nathan Paul, Dr. Eric Bouffet, Dr. David Dix, Dr. Stacey Marjerrison, Julie Constantin, JoAnn Duckworth, Loretta Anderson, Octavia Choi, are greatly appreciated for their generous help and support through this process. The contributions of Dr. Elena Tsangaris, Trisia Breitkopf, and Dr. Anne Klassen in the data analysis portion of this thesis are recognized and appreciated.

Section 1. Introduction

1.1 Head and neck cancer and its effect on facial differences

It is reported that approximately 650,000 new head and neck cancer (HNC) cases are diagnosed each year, making it the 6th most common type of cancer in the world (Parkin, Bray, Ferlay, & Pisani, 2005). Although the typical HNC patient is older, male, and often has comorbidities associated with smoking and drinking behaviours, there have been increased reports in recent years of HNC occurring in younger populations (<45 years of age) (Curado & Hashibe, 2009; Marur & Forastiere, 2008; Modh, Gayar, Elshaikh, Paulino, & Siddiqui, 2018). When data from several European countries were analyzed for squamous cell carcinoma (SCC) incidence among young adults (aged 20-39), it was found that 5.5% occurred among young adults (aged 20-39) between 1960 and 1994 (Annertz et al., 2002). The incidence of oral tongue SCC was found to have increased 5-fold among young men and 6-fold among young women, while only a 2-fold increase was observed in older age groups. Most common HNC diagnoses to occur in the pediatric population include lymphomas, rhabdomyosarcoma, and nasopharyngeal carcinoma (Cesmebasi et al., 2014; Sengupta et al., 2009). Given the complexity of child development and the importance of childhood years on cognitive, social, emotional, physical and behavioural development, treatment outcome cannot be solely measured by survival rate in this population. The diagnosis and treatment of HNC can have a serious impact on emotional, psychological, and physical health of individuals (Vickery, Latchford, Hewison, Bellew, & Feber, 2003). Children and youth, in particular, can experience long-term and late effects of childhood cancer, such as emotional effects including anxiety, depression, the fear of cancer returning, and the avoidance of healthcare into adulthood.

HNC and its various treatments can both lead to changes in facial appearance and facial function, impacting individuals’ health-related quality of life (HRQOL). A diagnosis of cancer itself can elicit fears of changes in self-image, confidence, and identity (Jones, Lund, Howard, Greenberg, & McCarthy, 2007). Patients undergoing surgery as a part of their treatment may experience visible facial differences that are difficult to conceal, such as not being able to show expression due to facial paralysis. Patients undergoing radiotherapy, brachytherapy, or chemotherapy and radiation must also endure differences in facial function, including soreness and dryness of the mouth and neck, leading to difficulties in eating and drinking (Krouse, Krouse, & Fabian, 1989). Other symptoms may include swelling, residual pain, decreased sensation, hoarseness, ear pain, enlarged cervical lymph nodes, nasal bleeding or blockage, and sores/ulcers in the mouth (Brockstein & Masters, 2010; List & Bilir, 2004; Marur & Forastiere, 2008). Adverse effects from such forms of HNC treatment can lead to changes in self-perception, perception of others, as well as limitations in social interactions and feelings of isolation (Herzon & Boshier, 1979; Krouse et al., 1989). Consequently, HNC survivors often require rehabilitative treatment following cancer treatment, including speech therapy, swallowing rehabilitation, maxillofacial rehabilitations, along with physical and occupational therapies (Eades, Chasen, & Bhargava, 2009).

Due to the visible nature of the disease and treatment, HNC has been argued to be the most psychologically traumatic cancer to experience (Björklund, Sarvimäki, & Berg, 2010; Howren, Christensen, Hynds Karnell, Van Liew, & Funk, 2013; Koster & Bergsma, 1990). Perceptions of facial appearance can have a strong impact on the social and emotional development of individuals (Langlois et al., 2000). Thus, differences in facial appearance can adversely shape personal development and expose individuals to risks for psychosocial difficulties and other impaired HRQOL (Topolski, Edwards, & Patrick, 2005). Such difficulties can arise from being stared at, avoided, and teased (Lawrence, Rosenberg, Mason, & Fauerbach, 2011; Strauss et al., 2007). The effect of visible facial differences on children and youth's

HRQOL is suggested to include difficulties with self-perception, emotional problems, and difficulty with social functioning (Hunt, Burden, Hepper, & Johnston, 2005; Stubbs et al., 2011; Topolski et al., 2005). Adverse social situations, such as bullying, have been shown to be associated with negative effects on mental health (Hunt, Burden, Hepper, Stevenson, & Johnston, 2007; Masnari et al., 2013).

1.2 Existing patient-reported outcome measures for patients with head and neck cancer

In order to capture and understand HRQOL in pediatric patients with HNC, a valid, reliable, and responsive patient-reported outcome measure (PROM) specific to this patient population is needed to measure outcomes from the patient perspective. Terwee et al. argued in the COSMIN Methodology for Assessing the Content Validity of PROMS User Manual that content validity is the most important measurement property of a PROM and the most challenging to assess (Terwee et al., 2018). Content validity is an indicator of how well the content of an instrument reflect the constructs the questionnaire aims to measure. In order to assess content validity, the COSMIN User Manual, the current gold standard for assessing content validity of PROMs, recommends systematically asking patients and professionals in the field about the relevance, comprehensiveness, and comprehensibility of the items. By understanding challenges patients have with their HRQOL, treatment options can be evaluated, and more suitable rehabilitative services could be offered, along with patient education materials and other sources of support. While there are various quality of life (QOL) measures available for pediatric oncology patients, a systematic review conducted by Anthony et al. revealed that these measures do not address facial differences in appearance and function, which are important aspects of HRQOL for the pediatric HNC population (Anthony et al., 2014).

A systematic review of existing HNC quality of life (QOL) assessment instruments by Ojo et al. highlighted the heterogeneity of measures used for measuring HNC QOL (Ojo et al., 2012). The authors found that the diversity and abundance of the many instruments available made it difficult to make sense of the findings across studies, even when content overlapped in certain dimensions (e.g. fatigue, mobility, sleep, and pain). Of the 57 total instruments identified, only 10 scales measured physical appearance, which can be dramatically altered by HNC and its treatment. Moreover, none of the 57 instruments included measured all the potential domains associated with HNC and its treatment identified through qualitative interviews with 26 patients with HNC (Albornoz et al., 2013). Missing important domains associated with HRQOL of HNC patients, such as facial appearance, facial function, and psychosocial impact, leads to a lack of content validity from inadequate comprehensive coverage of important domains.

A systematic review conducted by Wickert et al. identified 17 existing PROMs available for measuring HRQOL in pediatric patients with facial differences (Wickert et al., 2018) . The authors showed that PROMs used for children/youth with facial differences also had poor content validity. Of the 554 items included in the 17 instruments, only 19 measured facial appearance and 22 measured facial function. Since facial appearance and function are 2 important constructs to patients with facial differences due to HNC, this finding shows a lack of content validity for this patient population. To address this lack of content validity in available PROMs used for measuring the HRQOL in children and youth with facial differences, a new module, tentatively named the FACE-Q Craniofacial Module, was developed for children and youth with diverse conditions associated with visible and/or functional facial differences (Longmire et al., 2017). Since this new FACE-Q Craniofacial Module was originally designed for children and young adults with a broad range of craniofacial conditions, the current study aims to establish whether this instrument has content validity to be used for pediatric HNC patients with facial differences. It is hypothesized that while not all of the scales within the FACE-Q Craniofacial Module may be directly relevant to individuals with HNC, the instrument should be broadly applicable to and show evidence of content validity for this population.

