HHS Public Access

Author manuscript

Clin Trials. Author manuscript; available in PMC 2024 March 08.

Published in final edited form as:

Clin Trials. 2024 February ; 21(1): 6–17. doi:10.1177/17407745231205577.

Recommendations for Assessing Appearance Concerns Related to Plexiform and Cutaneous Neurofibromas in Neurofibromatosis 1 Clinical Trials

Vanessa L Merker[1], **Heather L Thompson**[2], **Pamela L Wolters**[3], **Frank D Buono**[4], **Cynthia M Hingtgen**[5], **Tena Rosser**[6], **Belinda Barton**[7], **Carolina Barnett**[8], **Taylor Smith**[9], **Diana Haberkamp**[10], **Miranda L McManus**[11], **Andrea Baldwin**[12], **Irene P Moss**[13], **Claas Röhl**[14], **Staci Martin**[3] **on behalf of the REiNS International Collaboration**

[1]Department of Neurology and Cancer Center, Massachusetts General Hospital, Boston, MA, USA

[2]Department of Communication Sciences and Disorders, California State University, Sacramento; Sacramento, CA, USA

[3]Pediatric Oncology Branch, National Cancer Institute, Bethesda, MD, USA

[4]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

[5]Neurological Health, Corewell Health, Grand Rapids, MI, USA

[6]Division of Neurology, Children's Hospital Los Angeles, Los Angeles, California, USA

[7]Faculty of Health, Discipline of Psychology, Southern Cross University, Coffs Harbour, NSW, Australia

[8]Division of Neurology, Department of Medicine, University Health Network and University of Toronto, Toronto, Ontario, Canada

[9]Department of Psychology and Child Development, California Polytechnic State University, San Luis Obispo, CA, USA

[10]Neurofibromatosis Midwest, Chicago, IL, USA

[11]Department of Biology, College of Charleston, Charleston, SC, USA

[12]Clinical Research Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

[13]Children's Tumor Foundation, New York, NY, USA

[14]NF Kinder, NF Patients United, Vienna, Austria

Abstract

Background/Aims: Individuals with neurofibromatosis 1 (NF1) may experience changes in their appearance due to physical manifestations of the disorders and/or treatment sequelae.

Corresponding Author: Vanessa L. Merker, PhD, Massachusetts General Hospital, 55 Fruit St., BHX 148, Boston, Massachusetts 02114, Phone: 617-724-6544, vmerker@mgh.harvard.edu.

Appearance concerns related to these physical changes can lead to psychological distress and poorer quality of life. While many NF1 clinical trials focus on assessing changes in tumor volume, evaluating patients' perspectives on corresponding changes in symptoms such as physical appearance can be key secondary outcomes. We aimed to determine if any existing patient-reported outcome (PRO) measures are appropriate for evaluating changes in appearance concerns within NF1 clinical trials.

Methods: After updating our previously published systematic review process, we used it to identify and rate existing PRO measures related to disfigurement and appearance. Using a systematic literature search and initial triage process, we focused on identifying PRO measures that could be used to evaluate changes in appearance concerns in plexiform or cutaneous neurofibromas clinical trials in NF1. Our revised Patient-Reported Outcomes Rating and Acceptance Tool for Endpoints then was used to evaluate each published PRO measures in five domains, including (1) respondent characteristics, (2) content validity, (3) scoring format and interpretability, (4) psychometric data and (5) feasibility. The highest-rated PRO measures were then re-reviewed in a side-by-side comparison to generate a final consensus recommendation.

Results: Eleven measures assessing appearance concerns were reviewed and rated; no measures were explicitly designed to assess appearance concerns related to NF1. The FACE-Q Craniofacial Module – Appearance Distress scale was the top-rated measure for potential use in NF1 clinical trials. Strengths of the measure included that it was rigorously developed, included individuals with NF1 in the validation sample, was applicable to children and adults, covered item topics deemed important by NF1 patient representatives, exhibited good psychometric properties, and was feasible for use in NF1 trials. Limitations included a lack of validation in older adults, no published information regarding sensitivity to change in clinical trials, and limited availability in languages other than English.

Conclusions: The Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) Patient-Reported Outcome working group currently recommends the FACE-Q Craniofacial Module - Appearance Distress scale to evaluate patient-reported changes in appearance concerns in clinical trials for NF1-related plexiform or cutaneous neurofibromas. Additional research is needed to validate this measure in people with NF1, including older adults and those with tumors in various body locations, and explore the effects of non-tumor manifestations on appearance concerns in people with NF1 and schwannomatosis.

Keywords

Neurofibromatosis 1; appearance concerns; disfigurement; patient reported outcomes; plexiform neurofibroma; cutaneous neurofibroma; physical appearance; patient-centered outcomes research; clinical trials

Background

Neurofibromatosis 1 (NF1) is a rare, autosomal dominant disorder associated with various physical deformations.[1, 2] As defined by recently revised diagnostic criteria, individuals with NF1 can have visible differences such as café au lait spots (coffee-colored birthmarks), skin fold freckles, cutaneous neurofibromas (benign tumors on the skin), plexiform neurofibromas (diffuse tumors that grow along nerves) and scoliosis (curvature of the

spine).[3] Among the most burdensome features of NF1-associated neurofibromas are their disfiguring appearance and negative impact on psychological well-being.[4] Research indicates that individuals with NF1 feel less attractive, less self-confident, and more insecure with their bodies compared to normative data.[5, 6]

A recent systematic review highlighted the significant increase in active and planned clinical trials for neurofibromatosis-associated tumors, including plexiform neurofibromas and cutaneous neurofibromas.[7] Plexiform neurofibromas occur in up to 50% of individuals with NF1 and can cause severe morbidity, including pain, disfigurement, motor dysfunction, and airway obstructions. Cutaneous neurofibromas occur in almost all adults with NF1, with the number of individual tumors ranging widely (from a few to multiple thousands) and typically increasing with age.[2] Given the significant impact of these tumors on appearance, the ability to measure the perceived disfigurement and appearance concerns of individuals with NF1 enrolled in clinical trials is paramount. Disfigurement was the most commonly reported tumor-related complication at baseline in a recent phase 2 trial of selumetinib for the treatment of inoperable plexiform neurofibromas in children with NF1, occurring in 88% of participants.[8] While subjective improvements in appearance were commonly reported by patients and parents in relation to their Global Impression of Change ratings,[8] no patient-reported outcome (PRO) measures specifically assessing appearance were included in the trial.[8]

PRO measures are commonly used to assess the benefit or risk of treatment from the patient's perspective.[9] The Food and Drug Administration requires that clinical trials demonstrating partial neurofibroma shrinkage also include evidence that treatment improves how patients feel or function as part of the drug approval process, making PRO measures an essential feature of NF1 clinical trials.[10] To date, most NF1 clinical trials that have incorporated PRO measures have focused on patients' health-related or disease-specific quality of life or specific symptoms such as pain and physical functioning.[11, 12] While these measures may include individual questions relating to physical appearance, PRO measures that specifically measure changes in disfigurement or appearance concerns generally have not been included in NF1 clinical trials. However, appearance-related PRO measures, such as a five-point rating scale to assess changes in the noticeability of vitiligo[13] and the Patient Reported Photonumeric Cellulite Severity Scale,[14] have been used as efficacy endpoints in clinical trials for other conditions leading to regulatory approvals by the Food and Drug Administration, highlighting the value and feasibility of appearance-related PRO measures.

