MDDT Summary of Evidence and Basis of Qualification Decision for
FACE-Q | Aesthetics©

Background

MDDT Name: FACE-Q | Aesthetics©

Submission Number: Q191161

Date of Submission: February 6, 2020

Contact:

Professor Anne Klassen
McMaster University
3N27, 1280 Main St W
Hamilton, ON
Canada, L8S 4L8
aklass@mcmaster.ca
Phone: 1-905-531-1033

Tool Description and Principle of Operation

FACE-Q | Aesthetics© is a patient-reported outcome (PRO) instrument that can be used to assess outcomes of aesthetic facial procedures from the patient's perspective. FACE-Q | Aesthetics© is comprised of 3 modules: Self-Perceived Facial Appearance, Health-Related Quality of Life (HRQL) and Adverse Effects of Treatment. Qualification of the FACE-Q | Aesthetics© MDDT included the following eleven scales determined to be most relevant to medical device regulatory decision making:

Qualified Self-Perceived Facial Appearance Scales:

1. Satisfaction with Cheeks
2. Satisfaction with Facial Appearance
3. Lines Between Eyebrows
4. Appraisal of Lines: Nasolabial Folds
5. Satisfaction with Lips
6. Satisfaction with Skin

Qualified Health-Related Quality of Life Scales:

1. Aging Appraisal
2. Early Life Impact of Treatment
3. Satisfaction with Outcome
4. Psychological Function
5. Social Function

Scales are transformed into scores ranging from 0 to 100 and can be used independently. Higher scores indicate a better outcome for all scales. Sufficient data are not available currently to make a determination on the change in score or difference in group scores that would indicate a clinically meaningful improvement or difference between groups.

QUALIFIED CONTEXT OF USE

The paper and electronic self-administered versions of the FACE-Q | Aesthetics© Satisfaction with Cheeks, Satisfaction with Facial Appearance, Lines Between Eyebrows, Appraisal of Lines: Nasolabial Folds, Satisfaction with Lips, Satisfaction with Skin, Aging Appraisal, Early Life Impact of Treatment, Satisfaction with Outcome, Psychological Function, and Social Function are used to quantify a subject's quality of life and satisfaction with facial aesthetic procedure(s).

Table 1 below outlines the specific qualified contexts of use for each of the 11 scales. The scales may be used as a co-primary, secondary, or ancillary endpoint in feasibility, pivotal, and post-approval studies to support the effectiveness of the device for proposed indications where the scale can meaningfully assess the clinical outcome.

The FACE-Q | Aesthetics© MDDT is not intended to be used as the sole primary endpoint. The eleven scales that have been qualified in this MDDT do not assess for safety or adverse events.

These scales may be used as a co-primary endpoint or composite endpoint with other clinically meaningful outcomes in studies where the benefits of the device cannot be directly measured by the FACE-Q | Aesthetics© scales alone. In addition, scales may be used to characterize changes from baseline when that is relevant to the study evaluation.

TABLE 1. QUALIFIED CONTEXT OF USE FOR EACH SCALE

Scale	Qualified Context of Use
1. Satisfaction with Cheeks	• Secondary endpoint •
2. Satisfaction with Facial Appearance 3. Lines Between Eyebrows 4. Appraisal of Lines: Nasolabial Folds 5. Satisfaction with Skin 6. Aging Appraisal 7. Early Life Impact of Treatment	• Co-primary or secondary endpoint

8. Satisfaction with Lips	• Co-primary or secondary endpoint •
9. Satisfaction with Outcome	• Secondary endpoint
10. Psychological Function 11. Social Function	• Ancillary endpoint

SUMMARY OF EVIDENCE TO SUPPORT QUALIFICATION

Extensive published literature, qualitative research and testing, and quantitative testing were submitted to support the qualification of the MDDT for the qualified context of use. This included published literature for the development of these scales as well as published literature from other parties who have used these scales in practice.

The scientific evidence provided in the qualification package demonstrated evidence of content validity, construct validity and reliability through the development and testing of this tool.

The tool was developed in three stages. During the Conceptual Framework Formation stage, interviews were performed with a cohort of patients who had surgical or minimally invasive procedures performed to the face. Words and phrases from these interviews were categorized into domains and themes and subthemes, leading to a framework that covered appearance, quality of life, and adverse effects.

During the Item Generation, Preliminary Scale Formation, and Pretesting phase, items from the first phase were used to form draft scales, and these were pretested with a small cohort of patients for feedback and were also shown to clinical experts for feedback.