Section 2. Methodology

The current study aims to assess the content validity of the FACE-Q Craniofacial Module for use in patients who had pediatric HNC using qualitative methods. The study was approved by applicable research ethics boards at participating hospitals in Ontario and British Columbia, Canada (Hamilton Integrated Research Ethics Board in Hamilton, The Hospital for Sick Children Research Ethics Board in Toronto, and the University of British Columbia Children's & Women’s Health Centre of British Columbia Research Ethics Board in Vancouver). Interviews and surveys were conducted in accordance with policies for ethical research involving human participants. Written and verbal informed consent were obtained for all participants over the age of 18. Assent for participants under 18 years of age as well as parent/guardian consent were also obtained.

2.1 Participants

We aimed for a consecutive sample at 3 pediatric cancer centres in Ontario and British Columbia (McMaster Children’s Hospital, The Hospital for Sick Children, BC Children’s Hospital) to recruit for the study. Inclusion criteria were as follows: Individuals between the ages of 8 and 29 who had a visible and/or functional facial difference as a result of pediatric head and neck cancer and/or its treatment; English speaking; and were able to participate in the cognitive interview (read, understand and discuss the scales). The aim of recruiting all eligible participants allowed for purposeful sampling and the ability to recruit a diverse sample in terms of age, gender, cancer diagnoses, and treatment.

In addition to patients, experts in the field of pediatric oncology in Canada, the United States, and the Netherlands were invited to participant in a survey through the research team's professional network of contacts.

2.2 Recruitment Procedure

Eligible participants were first identified by a member of the healthcare team at each recruitment site between June 2017 and August 2019. Potential participants were introduced to the study in person by a member of their healthcare team. Patients who expressed interest in participating in the study were then approached in person, by phone, or through email by a member of the research team to obtain informed consent and to schedule an interview. Informed consent was obtained from participants over the age of 18 at The Hospital for Sick Children and BC Children's Hospital, informed assent and consent from guardians were obtained for participants under the age of 18 at these sites. At McMaster Children's Hospital, informed consent was obtained from participants aged 16 and over, while informed assent and consent from guardians were obtained for participants under the age of 16. Interviews were conducted either in person or, for participants aged 12 years and older, by phone depending on participants' preference. A $50 prepaid Visa gift card was provided to participants to thank them for the time dedicated to the cognitive interview.

Expert feedback was obtained between September 2017 and November 2017. Invites were sent via email and feedback on the FACE-Q scales was obtained through a secure web-based survey via Research Electronic Data Capture (REDCap). A follow-up reminder was sent 7 days after the initial email invite.

2.3 The Instrument - The FACE-Q Craniofacial Module

The FACE-Q Craniofacial Module (FACE-Q) is a new PROM that measures appearance, facial function, and HRQOL of patients aged 8-29 years with craniofacial conditions. Development of the FACE-Q began with the development of the CLEFT-Q, a PROM designed to measure the HRQOL of children and youth with cleft lip/palate, the most common craniofacial anomaly. The CLEFT-Q was developed from qualitative data collected from 138 patients with cleft lip and/or palate (CL/P) from 6 countries and then field-tested with 2434 patients aged 8 to 29 years from 12 countries. The final version contains 12 scales and 1 checklist that measure appearance (face, jaws, lips, teeth, nose, nostril, cleft scar), function (speaking, eating/drinking), and HRQOL (speech distress, psychological, social and school function).

Semi-structured interviews were then conducted with 72 participants with ear anomalies, facial paralysis, skeletal conditions, or soft tissue conditions to develop new scales to cover concerns not measured by the CLEFT-Q. In addition, to determine if the CLEFT-Q scales were relevant to non-cleft craniofacial conditions, at the end of the interview, cognitive interviews were conducted with 69 of the participants who agreed to take part. All participants were asked to review 4 core CLEFT-Q scales as these were hypothesized to be relevant to all facial conditions (face, psychological, social and school function). In addition, participants were shown additional CLEFT-Q scales hypothesized to be relevant to specific subgroups of patients based on their condition. A total of 11 CLEFT-Q scales were determined to have content validity for use with other facial conditions (exception cleft lip scale).

The conceptual framework for the FACE-Q is outlined in Figure 1. Following the development of 13 new scales, participants (n=84) and experts (n=43) were invited to review the new scales to suggest missing content and provided feedback on the items, instructions and response options. The full set of 26 independently functioning scales is collectively called the FACE-Q. Of the 26 scales, the birthmark scale has been excluded for the current study due to its irrelevance to HNC patients.

2.4 Establishing Content Validity

Cognitive interviews were conducted with participants, which involved one-on-one semi-structured interviews with 1 of 3 trained qualitative interviewers using a cognitive interview guide (see Table 1). The cognitive interviewing approach used was adapted from Willis and is summarized in Table 2 (Willis, 2015). Objectives for cognitive interviewing were to understand participant comprehension of item wording, instructions, response options, as well as to identify potentially irrelevant, missing, or redundant content. Challenges with participant interpretation of the items, instructions, and response options were also noted. The "think aloud" approach, a recommended method for cognitive debriefing, was used so participants could verbalize their thoughts and interpretation of the scales (Willis, 2004). Verbal probing was used concurrently by the interviewers to further understand why certain aspects of scales might be unclear or difficult to answer. For instructions, items, and response options found to be ambiguous or difficult to answer by participants, participants were encouraged to suggest revisions to the wording. A set of 6 core scales were reviewed by all participants (face appearance, adverse treatment effects, appearance distress, psychological, school, and social). Additional scales (Table 3) were reviewed if applicable to participants based on their facial difference. All interviews, whether conducted in person or by phone, were audio recorded and transcribed verbatim before being coded and analyzed.

For expert feedback, a comment section was available at the end of each scale and experts were invited to provide written feedback on all aspects of the scales. Specifically, experts were invited to comment on the relevance and comprehensibility of the instructions, response options, and items of the FACE-Q to pediatric patients with HNC, as well as any missing content to be added to the scales.

Patient interview and expert feedback data were coded line-by-line by YW and all codes were entered into Microsoft Excel (2016) spreadsheets for analysis. All feedback was compiled and reviewed by the research team in order to determine the extent to which scales have content validity for use with patients with pediatric HNC.

Section 3. Results

3.1 Sample Characteristics

A total of 15 patients participated in cognitive interviews. The mean age of participants at the time of recruitment was 16 years (range 10 to 30 years), with most participants having completed treatment over 5 years prior to participating in the current study. The sample included 9 females and 6 males. Most participants were diagnosed with rhabdomyosarcoma, which is the primary HNC diagnosis in children. Experts in the field of pediatric oncology (N=50) were invited to participate in the study though an email invite via REDCap and 21 experts provided feedback. Most experts were Canadian pediatric oncologists. Patient participant demographics are summarized in Table 4, expert demographics are summarized in Table 5.

Qualitative feedback from all participants was obtained for the core scales (face appearance, face adverse effects, appearance distress, psychological, school, and social). A total of 17 additional scales were completed when found to be relevant to the participants' condition. Table 6 shows the total number of participants who provided feedback for each FACE-Q scale.

During both cognitive interviews and expert feedback, 1573 codes were developed from patients comments and 234 codes were developed from expert comments that related to the COSMIN criteria for judging content validity: comprehensiveness, comprehensibility, and relevance. These codes were flagged for review by the team.

3.2 Comprehensibility

The comprehensibility of the FACE-Q by patients with pediatric HNC was tested through participant understanding and appropriateness of its instructions (including recall period), items, response options.

3.2.1 Items

For the face appearance scale, 160 comments were coded for participant definitions of items, 50 comments were coded for difficulties participants had with items and the reasons. For the face adverse effects scale, 69 comments were coded for participant definitions of items and 12 comments were coded for difficulties with item comprehension and/or answering. The appearance distress scale had 71 codes for participant definitions of items and 9 codes for difficulties. A total of 51 participant definitions were coded for the psychology scale, along with 15 difficulties participants had. The school scale had 84 item definitions coded along with 3 difficulties. Lastly, the social scale had 110 codes for definitions and 8 difficulties coded. Items reviewed as part of the core set of scales and example quotes for interpretation are included in Table 7. Items flagged as unclear or difficult to answer in the core scales are summarized in Table 8.