Several terms are used to describe appearance-related constructs that may be assessed by PRO measures, including *appearance comparison*, *body satisfaction, body dissatisfaction*, *body image, disease visibility, disfigurement*, and *satisfaction with appearance. Body satisfaction*, *body dissatisfaction*, and *satisfaction with appearance* all refer to an individual's satisfaction or dissatisfaction with their general appearance or with specific body parts; the latter term has been used to measure social discomfort relative to acquired disfigurement from disease (e.g., Jewett et al., 2017). *Disease visibility* has been defined in NF1 literature as "the appearance of the person fully dressed and how readily symptoms could be perceived in impersonal interaction" (Ablon, 1996), and has been used to assess disability from skin disease (e.g., Chren et al., 1996). These terms overlap in that they assess an individual's

appearance but have differing applicability to an individual's self-reported assessment of their appearance as it is affected by a medical condition (rather than assessments of general attractiveness or physical fitness). Given the range of terms in the literature, we defined our focus broadly as *appearance concerns*, or an individual's self-reported concern with a visible difference that impacts their aesthetic, functional, and/or psychological well-being. This project aimed to identify and review existing PRO measures of appearance concerns with the goal of recommending a measure that would be appropriate for assessing changes in tumor-related disfigurement in NF1 clinical trials.

Methods

The Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) PRO working group was established to support the identification and use of appropriate PRO measures in neurofibromatosis and schwannomatosis clinical trials using a consensus-based and scientifically rigorous process.[11] REiNS PRO group members include clinicians and researchers with expertise in neurofibromatosis and schwannomatosis (including pediatrics, neurology, genetics, psychology, speech-language pathology, and health services research) and patient representatives (including adults with neurofibromatosis and schwannomatosis and caregivers of children with neurofibromatosis and schwannomatosis). We used a systematic process to identify, review, and rate existing PRO measures related to appearance concerns. This process was based on previously described procedures,[15–17] with modifications to the search strategy and rating forms to align our process with the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative and to formalize the process for patient representative input.[18, 19]

Development of updated PRO rating forms

We modified our previous rating form – the PRO Rating and Acceptance Tool for Endpoints[11] to: 1) explicitly identify domains related to content validity and scoring interpretability, 2) add relevant rating criteria within these two domains aligned with international standards,[19] and 3) remove "use in published studies" as a separate domain and instead extract relevant data from published studies to support ratings in other domains. We also created a companion user guide to systematize data extraction related to each PRO measure; this guide included detailed examples of extracted data and relevant instructional material from other published criteria to evaluate health measures.[19–21]

The revised PRO Rating and Acceptance Tool for Endpoints consists of five domains (Table 1), including: (1) respondent characteristics; (2) content validity (including assessment of the PRO measure development process and expert-assessed relevance, comprehensiveness, and comprehensibility of items as they pertain to neurofibromatosis and schwannomatosis clinical trials); (3) scoring format and interpretability (e.g., description of response options, type of scores available, availability of normative data, and rates of missing data and floor/ceiling effects in published studies); (4) psychometric data (reliability, structural validity, construct validity, criterion validity, and responsiveness); and (5) feasibility (cost, length, recall period used, ease of administration and scoring, languages available). Each criterion is rated on a scale of 0 (no or poor data) to 3 (solid published data supporting use in

neurofibromatosis and schwannomatosis trials) in increments of 0.25 points by individual REiNS group members. After a group discussion of each measure over videoconference, individuals may update their scores, which are then averaged to provide an overall group rating.

In addition, we created a companion PRO Rating and Acceptance Tool for Endpoints form for lay reviewers to facilitate patient representative involvement in the measure review process (Table 2). This form was iteratively refined based on feedback from our group's patient representatives to highlight the specific domains of the PRO Rating and Acceptance Tool for Endpoints form that benefit from patient representative input (i.e., content validity, adequacy of response options, and feasibility) and explain these domains in lay terms. Given that most PRO measures identified within relevant trial endpoint domains are not designed for or tested in neurofibromatosis and schwannomatosis patients, input from patient representatives is crucial to judge the content validity and feasibility of these measures for the neurofibromatosis and schwannomatosis population. Patient representatives were instructed to base their feedback on their own (or their child's) experience completing the measures, as well as how they believe the PRO measures may function for others with neurofibromatosis and schwannomatosis. Patient representatives' comments are featured during group discussion of each measure's content validity, scoring interpretability, and feasibility, and their numerical ratings are included in the overall group rating.

Identification of candidate PRO measures of appearance concerns

A preliminary search for PRO measures assessing disfigurement was conducted by one group member and results were presented to the group in a mini-review; as this process identified a limited number of relevant measures, we expanded our search to examine appearance concerns more broadly. Initial systematic search criteria adapted from Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative were used to identify a list of candidate measures for the construct of *appearance concerns*.[18] The instruments of interest were PRO measures, rather than parent or caregiver proxy forms, with a preference for self-administered questionnaires applicable across the lifespan, as plexiform and cutaneous neurofibroma clinical trials include both children and adults. An initial search of PubMed and PsycINFO databases on July 16, 2020 using the following search string [disfigurement OR "body image" OR appearance OR "body satisfaction" OR cosmetic out* OR "appearance satisfaction" AND questionnaire or survey or scale or instrument AND neurofibromatosis] yielded 49 manuscripts describing seven instruments: the Perceived Stigma Questionnaire, Social Comfort Questionnaire, Body Esteem Scale for Adolescents and Adults, Subjective Happiness Scale, Derriford Appearance Scale, The Self-Description Questionnaire I and Skindex. In reviewing the measures retrieved, many did not appear appropriate for our intended context of use. Therefore, we expanded the search and removed the key term of *neurofibromatosis* to determine if a wider search would yield instruments that would be more appropriate for our clinical population. Additional searches of PubMed and PsycINFO were conducted October 26 and November 4, 2020, yielding a total of 38 measures of appearance for potential full review (Figure 1; see Appendix A in the supplemental material for full search strategy and list of retrieved measures).

PRO measure rating and recommendation process

At least two group members reviewed each retrieved PRO measure to determine whether the measure assessed the intended construct of appearance concerns and could be applied to evaluate tumor-related disfigurement; any discrepancies were discussed until consensus was reached. All identified measures of appearance concerns that passed the triaging stage were reviewed and rated using our updated PRO Rating and Acceptance Tool for Endpoints forms for scientific and lay reviewers. The group then re-reviewed the three highest rated PRO measures in a detailed side-by-side comparison to ensure comparable numerical ratings and provide an opportunity to discuss the strengths and limitations of each measure.