Expert feedback was also obtained regarding the comprehensibility of the items. For the face appearance scale,1expert wondered whether “you need to ask about left and right side in

[the] profile question?" [EXP024]. This is currently unspecified in the question. For the appearance distress scale,1expert recommended rephrasing question 7 to "I avoid sharing photos of myself on social media" [EXP011] due to common social media applications frequently changing. As well, an expert recommended changing to word "avoid" to "try not to" and "dislike" to "do not like" to lower the reading level of certain items [EXP022]. Several experts commented on whether certain words in the face adverse effects scale might be to difficult for children and youth to comprehend, such as "tight", "tender" and "sore" [EXP004, 014].

3.2.2 Instructions and Response Options

The 15-item face appearance scale asked participants to answer thinking of how they look now, how much they liked their face (not at all, a little, quite a bit, very much). All 10 participants who provided feedback noted that the instructions were clear and "pretty self-explanatory" [COG005]. As well, the instructions were interpreted as intended (e.g. "I have to answer these following questions based on how much I like my features…" [COG008]). Most participants (7) thought the response options were good with "clear distinct categories" [COG016], while 2 participants would have preferred to have another response option between "a little bit" and "quite a bit". Most experts agreed that the instructions and response options were clear and fitting. However, 1 expert thought that "having the 'how much do you like' at the top for all subsequent questions looks a bit confusing" [EXP005] and another expert thought that the use of "…" at the beginning of each item could possibly "imply hostility in writing" [EXP012].

The 17-item face adverse effects of treatment scale measured how participants' faces felt during the past week. Questions were answered on a 4-point Likert scale rating the severity of symptoms (not at all, a little bit, quite a bit, very much). Participants were asked to answer based on the past week. If one side of their face felt better than the other, they were asked to answer for the side that feels worse. While 9 participants thought the instructions were clear, 2 participants had difficulties interpreting the instructions. As well, 5 participants found the response options to be clear and easy to answer, with 2 participants preferring another middle response option. No concerns were raised by experts regarding the instructions and response options for the face adverse effects scale.

All 4 of the HRQOL scales (appearance distress, psychological, school and social) used the same time frame (past week) and response options on a 4-point Likert scale (never, sometimes, often, always). The appearance distress scale asked participants about how they felt about how they look. All 8 participants who provided feedback found the instructions to be clear while 1 participant noted that it might be difficult to answer thinking of the past week as they "[hadn't] been out much, because of exams" [COG008]. Of those who provided feedback on the response options of the HRQOL scales, 9 participants noted that the response options are good, with 2 wanting a neutral option between "sometimes" and "often". No concerns were raised by experts regarding the instructions and response options for the appearance distress scale.

The psychological scale asked participants about how they felt, thinking of the past week. For the psychological scale, all 6 participants who offered feedback on the instructions noted that they were clear. For those who commented on the response options, 6 participants thought they were good and 2 participants would have preferred another middle response option. No concerns were raised by experts regarding the instructions and response options for the psychological scale.

The school scale aimed to ask participants about how they felt about their school life during the past week. A special note is included to ask participants to think about when they were last in school if they were not in school during the past week. All 8 participants who offered feedback on the school scale instructions thought the instructions were clear, with 1 participant specifically noting that “I was not in school this week, so I will think about when I was last in school” [COG007]. While 8 participants thought the response options made sense and were good for the school scale, 1 participant would have preferred a more neutral response option. As well, 1 expert noted that while the response options were appropriate for most of the items in the scale, items 1, 3, 4, 5 and 8 might have been better answered using “not at all, a little bit, quite a bit and very much” [EXP024].

Lastly, the social scale asked participants to think about their social life over the past week. All 7 participants who offered feedback on the instructions thought they were clear and all 10 participants who offered feedback on the response options thought they were good. No concerns were raised by experts regarding the instructions and response options for the social scale.

3.3 Comprehensiveness

All participants were invited to comment on any items they felt were missing from each scale they completed. After completion of the questionnaire, participants were asked if they felt anything was missing from the questionnaire overall. Of the core scales, there was 1 code for the facial appearance scale, 1 code for the facial adverse effects scale, 1 code for the appearance distress scale, 1 code for the psychological scale, 4 codes for the school scale and 2 codes for the social scale. Suggested items are summarized in Table 9. Experts were asked to comment on any key items they felt were missing across all of the FACE-Q scales. While no additional items were identified by experts for the core scales, 4 experts identified swallowing/dysphagia as an important item that was missing from the mouth function scale.

3.4 Relevance

The recall periods of either thinking of “now” or “the past week” were deemed appropriate by all of the study participants and experts. Response options were considered appropriate, feedback regarding response options has been summarized under the Comprehensibility subsection above. Redundancy of the items was also assessed by asking participants if they thought any items were too similar or repetitive (Table 10). Experts were also invited to comment on any redundant items. For the face adverse effects scale, 1 expert suggested that items asking about tenderness and uncomfortableness “are the same” [EXP012], while another 3 experts noted that “tender and sore are hard for patients to differentiate” [EXP014].

Participants were invited to comment on the relevancy of the scales and items by stating whether they measured an important issue for them. Most patient participants found the facial appearance, negative psychosocial function, psychological, school, and social scales to be relevant to them and measured important issues (see Table 11). However, only 6 out of the 13 participants who answered the face feel scale, meant to measure adverse treatment effects, found the scale to be relevant to their current condition. For those who found the scale irrelevant, participants reflected that it “would have been more applicable earlier on” [COG013].

The overall impression of the FACE-Q from patient participant and expert feedback is positive. General comments from patient participants and experts are summarized in Table 12.

Section 4. Discussion

Content validity of the FACE-Q Kids for patients with HNC was evaluated through cognitive interviews with patients and feedback from pediatric oncology experts. The core scales were answered by all participants and demonstrate overall content validity from feedback offered by both patients and experts. Specifically, the vast majority of feedback showed that the items of the 6 core scales were comprehendible, comprehensive and relevant to participants. In addition to the patient input, experts offered feedback on all scales included in the FACE-Q.

For the scales measuring specific facial appearance and functional issues, although there was limited data due to smaller numbers of patients providing feedback, the results suggest that pediatric patients with HNC understood the concepts within the items, the wording of instructions (including recall periods), and felt positively about the response options and scales overall. Experts were invited to comment on the non-core scales and generally found them to be appropriate for individualized conditions resulting from HNC. For patients with HNC with affected eye function, 1 expert noted that "this is detailed enough to really assess for fine defects in eyelid, lacrimal and visual function" [EXP008].

The FACE-Q Craniofacial Module currently consists of 4 major domains with 26 independently functioning scales. Modifications to the FACE-Q for the oncology-specific module included removing the birthmark scale from the facial appearance domain and included the facial adverse treatment effects scale as a core scale. The FACE-Q oncology-specific module, as a result, currently consists of 4 major domains, with 13 scales in the facial appearance domain measuring how much patients like their face overall and specific parts of their face, 5 scales in the facial function domain, 5 scales in the HRQOL domain, and 2 scales in the adverse treatment effects domain. Of the 25 scales, 6 core scales were found to have content validity for pediatric HNC patients with facial differences. These include the face appearance scale (15 items), the face adverse effects scale (17 items), the appearance distress scale (10 items), the psychological scale (10 items), the school scale for children and youth aged 8-18 who attend schools with other children (10 items), and the social scale (11 items). All scales are measured on a 4-point Likert scale. These scales are currently being field-tested in a long-term international follow-up study involving pediatric patients with HNC from the UK, the Netherlands, France and the USA.