When deciding on our ultimate recommendations, we focused on identifying PRO measures to evaluate change in appearance concerns within NF1 clinical trials of plexiform and cutaneous neurofibromas rather than for descriptive studies, clinical trials of general chronic illness, or studies of NF1 manifestations that represent less pressing clinical trial targets (such as café-au-lait macules). As no measures were specifically designed for people with NF1, the group's assessment of the relevance and importance of item content to individuals with NF1 was a priority. NF1 clinical trials for plexiform and cutaneous neurofibromas involve a wide age range of individuals, many of whom have learning disabilities, requiring easy to understand measures with content applicable across the lifespan. For this reason, we prioritized measures suitable for both children/adolescents and adults rather than a single age group. Psychometric properties adequate to support the use of measures as clinical trial endpoints and the feasibility of incorporating measures into multicenter, international clinical trials were also considered.

Results

Measures of disfigurement

The working group discussed several existing rating scales for disfigurement during the mini-reviews. However, none were found to be adequate for assessing tumor-related disfigurement in NF1 clinical trials. For example, several authors have modified the clinician-reported Ablon scale to a patient-reported format assessing the visibility of NF1 manifestations when clothed on a 3-point scale.[22, 23] Similarly, Chen et al. used a 9-point rating scale that asked observers to consider the visibility of the disfigured area[24] and Kleve et al. had participants rate how noticeable the disfigured area was to themselves and to others.[25] However, self-reports of disfigurement may encapsulate more than just how visible or noticeable a tumor is to others, such as when an area of disfigurement is typically covered by clothing but still bothersome to the individual. In addition, Lyford-Pike et al. used an 11-point scale on which observers rated the disfigurement of people with facial paralysis.[26] While the 11-point scale is potentially useful for NF1 clinical trials, it was not used to rate body parts other than the face, and there was no self-report version. For these reasons, we determined that a new disfigurement rating scale was necessary for plexiform and cutaneous neurofibromas clinical trials in NF1.

Measures of appearance concerns

Of the 38 identified measures, 27 were excluded during the initial triage process. Twenty-two measures were excluded because the item content was restricted to a specific condition (e.g., eating disorders, burns, etc.) or intervention (e.g., weight loss, cosmetic surgery such as breast reconstruction, etc.) that was not aligned with our project aim. Two measures were excluded because they focused on self-concept or happiness, respectively, rather than appearance concerns and two measures were excluded because they were not self-report. Finally, one non-validated scale that assessed self-reported disease visibility in patients with neurofibromatosis was excluded because the measure content and psychometric data were not published and thus was not evaluable.[27] Eleven measures assessing appearance concerns were advanced to full group reviews and ratings; their strengths and limitations are shown in Table 3.

The top three rated measures (the FACE-Q Craniofacial Module Appearance Distress Scale,[28] the Centre for Appearance Research Valence Scale,[29] and the Body Image Scale[30]) were re-reviewed head-to-head before arriving at a consensus recommendation for the FACE-Q, which is presented in detail below. While the Centre for Appearance Research Valence Scale (CARVAL) and Body Image Scale (BIS) had several strengths, both scales were developed for adults only and contained questions that may be less relevant to or appropriate for younger children. The Centre for Appearance Research Valence Scale was validated in a healthy, predominantly female white population, and is only available in English, limiting its versatility.[31] There was a paucity of other psychometric data on the Body Image Scale, and item content did not differentiate between disease impact and the impact of treatment/surgery, which could be confusing or irrelevant to NF1 study participants.[32]

FACE-Q Craniofacial Module – Appearance Distress Scale

The REiNS PRO working group recommended the FACE-Q Craniofacial Module – Appearance Distress scale to assess the psychosocial impact of any type of NF1-tumor related disfigurement in clinical trials for children and adults with NF1 (Table 4). This self-report PRO measure is one of the health-related quality of life scales in the FACE-Q Craniofacial Module. This module is part of the larger FACE-Q measure, which consists of several modules assessing outcomes for various conditions affecting the face, such as head and neck cancer, skin cancer, and paralysis (https://qportfolio.org/). Specifically, the FACE-Q Craniofacial Module (https://qportfolio.org/face-q/craniofacial/) was developed for individuals with conditions associated with a visible or functional facial difference. It consists of 27 independent scales that assess four domains: 1) appearance of specific parts of the face (e.g., nose, smile, eyes, lips), 2) functions of the face (e.g., speech, eating/drinking, breathing), 3) adverse events (e.g., ears, eyes, face), and 4) health-related quality of life (e.g., appearance distress, psychological, social).

The FACE-Q Craniofacial Module was developed for patients 8 to 29 years of age. The Appearance Distress scale of this module consists of 8 items rated on a 4-point scale (1=always to 4=never) that ask about social concerns, such as going out in public, meeting people, covering up, and people staring, as well as psychological issues, including feeling

unhappy, self-conscious, or disliking one's appearance, in the past week.[28, 33] Although these items were developed to assess facial differences, no part of the instructions or items specifically mention the face; thus, our group decided that it would be valuable to pilot this scale in NF1 to assess appearance-related distress associated with tumors in any body area. For example, the items (e.g., "I feel unhappy about how I look") could apply to tumors on the face, trunk, or limbs. Scores on each item are summed to produce a total raw score, which is converted to a 0–100 metric derived from Rasch analysis. Higher scores indicate better outcomes. This scale is brief, has a simple format, is easy to read (Flesch–Kincaid grade reading level=3.2), and is available in several languages.[28]

The authors conducted rigorous qualitative research to develop the content of the scales followed by Rasch measurement theory analysis to evaluate its psychometric properties in a large sample of 1,495 participants, including 31 with NF1 (2.9%).[33] The reliability of the health-related quality of life scales, including Appearance Distress, was high with Pearson Separation Index values ≥0.83 with and without extremes, and Cronbach alpha values ≥0.87 with and without extremes. The validity of the Appearance Distress scores was supported by moderate correlations (r = .37–.59) with the specific facial appearance ratings and higher correlations between scales within domains (e.g., health-related quality of life) than with scales in other domains.

Limitations of the FACE-Q Appearance Distress scale for use in NF1 trials identified by the REiNS PRO group are the lack of use and normative data in older adults (>29 years), no published information for this measure regarding sensitivity to change, and the need for translation into additional languages. In addition, the scale was developed for individuals with facial differences, so although the items appear relevant to any type of visual difference, this scale must be further evaluated in individuals with NF1 and disfigurement in the face and other body areas. The use of the FACE-Q requires the completion of a licensing agreement, and, due to copyright laws, no modifications to the items or scale can be made. Finally, there is no cost for non-profit academic organizations to use the scale, but for-profit companies must pay a licensing fee.

Discussion

The purpose of this study was to determine if any published PRO measures were appropriate for evaluating changes in appearance concerns in NF1 clinical trials. Given the substantial impact of NF1-related tumors on appearance, and the importance of documenting clinical benefit with tumor shrinkage in treatment trials, there is an unmet need to identify PRO measures to evaluate changes in NF1-related disfigurement. While some PRO measures developed specifically for NF1 include items related to disfigurement, none have appearance-specific subscales that are specific to plexiform and cutaneous neurofibromas tumors.[34, 35] In the absence of any measures developed specifically to measure tumor-related disfigurement in NF1 clinical trials, the REiNS PRO group rigorously reviewed and rated 11 existing measures of appearance concerns for their potential utility in NF1 clinical trials using its updated PRO Rating and Acceptance Tool for Endpoints methodology.

From this process, the REiNS group rated the FACE-Q Craniofacial Module Appearance Distress Scale as the most appropriate existing measure for use in NF1 clinical trials. Advantages of this measure include that it was rigorously developed, individuals with NF1 were included in the validation sample, it is suitable for children and young adults, it included items considered important by REiNS patient representatives, it has good psychometric properties, and it is feasible for use in clinical trials. The main limitation to its use with individuals with NF1 was that it was validated only in children and young adults with facial differences; however, the content of the items appears appropriate for any kind of appearance-related distress and for older adults. As such, it would be beneficial to validate the measure in individuals with non-facial plexiform and cutaneous neurofibromas and a wider age range that included older adults. To administer the FACE-Q Craniofacial Module Appearance Distress Scale in NF1 trials, participants would need to be instructed to focus on their appearance related only to their tumors, and not include other non-tumor-related conditions or temporary side effects of the study drug that might impact appearance, such as rashes from MEK inhibitors.

If validated in people with NF1, the FACE-Q Craniofacial Module Appearance Distress Scale could be used as a secondary endpoint to demonstrate potential treatment benefits related to improved tumor appearance in trials for plexiform neurofibromas or trials of systemic therapy for cutaneous neurofibromas. The Numerical Rating Scale-11 (NRS-11) and Pain Interference Index were similarly used to demonstrate treatment benefit in the domain of pain in the recent registration trial of selumetinib for the treatment of plexiform neurofibromas in children with NF1.[8] In plexiform neurofibroma trials, PRO measures for specific tumor-related complications (e.g., disfigurement, motor dysfunction) may be given only to those patients who endorse or are at high risk for these complications at baseline to reduce overall PRO measure administration burden.[36] For cutaneous neurofibroma trials, the FACE-Q and other measures of disfigurement may be used in conjunction with other measures, such as the cNF-Skindex, which assesses cutaneous neurofibroma-related quality of life more broadly (including effects on pain, pruritus, social functioning, and emotional functioning).[34] However, cutaneous neurofibroma eDelphi results presented at the REiNS 2022 Summer Meeting revealed that appearance was more often rated as a key cutaneous neurofibroma trial outcome by patients and family members than physical symptoms like pain and pruritus.[37] Therefore, an appearance-focused measure may be better suited to demonstrating patient-relevant treatment benefits. By using the FACE-Q or other measures that focus solely on appearance rather than multiple cutaneous neurofibroma-related quality of life domains, it may be easier to demonstrate change in PRO measure scores related to disfigurement in response to treatment.

Patient representatives offer a unique perspective that is integral to the work of the REiNS Collaboration and is critical for evaluating and developing outcome measures for use in neurofibromatosis and schwannomatosis clinical trials.[38] Patient representatives' input was crucial in defining our project's construct of appearance concerns, and in choosing and reviewing measures. In accordance with the recent cutaneous neurofibroma eDelphi results, from early in this process, patient representatives in the PRO working group strongly recognized that it is necessary to specifically assess disfigurement and appearance concerns rather than relying solely on broader health-related quality of life measures. They

ascertained that appearance is a central concern of patients that has a disproportionate negative impact on quality of life, and therefore should be considered separately, particularly in clinical trials focusing on tumors in NF1. Moreover, the patient representatives argued that an outcome measure focused on a patient's perception of their tumor-related appearance could capture significant impacts from treatment that might correspond to only subtle changes in tumor size, number of tumors, or summary measures of quality of life. Discussion with patient representatives confirmed that it was important that this perspective come from the patient directly rather than a parent or other caregiver, as appearance concerns may be judged differently depending on the informant; this supported our group's decision not to evaluate proxy-reported measures. Patient representatives also helped to identify items on candidate PRO measures that may be confusing, as well as items that may have unintended interpretations in the context of NF1 or were insensitively worded in a way that could potentially have an adverse impact on the patient's mental health.

Conclusions

While the FACE-Q Craniofacial Module Appearance Distress Scale shows promise for capturing appearance-related concerns in NF1 clinical trials, this measure needs to undergo a qualitative assessment with NF1 patient input, such as through focus groups and cognitive interviews.[39] The results of this assessment will determine whether this measure or a newly created tool should be evaluated in NF1 clinical trials to assess its sensitivity to changes occurring with treatment and determine its minimal clinically important difference.[40] Moving forward, a validated appearance concerns measure could be used as a secondary outcome in NF1 clinical trials. In addition to this work validating a measure of appearance concerns, future research could develop patient-reported ratings of disfigurement for plexiform and cutaneous neurofibroma and review potential PRO measures to assess the effects of non-tumor manifestations, such as scoliosis, on appearance and appearance concerns in people with NF1.

Acknowledgements

The authors acknowledge and thank additional members of the REiNS Patient-Reported Outcomes working group who participated in discussions to update the PRO Rating and Acceptance Tool for Endpoints forms and review the PRO measures assessing disfigurement and appearance concerns discussed in this article: Kim Koetsier, Carlos Romo, William Riter, Krizelle Alcantara, Brad Welling, Karin Walsh, Barbara Franklin, Jim Tonsgard, Jane Grabowski, and Kimberley Marrs. This manuscript is submitted on behalf of the REiNS International Collaboration, the full membership of which is listed in Appendix B.

Funding

This research was supported in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute. The authors also acknowledge the support of the Children's Tumor Foundation for publication of this manuscript.

Declaration of conflicting interests

VLM, PLW, and SM have a pending application for funding with the Neurofibromatosis Therapeutic Acceleration Program to validate measures of NF1-related disfigurement and appearance concerns based on the submitted work. The remaining authors declare no relevant conflicts of interest.

Appendix A.: Full search strategy and list of retrieved appearance measures

PubMed search completed on October 26th, 2020 – 110 hits

("satisfaction with appearance" OR "appearance concern") AND (patient[tiab] OR self[tiab] OR child[tiab] OR parent[tiab] OR carer[tiab] OR proxy[tiab]) AND ((report[tiab] OR reported[tiab] OR reporting[tiab]) OR (rated[tiab] OR rating[tiab] OR ratings[tiab]) OR (assessed[tiab] OR assessment[tiab] OR assessments[tiab]))) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])))

PsycINFO search completed on October 27th, 2020 & Nov 4, 2020 – 51 hits

- "satisfaction with appearance" OR "appearance concern"
- AND patient OR self OR child OR parent OR carer OR proxy
- AND ((report OR reported OR reporting) OR (rated OR rating OR ratings) OR (assessed OR assessment OR assessments))
- AND index OR indices OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR profile OR profiles OR scale OR scales OR score OR scores OR status OR survey OR surveys

After merging and de-duplicating search results, the following outcome measures related to appearance were identified. Measures advanced for further review and rating are presented with an asterisk; measures not advanced for review have a brief note on limitations.