Pediatric head and neck cancer patients experience changes to their HRQOL at the time of diagnosis, throughout treatment, and post treatment. It is important to consider and measure both the invisible (e.g. psychological functioning, social interactions) and the visible (e.g. facial appearance/function) concerns of patients when seeking to understand the impact of cancer and its treatment on the HRQOL of patients. Given that existing measures of HRQOL fail to ask about facial appearance and function and therefore lack content validity, the FACE-Q Craniofacial Module could be valuable in helping to understand the HRQOL of pediatric patients with HNC through various domains and to understand their self-perception. This information would be valuable in assessing these children's needs for support and evaluating patient-important outcomes after diagnosis and treatment of HNC. Furthermore, the FACE-Q could offer a structured way for clinicians and social care professionals to discuss patient quality of life across various domains of life at follow up appointments.

4.1 COSMIN Criteria

FACE-Q Kids was assessed using the criteria outlined by COSMIN, the current gold standard approach for evaluating content validity of PROMs. The 3 sections evaluated include comprehensibility, comprehensiveness and relevance.

4.1.1 Comprehensibility

Instructions (including recall period), items, and response options for the core scales were reviewed by all patient participants and experts and were generally understood as intended. A few items were found to be confusing or difficult to answer and were flagged for review by the research team (summarized in Table 8). These items may be modified, reworded, or removed during future revisions. Difficulty with questions 10 and 11 in the face appearance scale may be due to the questions being included in the questionnaire for adult patients with facial paralysis. General comments on comprehension and interpretation of items can be found in Table 7, indicating that the items are appropriately worded and understood by pediatric HNC patients.

4.1.2 Comprehensiveness

Both patient participants and experts were asked whether they thought any concepts or items were missing from the scales they assessed. The only key concept deemed missing after analysis was completed is a swallowing/dysphagia section in the face adverse effects scale.

4.1.3 Relevance

Relevance of items and scales within hypothetical constructs was demonstrated through expert feedback and by asking patient participants whether they thought any of the items did not belong with the rest. Each of the 4 domains included in the FACE-Q Kids were assessed: facial appearance, facial function, quality of life, and adverse treatment effects. Relevance of items to the target population of interest was also assessed through expert feedback and cognitive interviews, the data are summarized in Tables 11 and 12. Response options were also assessed by both patient participants and experts, with the majority of responses indicating that they "make sense" and are "good". The recall period of "one week" was assessed by both participants and experts and deemed to be appropriate. Redundancy was assessed by directly asking participants. It was found that 9 participants thought the questions asking "… how balanced your face looks?” and “… how symmetric your face looks?” were too similar in the face appearance scale. Redundant items will be addressed through future revisions to the scales.

4.2 Limitations

A limitation of this study is that only participants with 5 distinct HNC diagnoses were included due to the small patient pool available. Other common HNC diagnoses in the pediatric population include oral cancer and salivary gland cancer. This prevented the gathering of an adequate amount of participant feedback for all 25 scales included in the FACE-Q oncology-specific module, since participants were only asked to complete facial appearance and function scales relevant to their specific conditions. Although all participants were asked to complete the 6 core scales, not all 6 were completed by a few participants. As well, although a consecutive sampling technique was intended, this was difficult to ensure due to having multiple recruitment sites and having different members of the healthcare teams as contacts. Further research could be conducted that includes greater diversity of patient diagnoses to assess the content validity of the scales not included in the core set. It should be noted that the FACE-Q scales aims to assess outcomes from the patient perspective and does not assess or capture objective measures of facial appearance/function differences that are also important outcomes. Another limitation of the current study is the inclusion of only 2 patients in the 8-12 age range. The inclusion of participants with a diverse range in age from 13-30 and educational backgrounds in the HNC sample and the original sample of craniofacial participants has ensured that the FACE-Q is understandable to children with varying literacy levels and uses child-friendly language. Finally, the current study lacks feedback from professionals of certain relevant disciplines, including speech language pathologists, physical and occupational therapists, and reconstructive surgeons. While dentists and orthodontists were not included for expert feedback, the dental scales included in the FACE-Q evidenced content validity in a separate study of patients seeking orthodontic care (Tsangaris et al., 2017).

4.3 Next Steps

Next steps in research to provide a means to measure outcome in pediatric HNC is to test the scales and examine their psychometric performance. Data obtained from psychometric testing will be used along with the qualitative data obtained in the current study to further revise the instrument. To this end, the FACE-Q scales were provided to a research team conducting an international follow-up study of pediatric HNC survivors. Data from this sample has been added to the field-test sample of craniofacial patients, some of whom also have malignant and benign HN tumours, into the FACE-Q field-test study. The full sample of over 2000 participants includes 188 participants with HNC. Psychometric analysis will be conducted to determine the performance of the scales overall as well as in the HNC patients. A range of psychometric tests will be performed to examine reliability and validity of the scales. In addition, the association between FACE-Q scale scores and the severity of patient facial difference (appearance and function) will be examined to evaluate the discriminative validity of the questionnaire and its future potential to measure change.

Section 5. Conclusion

The FACE-Q evidenced content validity for core scales as well as limited evidence that the remaining scales covered issues relevant to patients with specific facial differences. The FACE-Q is a comprehensive, easy to understand and relevant outcome measure of subjective HRQOL appearance and facial function. This comprehensive modular instrument shows promise for use as both as an outcomes measure and tool for measuring patient HRQOL in clinical practice. Assessment of the psychometric properties of the new measure is forthcoming as part of the international FACE-Q field-test study.

References

Albornoz, C. R., Pusic, A. L., Reavey, P., Scott, A. M., Klassen, A. F., Cano, S. J., … Matros, E. (2013). Measuring health-related quality of life outcomes in head and neck reconstruction. *Clinics in Plastic Surgery*, *40*(2), 341–349.

Annertz, K., Anderson, H., Biörklund, A., Möller, T., Kantola, S., Mork, J., … Wennerberg, J. (2002). Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. *International Journal of Cancer*, *101*(1), 95–99.

Anthony, S. J., Selkirk, E., Sung, L., Klaassen, R. J., Dix, D., Scheinemann, K., & Klassen, A. F. (2014). Considering quality of life for children with cancer: a systematic review of patient-reported outcome measures and the development of a conceptual model. *Quality of Life Research*, *23*(3), 771–789.

Björklund, M., Sarvimäki, A., & Berg, A. (2010). Living with head and neck cancer: A profile of captivity. *Journal of Nursing and Healthcare of Chronic Illness*, *2*(1), 22–31.

Brockstein, B., & Masters, G. (2010). *Head and Neck Cancer*. Springer US. Retrieved from https://books.google.ca/books?id=HJoHkgAACAAJ

Cesmebasi, A., Gabriel, A., Niku, D., Bukala, K., Donnelly, J., Fields, P. J., … Loukas, M. (2014). Pediatric head and neck tumors: an intra-demographic analysis using the SEER* database. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, *20*, 2536–2542. https://doi.org/10.12659/MSM.891052

Curado, M. P., & Hashibe, M. (2009). Recent changes in the epidemiology of head and neck cancer. *Current Opinion in Oncology*, *21*(3), 194–200.

Eades, M., Chasen, M., & Bhargava, R. (2009). Rehabilitation: long-term physical and functional changes following treatment. In *Seminars in oncology nursing* (Vol. 25, pp. 222–230). Elsevier.

Herzon, F. S., & Boshier, M. (1979). Head and neck cancer—Emotional management. *Head & Neck Surgery*, *2*(2), 112–118. https://doi.org/10.1002/hed.2890020206

Howren, M. B., Christensen, A. J., Hynds Karnell, L., Van Liew, J. R., & Funk, G. F. (2013). Influence of pretreatment social support on health-related quality of life in head and neck cancer survivors: Results from a prospective study. *Head & Neck*, *35*(6), 779–787.