1. Ablon's Visibility Scale – not a self-reported measure
2. Body Attitude Test – developed for individuals with eating disorders
3. Body Concealment Scale for Scleroderma – developed for individuals with scleroderma
4. Body Esteem Scale; Body Esteem Scale Revised – developed for individuals with eating disorders
5. Body Image Scale (BIS)*
6. Body Images Coping Strategies Inventory – developed for other aspect of body image (e.g., body image disturbances)
7. Body Image Quality of Life Inventory (BIQLI)*
8. Body Image Concern Inventory (BICI)*
9. Body Image Disturbance Questionnaire – developed for individuals with craniofacial concerns

10. Body Image Questionnaire – developed for other aspect of body image (e.g., body image with focus on gender expression)
11. BODY-Q*
12. Body Satisfaction Visual Analog Scale – developed for individuals with eating disorders
13. Body Satisfaction Scale – developed for individuals with eating disorders
14. Body Uneasiness Test - developed for individuals with eating disorders
15. BREAST-Q – developed specifically for breast reconstruction.
16. Brief Satisfaction With Appearance Scale for Systemic Sclerosis (Brief-SWASS) – developed for individuals with systemic sclerosis
17. CARVAL*
18. CARSAL*
19. Derriford 24 (DAS24)*
20. Derriford 59 (DAS59)*
21. Eating Disorder Inventory (EAT-3) – developed for individuals with eating disorders
22. FACE-Q*
23. Granström's Scale – measure items not published/validated
24. Human Figure Drawing Test - not a self-report measure
25. Multidimensional Body-Self Relations Questionnaire-Appearance Scales (MBSRQ-AS)*
26. Multidimensional Body-Self Relations Questionnaire-Body Areas Satisfaction Scale – items assess physical attractiveness
27. Oral Health Impact Profile – developed for individuals with dental concerns
28. Patient Scar Assessment Questionnaire – assesses patients with scarring
29. Perceived Stigmatization Scale (PSQ) – developed for individuals with burns
30. Piers-Harris Children's Self-Concept Scale 2 – child/adolescent forms only; also evaluates self-concept development rather than appearance
31. Satisfaction with Appearance Scale (SWAP); Adapted SWAP – developed for individuals following traumatic brain injury
32. SCINEXA – assesses intrinsic aging
33. Self-Perception Profile for Children, Self-Perception Profile for Adolescents and Harter's Self-Perception Profile for Adults (SPP)*
34. Sociocultural Attitudes Toward Appearance Questionnaire - assesses weight loss

35. Social Physique Anxiety Scale – developed for other aspect of body image (e.g., anxiety developed when people feel their appearance is judged by others)
36. Skindex – developed for dermatology conditions
37. State Body Satisfaction – assesses with loss
38. The Subjective Happiness Scale – child/adolescent forms only; also evaluates happiness rather than appearance

Appendix B.: Members of the REINS International Collaboration

The Response Evaluation in Neurofibromatosis and Schwannomatosis (REINS) International Collaboration acknowledges the efforts of all current working group members, who include:

Sarah Adsit, Shivani Ahlawat, Krizelle Alcantara, Rosemary Anderson, Kara Anstett, Lisa Aouad, Rob Avery, Carolina Barnett-Tapia, Belinda Barton, Catherine Batlle, David Bedwell, Tammy Benson O'Brian, Dale Berg, Christina Bergqvist, Yemima Berman, Chetan Bettegowda, Jaishri Blakeley, Rob Brainin, Sara Brebbia Dirksen, Frank Buono, Susan Buono, Wenli Cai, Cynthia Campen, Ashley Cannon, Kelly Carpenter, Alexandra Cellucci, Nicolas Champollion, Long-Shen Chang, Peter de Blank, Allison del Castillo, Laura Didier, Eva Dombi, Gregg Erickson, Gareth Evans, Khaled Ezzedine, Onno Faber, Laura Fertitta, Michael Fisher, Barbara Franklin, Edith Garrett, Madalyn Gibson-Williams, Marco Giovannini, Dani Glad, Deborah Gold, Adam Goodkind, Kayo Goto, Jane Grabowski, Vito Grasso, Andrea Gross, Lorenza Grotti, David Gutmann, Diana Haberkamp, Kristine Haebich, Cindy Hahn, Oliver Hanemann, Cynthia Hingtgen, Matt Hocking, Eric Hunt, Maureen Hussey, Kelly Janke, Jennifer Janusz, Justin Jordan, Valerie Kaer, Matthias Karajannis, Kim Keeling, Bob Kesterson, Aerang Kim, Hannah Kim, Yoori Kim, Bonnie Klein-Tasman, Laura Klesse, Kimberley Koetsier, Edina Komlodi-Pasztor, Noelle Larson, Andre Leier, Erica Leif, Andres Lessing, Robert Listernick, Elana Loftspring, Sharon Loftspring, K. Ina Ly, Linda Manth, Staci Martin, Sherri McKnight, Miranda McManus, Vanessa Merker, Elizabeth Morehouse, Irene Moss, Lara Mukabenov, Patricia Newman, Beverly Oberlander, Jonathan Payne, Drea Peterson, Dominque Pichard, Rene Pierpoint, Scott Plotkin, Dorothy Pope, Linda Popplewell, Natalie Pride, Marcus Ratley, Steven Rhodes, Jonathan Rios, Inka Ristow, Bill Riter, Claas Rohl, Carlos Romo, Adam Rosenberg, Tena Rosser, Jessica Samblanet, Kavita Sarin , Herb Sarnoff, Natalie Schneider, Elizabeth Schorry, Mary Sell, Steven Sheard, Dani Silverman, Brittany Simpson, Melissa Smith, Connie Sorman, Verena Staedtke, Pete Stavinoha, Anat Stemmer-Rachamimov, David Stevenson, Taylor Sundby, Heather Thompson, Nicole Ullrich, Meena Upadhyaya, Marjan van de Vrie, Dave Viskochil, Jeremie Vitte, James Walker, Jo Wallace, Deann Wallis, Karin Walsh, Shannon Weaver, Hannah Weisman, Bradley Welling, Brigitte Widemann, Tracy Wirtanen, Pierre Wolkenstein, Pamela Wolters, Mary Achtziger Wong, Jiangbing Zhou.