Hunt, O., Burden, D., Hepper, P., & Johnston, C. (2005). The psychosocial effects of cleft lip and palate: a systematic review. *European Journal of Orthodontics*, *27*(3), 274–285.

Hunt, O., Burden, D., Hepper, P., Stevenson, M., & Johnston, C. (2007). Parent reports of the psychosocial functioning of children with cleft lip and/or palate. *The Cleft Palate-Craniofacial Journal*, *44*(3), 304–311.

Jones, E., Lund, V. J., Howard, D. J., Greenberg, M. P., & McCarthy, M. (2007). Quality of life of patients treated surgically for head and neck cancer. *The Journal of Laryngology & Otology*, *106*(03), 238–242. https://doi.org/10.1017/S0022215100119152

Koster, M., & Bergsma, J. (1990). Problems and coping behaviour of facial cancer patients. *Social Science & Medicine*, *30*(5), 569–578.

Krouse, J. H., Krouse, H. J., & Fabian, R. L. (1989). Adaptation to Surgery for Head and Neck Cancer. *The Laryngoscope*, *99*(8), 789???794. https://doi.org/10.1288/00005537-198908000-00004

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390.

Lawrence, J. W., Rosenberg, L., Mason, S., & Fauerbach, J. A. (2011). Comparing parent and child perceptions of stigmatizing behavior experienced by children with burn scars. *Body Image*, *8*(1), 70–73.

List, M. A., & Bilir, S. P. (2004). Functional outcomes in head and neck cancer. In *Seminars in radiation oncology* (Vol. 14, pp. 178–189). Elsevier.

Longmire, N. M., Riff, K. W. Y. W., O'Hara, J. L., Aggarwala, S., Allen, G. C., Bulstrode, N. W., …

Marucci, D. (2017). Development of a new module of the FACE-Q for children and young adults with diverse conditions associated with visible and/or functional facial differences. *Facial Plastic Surgery*, *33*(05), 499–508.

Marur, S., & Forastiere, A. A. (2008). Head and neck cancer: changing epidemiology, diagnosis, and treatment. In *Mayo Clinic Proceedings* (Vol. 83, pp. 489–501). Elsevier.

Masnari, O., Schiestl, C., Rössler, J., Gütlein, S. K., Neuhaus, K., Weibel, L., … Landolt, M. A. (2013). Stigmatization predicts psychological adjustment and quality of life in children and adolescents with a facial difference. *Journal of Pediatric Psychology*, *38*(2), 162–172. https://doi.org/10.1093/jpepsy/jss106

Modh, A., Gayar, O. H., Elshaikh, M. A., Paulino, A. C., & Siddiqui, F. (2018). Pediatric head and neck squamous cell carcinoma: Patient demographics, treatment trends and outcomes. *International Journal of Pediatric Otorhinolaryngology*, *106*, 21–25.

Ojo, B., Genden, E. M., Teng, M. S., Milbury, K., Misiukiewicz, K. J., & Badr, H. (2012). A systematic review of head and neck cancer quality of life assessment instruments. *Oral Oncology*, *48*(10), 923–937.

Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics, 2002. *CA: A Cancer Journal for Clinicians*, *55*(2), 74–108.

Sengupta, S., Pal, R., Saha, S., Bera, S. P., Pal, I., & Tuli, I. P. (2009). Spectrum of head and neck cancer in children. *Journal of Indian Association of Pediatric Surgeons*, *14*(4), 200–203. https://doi.org/10.4103/0971-9261.59601

Strauss, J. S., Krowchuk, D. P., Leyden, J. J., Lucky, A. W., Shalita, A. R., Siegfried, E. C., … Sieck, C. K. (2007). Guidelines of care for acne vulgaris management. *Journal of the American Academy of Dermatology*, *56*(4), 651–663.

Stubbs, T. K., James, L. E., Daugherty, M. B., Epperson, K., Barajaz, K. A., Blakeney, P., … Kagan, R. J. (2011). Psychosocial impact of childhood face burns: a multicenter, prospective, longitudinal study of 390 children and adolescents. *Burns*, *37*(3), 387–394.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., … Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*, *27*(5), 1159–1170.

Topolski, T. D., Edwards, T. C., & Patrick, D. L. (2005). Quality of life: how do adolescents with facial differences compare with other adolescents? *The Cleft Palate-Craniofacial Journal*, *42*(1), 25–32.

Tsangaris, E., Wong Riff, K. W. Y., Goodacre, T., Forrest, C. R., Dreise, M., Sykes, J., … Klassen, A. F. (2017). Establishing Content Validity of the CLEFT-Q: A New Patient-reported Outcome Instrument for Cleft Lip/Palate. *Plastic and Reconstructive Surgery. Global Open*, *5*(4), e1305–e1305. https://doi.org/10.1097/GOX.0000000000001305

Vickery, L. E., Latchford, G., Hewison, J., Bellew, M., & Feber, T. (2003). The impact of head and neck cancer and facial disfigurement on the quality of life of patients and their partners. *Head & Neck*, *25*(4), 289–296. https://doi.org/10.1002/hed.10206

Wickert, N. M., Wong Riff, K. W. Y., Mansour, M., Forrest, C. R., Goodacre, T. E. E., Pusic, A. L., & Klassen, A. F. (2018). Content validity of patient-reported outcome instruments used with pediatric patients with facial differences: a systematic review. *The Cleft Palate-Craniofacial Journal*, *55*(7), 989–998.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.

Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press.

Figure 1: Conceptual framework of FACE-Q Craniofacial Module

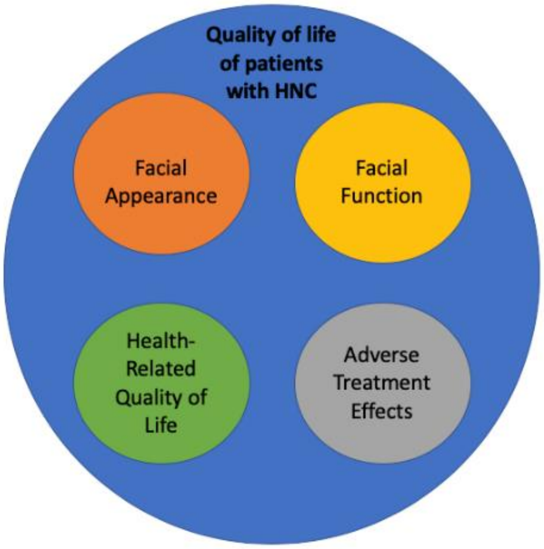

Table 1: Cognitive interview guide

Part 1: Instructions

1. What are the instructions asking you to do? Please explain to me in your own words.
2. What does time frame mean to you? What days do you include? (e.g., past week)
3. Are there any words we should change to make the instructions easier?

Part 2: Items

1. In your own words, what do you think this item is asking?
2. What do you think of when answering this item?
3. Are any words difficult to understand?
4. Was this item hard or easy to answer? Why?
5. Does this item measure an important issue for you?
6. What do you think about the response choices?

Part 3: At the end of a scale

1. In your own words, what is this group of items asking about?
2. Does this group of items measure an important issue for you?
3. Are there any items that don't "belong" with the rest, missing, redundant?

Part 4: At the end of the interview

1. What are your overall thoughts about our questionnaire?
2. Is there anything we forgot to ask that is important to patients with facial differences?
3. Is there anything we should change about our questionnaire?
4. Is there anything else that you would like to add or comment upon?

Table 2: Cognitive Interview Approach

Cognitive Interviewing	**Steps Outlined by Willis**	**Modified Steps for Present Study**
Purpose	To gather information about the functioning of the survey questions before to finalizing for field testing	None
Sample Size	Typically small and may consist of several rounds (approximately 10 interviews per round)	1 round of interviews with 15 participants
Recruitment	Participants obtained through a variety of recruitment strategies to produce variations in the types of individuals recruited	Participants were recruited during their clinic visits, were contact through email, or were contacted by telephone
Interviewers	Generally a small number (1–4) of highly trained cognitive interviewers	Most interviews were conducted by Y.W., some interviews were conducted by 2 other trained interviewers
Materials presented	A cognitive interview guide consisting of the survey questionnaire along with probe questions	None
Method	Flexible administration that relies heavily on probe questions to enhance the quality of the questionnaire	None
Analysis	Qualitative analysis of responses based on interviewer notes or recording of verbatim interviews	Qualitative analysis conducted based recording of verbatim interviews

This table was adapted from Willis GB. Analysis of the cognitive interview in questionnaire design: understanding qualitative research. Toronto: Oxford University Press; 2015 (Willis, 2015).

Table 3: All FACE-Q Kids scales, core scales are highlighted

Facial Appearance Scales	Facial Function Scales	Health-Related Quality of Life Scales	Adverse Treatment Effects
Face	Face	Appearance distress	Face
Cheeks	Eyes	School	Ears
Chin	Mouth	Social	
Ears	Breathing	Psychological	
Eyes/Eyelids	Speaking	Speech distress	
Forehead/Eyebrows			
Head			
Jaws			
Lips			
Teeth			
Nose			
Nostrils			
Smile			

Table 4: Participant demographics

Sample N=15	
Age in years	
8-12	2
13-17	10
18-30	3
Gender	
Male	6
Female	9
Diagnosis	
Rhabdomyosarcoma	8
Hodgkin Lymphoma	1
Brain tumour	3
Thyroid tumour	1
Soft tissue sarcoma	1
Unspecified	1
Time since treatment completion	
0-3 years ago	0
4-5 years ago	1
More than 5 years ago	10
Unsure	4

Table 5: Expert demographics

Sample N=21	
Country	
Canada	17
USA	3
The Netherlands	2
Occupation	
Oncologist	9
Nurse practitioner	3
Social worker	1
Otolaryngologist	2
PhD Trainee	1
Researcher	1
Psychologist	3
Training HNC surgeon	1
Other (Unspecified)	1
Main focus	
Pediatric cancer	10
Pediatric HNC	7
Not applicable	4

Table 6: Number of participants who provided feedback for each scale

Scales	**Rhabdomyosarcoma**	**Hodgkin Lymphoma**	**Brain tumour**	**Thyroid tumour**	**Soft tissue sarcoma**	**Unspecified**	**Total**
Facial Appearance							
Cheeks	3				1		4
Chin	2						2
Ears	1	1					2
Eyes eyelids	1		3	1			5
Forehead eyebrows			3	1		1	5
Jaws	3	1					4
Lips	2			1		1	4
Smile	4			1		1	6
Teeth	1	1	1				3
Face	8	1	3	1	1	1	15
Head	1		2				3
Nose	1		1		1		3
Nostrils	1						1
Facial Function							
Face	3	1				1	5
Eyes			3	1	1	1	6
Breathing	1		2				3
Mouth	2					1	3
Speaking					1		1
Health-Related Quality of Life							
Negative psychosocial function	8	1	2	1	1		13
Psychological	7	1	3	1	1		13
School	6		2	1	1		10
Social	6	1	3	1	1		12
Speech distress							0
Adverse Treatment Effects							
Face feel	8	1	2	1	1		13
Ears feel							0

Table 7: Items in core scales and example quotes showing comprehension interpretations

Face appearance scale Instructions: HOW DOES YOUR FACE LOOK? Answer each question by circling one number. Please answer thinking of how your face looks NOW.	
Items	Example quotes for interpretation
1. …how your face looks when you look your best?	(9) When I look in a mirror and I had a shower and a good sleep and I look more awake. (10) When I dress up and put makeup on. (11) After a shower and I actually got ready with like my best clothes.
3. …the shape of your face (eg, round or oval)?	(2) If your face have any points or square. (4) Thinking of if it's round or oval. (9) One side looks a bit more in than the other, so it kind of looks like a C.
5. …how your face looks when you smile?	(1) When I smile I can’t move the left side of my smile. So it’s just, it looks weird compared to other people. (5) Because of the radiation and chemo, I lost all my teeth in the bottom jaw. So when I smile, my bottom jaw tends to crunch up my lip and I kind of look like a toothless old man in a sense. (6) I don't like how my face looks when I smile because I need a lot of orthodontal plastic surgery. I don't know how much would have been there without the treatment, like if I would have a slightly underdeveloped jaw regardless, but it's worse because of the treatment.
7. …how your face looks up close?	(3) I’m thinking is up close like having a face-to-face conversation with someone? (5) So it would be up close and personal, so I would be looking deeply into the mirror. So for someone looking deeply. And sort of at that point, I can kind of pick out all those little imperfections that I may have. (12) When someone is taking a close picture of you, I just, I personally don’t like my face being too close. Notice eye being droopy from eye right away.
9. …how your face looks when it’s relaxed (still)?	(5) My opinion of how your face looks when you are at a calm state of mind. (9) When I’m relaxed, one side is more droopy than the other side. (13) When it’s just relaxed, I usually look kind of mad. (16) Neutral face, pretty normal.
11. …how balanced your face looks?	(7) There is a bit of imbalance between my left and right side. I’m not entirely sure why. It’s just my left eye seems a little bit baggier. (10) Everything is even and with each other.

	(11) One side looks bigger than the other, I do have a pretty big forehead. My nose is big too. (13) Balanced doesn't mean symmetric, it means that it fits. I think evenness. Like the eyes are the same shape and the lip…the bottom and the top lip are around the same.
13. …how well both sides of your face match (look the same)?	(2) Half my face is extremely bumpy and half of it isn't. (3) Generally my face matches on both sides except for the eye. (8) I have marks on the left side of my face that I don't have on the right side of my face.
15. …how your face looks overall?	(7) I'm pretty happy with the shape of my face. (13) It's unbalanced and unsymmetrical (eyes and lips don't tend to match with each other well).
Face adverse effects scale Instructions: HOW DOES YOUR FACE FEEL? Answer each question by circling one number. Please answer thinking of the PAST WEEK. NOTE: Does one side of your face feel better than the other? If yes, please answer thinking of the side that feels worse.	
2. My face feels tender.	(6) Tender I guess would be like if my face felt super sensitive all the time. (7) I kind of think of something like that's a little bit more vulnerable. So something like a just healed wound. (10) Tender is just like when you kind of slightly touch it and it just kind of hurts.
4. My face feels uncomfortable.	(8) Uncomfortable as in if I were to touch it and then compare it to the other side, you can feel my bone way more easily than the other side. So it feels uncomfortable. So if I moisturize my face, it feels uncomfortable to rub in the cream on the left side of my face. (9) Face feels tender, feels a bit uncomfortable on one side. (14) Left cheek feels weird when I'm breathing out, can feel left cheek pushing up and don't feel in right cheek.
6. My face feels firm when I touch it.	(5) When you make certain facial expressions and your face can tense up and kind of flex in a sense from parts of your eyebrows and whatnot. (6) I guess if you had some sort of treatment that made your face sort of tighter in one area, then that's what that would mean to me. I have a lot of scar tissue back here, because that's where my radiation was. So... Maybe it feels kind of firm on the inside of my mouth, but my face itself, that doesn't apply. (8) Like hard I guess. Like solid. And it's a little bit solid, I guess. Because I can feel my bone.
8. My face feels tight.	(5) I would refer to this one as my face after it's sunburned, all the skin is really tight, as if my skin is dehydrated and completely stretched and pulled.