References

1. Gutmann DH, Ferner RE, Listernick RH, et al. Neurofibromatosis type 1. Nat Rev Dis Primers 2017; 3: 17004. [PubMed: 28230061]

2. Plotkin SR and Wick A. Neurofibromatosis and Schwannomatosis. Semin Neurol 2018; 38: 7385.
3. Legius E, Messiaen L, Wolkenstein P, et al. Revised diagnostic criteria for neurofibromatosis type 1 and Legius syndrome: an international consensus recommendation. Genetics in Medicine 2021; 23: 1506–1513. [PubMed: 34012067]
4. Chamseddin BH and Le LQ. Management of cutaneous neurofibroma: current therapy and future directions. Neuro-Oncology Advances 2019; 2: i107–i116. [PubMed: 32642736]
5. Smith KB, Wang DL, Plotkin SR, et al. Appearance concerns among women with neurofibromatosis: examining sexual/bodily and social self-consciousness. Psycho-Oncology 2013; 22: 2711–2719. [PubMed: 23873824]
6. Barke J, Coad J and Harcourt D. The Role of Appearance in Adolescents' Experiences of Neurofibromatosis Type 1: A Survey of Young People and Parents. Journal of Genetic Counseling 2016; 25: 1054–1062. [PubMed: 27001554]
7. Roman Souza G, Abdalla A and Mahadevan D. Clinical trials targeting neurofibromatosis-associated tumors: a systematic review. Neuro-Oncology Advances 2022; 4: vdac005.
8. Gross AM, Wolters PL, Dombi E, et al. Selumetinib in Children with Inoperable Plexiform Neurofibromas. N Engl J Med 2020; 382: 1430–1442. [PubMed: 32187457]
9. Food and Drug Administration. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims (2009, accessed 3 February 2023).
10. Gross AM and Widemann BC. Clinical trial design in neurofibromatosis type 1 as a model for other tumor predisposition syndromes. Neurooncol Adv 2020; 2: i134–i140. [PubMed: 32642739]
11. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes in neurofibromatosis and schwannomatosis clinical trials. Neurology 2013; 81: S6–14. [PubMed: 24249806]
12. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes of pain and physical functioning in neurofibromatosis clinical trials. Neurology 2016; 87: S4–S12.
13. Batchelor JM, Tan W, Tour S, et al. Validation of the Vitiligo Noticeability Scale: a patient-reported outcome measure of vitiligo treatment success. Br J Dermatol 2016; 174: 386–394. [PubMed: 26409256]
14. Kaufman-Janette J, Joseph JH, Kaminer MS, et al. Collagenase Clostridium Histolyticum-aaes for the Treatment of Cellulite in Women: Results From Two Phase 3 Randomized, Placebo-Controlled Trials. Dermatol Surg 2021; 47: 649–656. [PubMed: 33840781]
15. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes in neurofibromatosis and schwannomatosis clinical trials. Neurology 2013; 81: S6–14. [PubMed: 24249806]
16. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes of pain and physical functioning in neurofibromatosis clinical trials. Neurology 2016; 87: S4–S12.
17. Wolters PL, Vranceanu AM, Thompson HL, et al. Current Recommendations for Patient-Reported Outcome Measures Assessing Domains of Quality of Life in Neurofibromatosis Clinical Trials. Neurology 2021; 97: S50–S63. [PubMed: 34230198]
18. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research 2009; 18: 1115–1123. [PubMed: 19711195]
19. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018; 27: 1159–1170. [PubMed: 29550964]
20. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60: 34–42. [PubMed: 17161752]
21. Lewis CC, Mettert KD, Stanick CF, et al. The psychometric and pragmatic evidence rating scale (PAPERS) for measure development and evaluation. Implementation Research and Practice 2021; 2: 263348952110373.
22. Page PZ, Page GP, Ecosse E, et al. Impact of neurofibromatosis 1 on Quality of Life: A cross-sectional study of 176 American cases. American Journal of Medical Genetics Part A 2006; 140A: 1893–1898.

23. Fleming J, Morgan O, Wong C, et al. Characterization of health concerns in people with neurofibromatosis type 1. Molecular Genetics & Genomic Medicine 2023; 11: e2077. [PubMed: 36444392]
24. Chen SC, Huang BS, Lin CY, et al. Psychosocial effects of a skin camouflage program in female survivors with head and neck cancer: A randomized controlled trial. Psychooncology 2017; 26: 13761383.
25. Kleve L, Rumsey N, Wyn-Williams M, et al. The effectiveness of cognitive-behavioural interventions provided at Outlook: a disfigurement support unit. Journal of evaluation in clinical practice 2002; 8: 387–395. [PubMed: 12421388]
26. Lyford-Pike S, Helwig NE, Sohre NE, et al. Predicting Perceived Disfigurement from Facial Function in Patients with Unilateral Paralysis. Plast Reconstr Surg 2018; 142: 722e–728e.
27. Granström S, Langenbruch A, Augustin M, et al. Psychological burden in adult neurofibromatosis type 1 patients: impact of disease visibility on body image. Dermatology 2012; 224: 160–167. [PubMed: 22508340]
28. Klassen A and Wong K. FACE-Q Craniofacial User Guide, https://qportfolio.org/wpcontent/uploads/2021/07/FACE-Q-CRANIOFACIAL-USERS-GUIDE-1.pdf (2020, accessed 26 September 2022).
29. Moss TP and Rosser BA. The Moderated Relationship of Appearance Valence on Appearance Self Consciousness: Development and Testing of New Measures of Appearance Schema Components. PLoS ONE 2012; 7: e50605. [PubMed: 23226326]
30. Hopwood P, Fletcher I, Lee A, et al. A body image scale for use with cancer patients. Eur J Cancer 2001; 37: 189–197. [PubMed: 11166145]
31. Moss TP, Lawson V and White P. Salience and Valence of Appearance in a Population with a Visible Difference of Appearance: Direct and Moderated Relationships with Self-Consciousness, Anxiety and Depression. PLoS One 2014; 9: e88435. [PubMed: 24516660]
32. Melissant HC, Neijenhuijs KI, Jansen F, et al. A systematic review of the measurement properties of the Body Image Scale (BIS) in cancer patients. Supportive Care in Cancer 2018; 26: 1715–1726. [PubMed: 29532245]
33. Klassen AF, Rae C, Riff W, et al. FACE-Q craniofacial module: Part 2 Psychometric properties of newly developed scales for children and young adults with facial conditions. Journal of Plastic, Reconstructive & Aesthetic Surgery 2021; 74: 2330–2340.
34. Fertitta L, Bergqvist C, Armand ML, et al. Quality of life in neurofibromatosis 1: development and validation of a tool dedicated to cutaneous neurofibromas in adults. Journal of the European Academy of Dermatology and Venereology 2022; 36: 1359–1366. [PubMed: 35412677]
35. Nutakki K, Varni JW and Swigonski NL. PedsQL Neurofibromatosis Type 1 Module for children, adolescents and young adults: feasibility, reliability, and validity. Journal of Neuro-Oncology 2018; 137: 337–347. [PubMed: 29273891]
36. Gross AM, Glassberg B, Wolters PL, et al. Selumetinib in children with neurofibromatosis type 1 and asymptomatic inoperable plexiform neurofibroma at risk for developing tumor-related morbidity. Neuro Oncol 2022; 24: 1978–1988. [PubMed: 35467749]
37. Bergqvist C, Berman Y, Cannon A, et al. REiNS Mini-symposium: Consensus meeting to define the minimal set of outcomes to be reported for clinical trials of cutaneous neurofibromas, https://ccrod.cancer.gov/confluence/display/REINS/Presentations?preview=/141000929/262045933/9.pdf (2022, accessed 20 December 2022).
38. Merker VL, Lessing AJ, Moss I, et al. Enhancing Neurofibromatosis Clinical Trial Outcome Measures Through Patient Engagement: Lessons From REiNS. Neurology 2021; 97: S4–S14. [PubMed: 34230208]
39. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. Value Health 2011; 14: 978–988. [PubMed: 22152166]
40. Jaeschke R, Singer J and Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989; 10: 407–415. [PubMed: 2691207]

41. Klassen AF, Rae C, Gallo L, et al. Psychometric Validation of the FACE-Q Craniofacial Module for Facial Nerve Paralysis. Facial Plast Surg Aesthet Med 2022; 24: 1–7. [PubMed: 33826408]

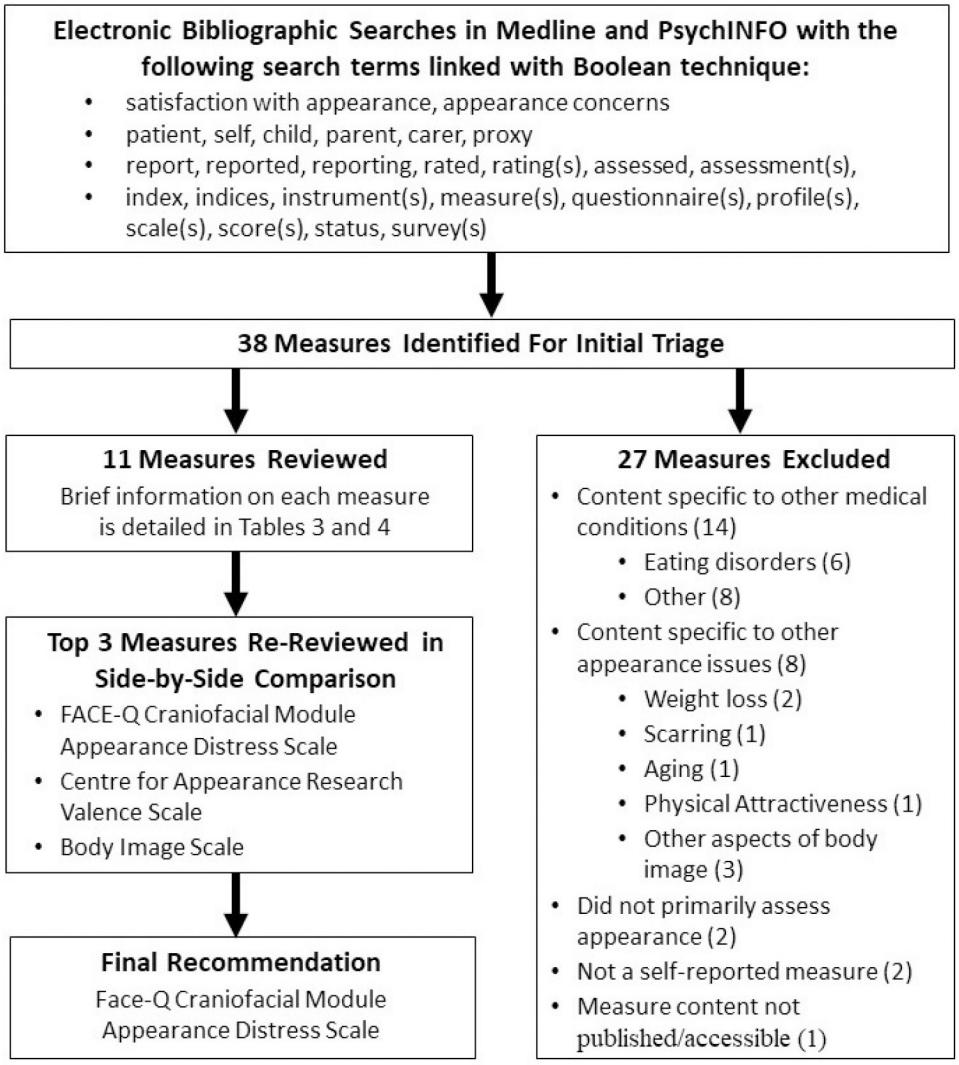

Figure 1.
Appearance Concerns PRO Measure Search, Review, and Rating Process

Table 1.

Updated PRO Rating and Acceptance Tool for Endpoints review criteria

Criteria[a]	**Measure Characteristics**[b]
Patient Characteristics	• Age range • Reporting type (e.g., self vs. proxy report)
Content Validity	*PRO Measure Development Process:* Describe whether and in what populations there was: • Concept elicitation with patients to generate/select items • Cognitive interviews or other pilot testing to refine measure • Any additional content validity studies *Expert-Assessed Content Validity for Neurofibromatosis/Schwannomatosis Trials:* • Description of domains and number of items in each (e.g., physical, social, emotional, cognitive) For the intended clinical trial context, are items • Relevant? • Comprehensive? • Comprehensible?
Scoring Format and Interpretability	• Item response wording and format (e.g., Likert scale, visual analog scale) • Types and range of scores available (e.g., raw/standardized, domain/total) • Normative and other reference groups (e.g., general, neurofibromatosis/schwannomatosis, other; number of subjects) • Missing Data (i.e., % of items and % of scores missing in published studies) • Floor/ceiling effects (i.e., % of study sample scoring highest or lowest score in published studies)
Psychometric Data	• Factor Analysis/Structural Validity • Reliability (e.g., internal consistency, test/retest) • Construct Validity (e.g., known groups, convergent, discriminative) • Criterion Validity (*for comparing short forms to full length measures only) • Responsiveness (including minimal clinically important difference, if available)
Feasibility	• Cost • Length (number of items and completion time) • Recall period assessed • Ease of administration (e.g., self vs. interviewer administered; assessor burden in administration and interpretation): • Original Language and available translations

[a] Each domain is scored on a scale of 0–3 in 0.25 increments, which are then averaged to produce a total score.

[b] After each group review of a measure, the group's overall impression of the pros/cons of the measure as it applies to neurofibromatosis/schwannomatosis clinical trials is also recorded.

Table 2.