	(16) Tight would be if I maybe moved my mouth over to one side of my face. Tightness would be more resistance.
10. My face feels tingly (pins and needles feeling).	(4) Like a bzzzz feeling.
12. There is blood or other fluid coming from my face.	(5) I tend to have very dry nose a lot of the time and when I blow my nose, I tend to have…it erupted something in my nasal cavity and then I end up having blood when I blow my nose. (7) I think of open sores or a cut that won't heal, or something like that. (15) Eye leaks without crying.
14. It is hard for me to eat or drink.	(5) Because of the fact that I have only 3 bottom teeth which are molars, I do have trouble eating. I do not have trouble drinking though. (9) Not because of my face (chewing), it's more from the back of my throat when I lost some nerve damage or something like that (loss of sensation in throat when swallowing).
16. My eyes bother me (eg, itchy, dry, red, watery).	(5) My eyes tend to be very itchy. I tend to have a buildup of sort of what I call gunk in my eyes, I tend to get that all throughout the day to the point where I can actually get blurred vision from it. (9) my eyes do get really itchy, but it's because I have allergies and they are not really dry
Appearance distress scale Instructions: HOW DO YOU FEEL ABOUT HOW YOU LOOK? Answer each question by circling one number. Please answer thinking of the PAST WEEK.	
1. I get upset when people stare at me.	(3) I've pretty much gotten used to all the staring so I don't really care about it anymore. (8) I don't like the feeling of people staring at me, because I assume they might be looking at an impurity or maybe talking behind my back.
3. I avoid having my photo taken.	(3) Usually when I have photos taken, I like to know about them. People will surprise take them, and I don't like that so I avoid it. (6) I hate pictures. I've always not liked pictures of myself. Cus I'm not photogenic, at all. But, part of it is, I notice the asymmetry more when I see a picture of me.
5. I avoid looking at myself in a mirror.	(8) If I'm at my house, then I'm by myself. it's not awkward for me to look in the mirror. But if I were to go to the girls bathroom and fix my hair, and I was looking at myself and I notice that there is someone else also looking

	in the mirror, it’s awkward for me to look in the mirror. (13) Sometimes I just feel frustrated and I don’t want to look.
7. I avoid sharing photos of myself (eg, Instagram or Snapchat).	(7) It gives the example of Instagram and Snapchat, but I more feel like sharing pictures of myself at places I visited with my friends (on my phone). (8) I think sometimes I avoid sharing. I would post pictures online. I guess sometimes if you compare it to someone who is always posting and then there is me and I post like once in a year or a couple of times in a year. (11) Maybe embarrassing pictures, I just don't send or post them.
9. I feel self-conscious about how I look.	(4) It means like you worry about what other people think about you. (10) My eyes are different, have gotten used to feeling self-conscious. (12) Sometimes I feel like I don't look the best, so I'm scared of what other people think about me.
Psychological scale Instructions: HOW DO YOU FEEL? Answer each question by circling one number. Please answer thinking of the PAST WEEK.	
2. I enjoy life.	(7) I enjoy life is kind of like your day-to-day doings rather than the space you are at. (11) Playing outside, enjoy life. Be thankful that I had cancer and I’m alive right now. (11) make the best of everything that you have.
4. I feel okay about myself.	(7) I feel OK about myself is I’m comfortable with where I am right now. I feel I can make improvements to myself, but I'm pretty OK with where I am.
6. I am proud of myself.	(8) Often I’m proud of myself, because I usually do something new. Like I got into this new art, I was happy about myself because I was able to create it. (15) Feel like doing something and feel good about it.
8. I feel confident.	(6) My insecurity about my asymmetry and also it has to do with going through treatment. I was in treatment for a year. So I was taken out of school, so I was away from social life, I was away from my friends for a while. So I think that made a difference in my confidence. (8) When I’m on my own I feel confident, but when I’m with other people, maybe I feel a little less confident because I might compare myself to them. (12) Self-esteem.
10. I feel good about how I look.	(7) That’s mostly related with my whole body rather than my face.

	(8) There are times I might wake up and my hair is in a mess so then I feel rumpled.
School scale Instructions: HOW IS YOUR SCHOOL LIFE? Answer each question by circling one number. Please answer thinking of the PAST WEEK. NOTE: if you were not in school this week, think about when you were last in school.	
1. I like seeing my friends at school.	(8) Always, I don’t like to be alone at school. (13) Yeah always. They are good people. (15) Don't have a lot of friends, don't really have friends.
3. I feel accepted at school.	(7) I feel accepted at school is a very general thing, where you feel you belong there or at least you are part of the student population. (8) I don’t have enemies at school and there is no one I’m against, so I always feel accepted. (9) Friends or teachers like me. And I belong there.
5. I am happy at school.	(7) I'm not entirely happy to actually be in school, but I enjoy being with my friends and stuff. (8) Usually when there are no tests or quizzes or anything to stress me out. (13) School is my favorite place to be.
7. Other students listen to me when I talk.	(8) Often, excluding in my group work when we are forced to work with people who don’t want to work, so they might ignore you. (10) I have to say it a few times to actually get someone’s attention to listen to what I’m trying to say.
9. It’s easy for me to make friends.	(6) I've always been a little bit shy and I've definitely gotten more shy after my treatment. (7) I have lots of friends at school. (8) I’m not a person who likes talking. I find sometimes I’m isolated. I isolate myself, so it’s harder for me to make friends in a new class. (10) When I try to make friends, sometimes [they don’t want] to be friends because of the way I look.
Social scale Instructions: HOW IS YOUR SOCIAL LIFE? Answer each question by circling one number. Please answer thinking of the PAST WEEK.	
1. My friends accept me.	(8) I think they always accept me. They don’t tell me that there is anything wrong with me. (10) I have friends that actually care and don’t care about the way I look. (13) I think they… get me. They understand my humour at least.

3. People listen to what I have to say.	(5) Sometimes people underestimate me or people don't want to take your opinion. But as far as a regular everyday conversation, people do listen to me always. (6) I tend to be shy, so sometimes maybe I'm not assertive enough, and I think that had to do with my treatment. (14) People understand what I have to say. So in general people are usually interested in what I have to say.
5. I like being with other people.	(5) I tend to be alone a lot so I like company. (7) I mean if they are my friends, then yeah. (8) That depends, because if it's with friends, then always or often. But if it's with strangers, then sometimes. When reading question, think more about strangers/people I don't know very well.
7. I feel like I fit in.	(9) Like fit in with the majority of people. Not just friends, but the majority of people. (14) Well it's just how people treat me. If I don't fit in, I'm just sitting there and not doing anything.
9. I feel the same as other people my age.	(5) I feel that I went wrong somewhere. And I'm catching up. (6) I think that's a given for anyone who goes through the kind of thing that I went through when I was young. I definitely don't feel the same as other people my age. (13) With the things I've gone through, not everyone can understand. (16) Older sisters always told me that I seem a lot older than I am. And I think ever since I was sick when I was 10, I've kind of been like that. I just skipped the whole juvenile attitude.
11. It's okay if people ask me about my face.	(8) Maybe it's like a friend I'm OK with. But if it's a stranger, then not really. (12) Depending on the person, sometimes people say it to try to make fun of you and I'm not going to answer you, because you don't bother me. (14) Like it depends on how they ask me. I feel if they are rude and ask me, then I don't tell people why my face looks different.
Participant IDs are indicated by the numbers in brackets.	