Adapted PRO Rating and Acceptance Tool for Endpoints criteria for patient representatives

Criteria[a]	Measure Characteristics[b]
Content for neurofibromatosis and schwannomatosis clinical trials	• Does this questionnaire assess domains (topics) and items (questions) relevant to individuals with neurofibromatosis or schwannomatosis? • Is the content something that is likely to change over time after an intervention in a clinical trial? • Are there any topics that are important [for measuring change in neurofibromatosis/schwannomatosis clinical trials] that are missing? • Are the questions worded clearly and easy to understand? Were there any technical/medical terms that you didn't understand?
Scoring	• Are the response options and/or rating scale easy to understand? (For example, 0–10 scale or choices of "always/sometimes/never") • Are people likely to miss or skip any questions? For example, is it easy to accidentally skip a question? Did you skip some questions because they don't apply to you or you didn't know how to answer?
Feasibility	• Are the instructions easy to understand? Would you be able to fill out this questionnaire if it was given to you with no explanation? • Did you know what time period to consider when answering? Can you remember the answers to the question using this time period? • How long did it take to complete this questionnaire? _____ minutes? Is this a reasonable amount of time for an neurofibromatosis/schwannomatosis clinical trial?

[a]Patient representatives are provided with instructions for completing the form, and requested to provide both narrative comments and numerical ratings (scored on the same 0–3 scale as the main PRO Rating and Acceptance Tool for Endpoints form).

[b]There is also space at the end of the form to record any additional comments about the use of the questionnaire in neurofibromatosis/schwannomatosis clinical trials.

Table 3.

List of Additional PRO Measures Reviewed to Assess Tumor-related Appearance Concerns in NF1 Clinical Trials

Name of measure	Age range	Domain and Number of Items	Strengths for use in NF1 clinical trials	Limitations for use in NF1 clinical trials
Body Image Scale (BIS) [Hopwood et al., 2010]	≥18 years	Assessment of body image changes in patients with cancer; 10 items	Had patient input during development. Items relevant for NF1. Studies in different populations. Easy to administer. Multiple languages available.	Not studied in NF1. Only adults, not suitable for children or adolescents. No MCID available.
Body Image Quality of Life Inventory (BIQLI) [Cash and Fleming, 2001]	≥18 years	Impact of body image on aspects of life; 19 items	Studied in multiple countries, brief and simple to administer. Multiple languages available.	Several items geared towards weight and eating disorders, not applicable for NF1 trials. Not studied in NF1.
Body Image Concern Inventory (BICI) [Littleton et al., 2005]	≥18 years	Assessment of dysmorphic concern; 19 items	Studied in multiple countries and across different diseases and concerns. Items relevant to NF1. Good psychometric data. Multiple languages available.	Patients not involved in item development. Not studied in NF1. MCID not available.
BODY-Q [Klassen et al., 2016]	≥18 years	Assessment of appearance, patient experience regarding care, and quality of life; 8 items	Brief and simple to administer. Good methodology with scales refined through patient interviews. Free of charge for non-profit organizations and for clinical care. Multiple languages available.	Four-point Likert scale. Separate questions individually sampled, no overall analysis for assessments.
Centre for Appearance Research Valence Scale (CARVAL) [Moss and Rosser, 2012]	≥18 years	Measures how positively or negatively an individual measures their own appearance; 8 items	Brief and simple to administer. Items are relevant to the NF1 population. Good psychometrics. Free of charge.	Validated in healthy university student population that was predominantly female and white. Not studied in NF1. Only available in English.
Centre for Appearance Research Salience Scale (CARSAL) [Moss and Rosser, 2012]	≥18 years	Measures the extent to which appearance is important to a person; 5 items	Brief and simple to administer. Items are relevant to the NF1 population. Good psychometrics. Free of charge.	Validated in healthy university student population that was predominantly female and white. Not studied in NF1. Only available in English.
Derriford 24 (DAS24) [Moss, Harris, Carr, 2004]	Developed for ≥18 years; has been used in children aged 9+	Assesses adjustment to problems of disfigurement and visible difference; 24 items + 6 items about a specific feature	Brief, easy to administer. Multiple languages available.	Some items not applicable to NF1. Not as strongly validated as DAS59.
Derriford 59 (DAS59) [Harris, 1982]	≥16 years	Assesses appearance related distress and problems; 59 items + 4 items about a specific feature	Well validated, good psychometric data. Has been used in NF1 (e.g., Smith et al., 2013). Multiple languages available.	Some items not applicable to NF1. Long form, high reading level, concerning for individuals with NF1 and learning disabilities.
Multidimensio nal Body-Self Relations Questionnaire (MBSRQ) – Appearance Scales [Cash, 2018]	≥ 15 years	Assesses self-attitudinal aspects of body-image and satisfaction with discrete aspects of appearance; 34 items	Content relevant to adolescents and adults with NF1. Comprehensible, easy to administer and score. Multiple languages available.	Patients not clearly involved in item development. Not studied in NF1.
Self-Perception Profile (SPP) for Children, Adolescents, and Adults [Harter, 2012a; Harter, 2012b; Messer & Harter, 1986]	8–60 years	Assesses perceived competence in physical appearance; 36–50 items depending on whether profile for children, adolescents, or adults is used	Free of charge. Children version used in two NF1 studies. Multiple languages available.	Many scales for different ages. Response format difficult to understand, especially for children who are poor readers. Limited psychometric data for the adult scale and no data on responsiveness.

Table 4.

Evaluation of the FACE-Q Craniofacial Module Appearance Distress Subscale

Rating Criteria (ratings)	Measure Characteristics
Patient Characteristics (2.5)	• Age range is 8 to 29 years[33] – Self-report form – No observer-report form • The larger FACE-Q Craniofacial Module is validated for patients with facial paralysis, ages 8 to 81 years[41]
Content Validity (2.5)	• 8 items • Item content: Social concerns, such as going out, meeting people, covering up, and people staring, as well as psychological issues, including feeling unhappy, self-conscious, or disliking one's appearance • Recall period is over the past week • Followed recommended guidelines for PRO development using a multiphase mixed methods approach, including qualitative data • https://qportfolio.org/face-q/craniofacial/
Scoring Format and Interpretability (2.5)	• Items rated on a 4-point scale (1=always to 4=never) • Scores on each item are summed to produce a total raw score, which is converted to a score ranging from 0–100 based on Rasch analysis • Higher scores indicate better outcomes
Psychometric Data (2.5)	• Evaluated using the modern psychometric approach of Rasch Measurement Theory (RMT) • Psychometric properties studied in a large sample of 1495 participants with a range of facial conditions, including 31 with NF1 (2.9%) • Reliability of the FACE-Q scales, including Appearance Distress, was high (Pearson Separation Index values ≥0.83; Cronbach alpha values ≥0.87). • Validity supported by: 1) lower scores in subjects with a major difference in appearance/function; 2) moderate correlations with the specific facial appearance ratings; 3) higher correlations obtained between scales within domains (e.g., health-related quality of life) than with scales in other domains. • No minimal clinically important difference was available
Feasibility (2.75)	• Brief (5 minutes to complete), simple format, and easy to read (3.2 grade reading level) • English, Spanish, French, Dutch, and Portuguese; with permission, it can be translated into other languages • Free of charge for non-profit academic research and clinical care; a licensing fee required for use by "for-profit" organizations • To obtain a license use the following link: https://research.mcmaster.ca/industry-and-investors/technologiesavailable-for-licensing/questionnaire-request-form/