Table 8: Items with comments suggesting problems with comprehension that were flagged for review in the core scales

Facial appearance scale Instructions: HOW DOES YOUR FACE LOOK? Answer each question by circling one number. Please answer thinking of how your face looks NOW.	
10. …how rested your face looks (not tired)?	(1) I'm confused about rested. *– 13 yo, F.* (2) I do not understand that one. *– 13 yo, M.* (15) Confused from reading question. *– 22 yo, F.*
11. …how balanced your face looks?	(2) That one does not make any sense. *– 13 yo, M.* (5) Are you ears considered part of your face? I suppose so, right? *– 18 yo, F.*
14. …how your face looks from the side (your profile)?	(1) It depends which side. *– 14 yo, F.* (4) What does that mean? Which side? *– 30 yo, M.* (5) Depends on which side. *– 18 yo, F.* (15) Not sure whether to answer using preferred side or not preferred side, thought of preferred side reading [the] question. *– 22 yo, F.*
Face feel scale Instructions: HOW DOES YOUR FACE FEEL? Answer each question by circling one number. Please answer thinking of the PAST WEEK. NOTE: Does one side of your face feel better than the other? If yes, please answer thinking of the side that feels worse.	
2. My face feels tender.	(8) Do you mean, like, soft? *– 14 yo, M.* (15) What does this mean? *– 22 yo, F.*
6. My face feels firm when I touch it.	(4) Does that mean it's stable? *– 30 yr, M.* (5) I don't know what they would consider a perfectly normal response to that out of the responses they give me to choose from. *– 18 yo, F.* (6) I'm not entirely sure what is meant by that, I guess. It's not relevant to me so I don't know what it means as much. *– 16 yo, F.* (15) What does this mean? *– 22 yo, F.*
Negative psychosocial function scale	

Instructions: HOW DO YOU FEEL ABOUT HOW YOU LOOK? Answer each question by circling one number. Please answer thinking of the PAST WEEK.	
7. I avoid sharing photos of myself (eg, Instagram or Snapchat).	(8) I was thinking more of posting. I would avoid posting pictures of myself or exposing myself to the internet, I guess. *– 14 yo, M.*
School scale Instructions: HOW IS YOUR SCHOOL LIFE? Answer each question by circling one number. Please answer thinking of the PAST WEEK. NOTE: if you were not in school this week, think about when you were last in school.	
5. I am happy at school.	(6) Is anyone really that happy at school? *– 16 yo, F.*
Social scale Instructions: HOW IS YOUR SOCIAL LIFE? Answer each question by circling one number. Please answer thinking of the PAST WEEK.	
4. People treat me the same as everyone else.	(8) I think racial wise, maybe. Lots of people are racist, so they might act differently towards one person versus another person. So then I think they always treat me the same, the way they would treat someone else the same. *– 14 yo, M.*
6. I feel confident when I go out (like to a party).	(8) When I go to a party and I'm with friends, I feel confident when I go out, because I'm thinking of when I'm with my friends. *– 14 yo, M.* (10) It's more than just my friend group and there will be people there that don't really accept it. *– 10 yo, M.*
10. It's okay when people look at my face.	(14) Looking? Like how long do they look? *– 14 yo, F.* (14) Maybe if you change the question to stare then it would be easier to answer that question. *– 14 yo, F.*
11. It's okay if people ask me about my face.	(8) Maybe it's like a friend I'm OK with. But if it's a stranger, then not really. *– 14 yo, M.* (10) Depending on how your ask or the way you ask. If someone is going to be rude about it or mean about it, I'm not going to answer. But if they are just kind of curious or just want to know, but they are not being rude, I'll tell them. *– 10 yo, M.* (14) Like it depends on how they ask me. I feel if they are rude and ask me, then I don't tell people why my face looks different. *– 14 yo, F.*

	(15) Depends on what you are asking me about, asking nicely. – *22 yo, F.*
Participant IDs are indicated by the numbers in brackets.	

Table 9: Additional items suggested by patient participants

Scale	**Suggested Item**
Face appearance	(8) How do you feel about where you got treatment?
Face adverse effects	(8) One side of my face feels different [from] the other side.
Appearance distress	(5) [I] don't worry about what other people think or what I may look like.
Psychological	(6) Does how you look stop you from doing things in life?
School	(4) How happy are you with how many friends you have?
	(6) Something more specific regarding "if kids feel like they're not joining things or they're not doing things at school because of insecurities"
	(12) Do you have to deal with a learning disability or are you a slow learner [compared with] students in your class?
	(15) How I get bullied as an add-on question.
Social	(5) Duplicate these questions onto a professional scale for people who work.
	(15) How do you feel about how kids treat you?
Participant IDs are indicated by the numbers in brackets.	

Table 10: Items flagged as redundant by participants

Scale	**Items**	**Number of participants who found items to be redundant**
Face appearance	1. …how your face looks when you look your best? 2. …how your face looks when you are ready to go out (like to a party)?	3
	5. …how your face looks when you smile? 6. …how your face looks when you laugh?	1
	9. …how your face looks when it's relaxed (still)? 10. …how rested your face looks (not tired)?	2
	11. …how balanced your face looks? 12. …how symmetric your face looks? 13. …how well both sides of your face match (look the same)?	7
	11. …how balanced your face looks? 12. …how symmetric your face looks?	2
	12. …how symmetric your face looks? 13. …how well both sides of your face match (look the same)?	3
Face adverse effects	2. My face feels tender. 5. My face feels sensitive when I touch it.	1
Appearance distress	2. My face feels tender. 4. My face feels uncomfortable.	3
	6. My face feels firm when I touch it. 8. My face feels tight.	2
Psychological	1. I am happy with my life. 2. I enjoy life. 3. I feel happy.	2
	1. I am happy with my life.	1

	3. I feel happy.	
	4. I feel okay about myself. 9. I feel great about myself.	1
	4. I feel okay about myself. 10. I feel good about how I look.	1
	5. I believe in myself. 8. I feel confident.	2
	7. I like myself. 9. I feel great about myself.	1
	9. I feel great about myself. 10. I feel good about how I look.	1

Table 11: FACE-Q Kids Oncology Module Scales and Relevance

Scale	**Relevant to participants**	**Example Quote for Relevance**
Facial Appearance	11/15	(7) Measures something important, this is a good indicator of how I feel generally about the shape and look of my facial features and my head
Face feel	6/13	(12) Scale is trying to ask about how your treatments have affected your face and how your face feels and how you feel about your face.
Negative psychosocial function	10/13	(9) Scale is trying to ask if I'm self-conscious about my face. How I feel. How do I feel about how other people feel about my face.
Psychological	10/13	(7) Scale measures self-worth and your kind of mental quality of life. And how you feel every day. So I think those are all very important things to measure.
School	8/10	(7) Scale measures something important for them, It shows if I am really enjoying school life, because school from elementary to high school is a very large part of my life. It's 8 hours of the day for most of the year. If I'm not really enjoying school, then the rest of my life and my quality of life would end up suffering greatly.
Social	9/12	(13) Scale is trying to ask about how people feel around others and their social life.
Participant IDs are indicated by the numbers in brackets.		

Table 12: General comments regarding the FACE-Q Kids Oncology Module

Patient participants
(16) I think it… just made me think a lot more about all the different struggles that people experience from the same thing.
(15) Liked the questionnaire a lot because I got to express myself, I got to express how I feel about myself and someone was actually listening.
(11) I would say…I feel a little bit better about how I look. Because like…kind of like shows stuff that I'm feeling.
(8) I think I enjoyed it and I think it was a good questionnaire. I think you're researching right. So I think it's a good method I guess, to know how children are feeling. How they are feeling about the way they are after the cancer.
(7) I thought it was pretty good at measuring a lot of factors that would affect my quality of life.
(6) I guess I would say, if anything, it made me feel like maybe I shouldn't care as much about how I look. Because... it's not that important.
Expert Input
(4) These would work for body dysmorphia [as well], drawings/illustrations are very helpful
(8) This is detailed enough to really assess for fine defects in eyelid, lacrimal, and visual function.
(8) This is a great questionnaire set
(17) Great questions!
(19) Looks good
Participant IDs are indicated by the numbers in brackets.