In the third phase of Field Testing, Scale Construction, and Psychometric Evaluation, the FACE-Q | Aesthetics© field-test was performed in multiple sites for various procedures. These results were published in a series of 9 papers. Psychometric analyses were performed using Rasch measurement theory (RMT) analyses. This allowed the data to then be used to measure a concept, and also measure change. The RMT analyses supported summing items to form a total score for appearance and quality of life scales. The psychometric analyses demonstrated acceptability, reliability and validity.

DISCUSSION OF THE EVIDENCE STRENGTH TO SUPPORT QUALIFICATION

FACE-Q | Aesthetics© has a history of use in clinical trials evaluating device-based aesthetic treatment of various regions of the face to inform regulatory decisions. This experience was considered during the review in addition to the data submitted in the Qualification Package. The developer also submitted peer-reviewed publications that demonstrate evidence of change, including some publications that evaluated populations that are different from the ones in which these scales were developed; this demonstrated the use of these scales in more diverse populations in terms of aesthetic procedure, ethnicity, and gender. The multiple sources and types of evidence submitted provide confidence in the accuracy and meaning of the scores and ability of these scales to detect change.

All of the qualified appearance and quality of life scales were supported by the same basic evidence. The evidence shows that the items and response options function as they should, and the score produced by each scale is adequately reliable. This baseline information supports their use in a clinical study. Many of the scales also have evidence of responsiveness, or the ability to detect change. This helps support their use and interpretation as endpoints. While there is evidence of responsiveness and evidence that is suggestive of a meaningful difference estimate, sufficient data are not available currently to make a determination on clinically meaningful difference estimates.

Most of the scales have sufficient supportive evidence for use as a co-primary endpoint or secondary endpoint in a clinical study. As discussed, however, it is unknown what constitutes a clinically meaningful change, which may affect the ability to determine success criteria. An additional limitation is the potential use of the scales in regulatory decision-making encompasses many types of device-based procedures and patient populations. However, the data on responsiveness is derived from a limited number of procedures and patient demographics. For example, the participants in the validation studies were primarily self-identified as women (over 85%), Caucasian (67-100% depending on the study), and underwent surgical procedures, such as facelift procedures. A summary of the demographics for the field studies to support each scale is presented in Table 2 below. The validity of these scales for devices or patient populations that have not been studied is not known, especially if the effect of the proposed device or device-based procedure could reasonably be thought to differ from a surgical procedure such as facelift procedure. Sponsors should engage with FDA to determine the if the MDDT is being used within the qualified context of use to support their proposed clinical study design and proposed indications for use.

TABLE 2. FIELD STUDY DEMOGRAPHICS FOR EACH QUALIFIED SCALE

Scale	Field Study Demographics	Recommendation
1. Satisfaction with Facial Appearance	• Study 1: 323 (86%) women and 54 (14%) men; 67% white non-Hispanic; variety of procedures • Study 2: 86 (86%) women and 10 (14%) 100% white non-Hispanic; 100% facelifts.	Co-primary or Secondary
2. Aging Appraisal	• Age 28-89, 251 (87%) women, 30 men (10%), and 7 not reported • 87% were white non-Hispanic • Procedures: 217 facelift, 67 blepharoplasty, and 4 other	Co-primary or Secondary
3. Satisfaction with Cheeks 4. Appraisal with Lines: Nasolabial Folds	• Age range 36-77; 205 (92%) women and 18 (8%) men; 210 (93%) subjects are Caucasian • 225/225 subjects underwent facelift	Secondary (cheeks) Co-primary or Secondary (NLF)
5. Satisfaction with Outcome 6. Psychological Function 7. Social Function 8. Early Life Impact of Treatment	• Three studies: Ages 18-89; total of 621 (90%) women, 68 men (10%) • Ethnicities were not reported • Diverse range of procedures	Secondary (Outcome) Co-primary or Secondary (Early Life Impact) Ancillary (Psychological, Social)
9. Satisfaction with Skin 10. Satisfaction with Lips 11. Lines: Between the Eyebrows	• Ages 18-80; 429 (85%) women, 46 (9%) men, 28 not reported • 71% were white non-Hispanic • Diverse mix of procedures	Co-primary or Secondary

ASSESSMENT OF ADVANTAGES/DISADVANTAGES OF QUALIFICATION

Assessments of Advantages of Using the MDDT

Aesthetic surgical and minimally invasive procedures are increasingly common in the United States. Along with this increasing demand, patient-reported outcomes instruments are necessary to understand the patient perspective and also incorporate

the patient perspective into clinical study design.

The FACE-Q | Aesthetics© developers published a systematic review of PRO instruments for facial aesthetic surgery and/or nonsurgical facial rejuvenation. From their review, they found no PRO instruments that satisfied best practice guidelines for PRO instrument development and validation. The developers concluded that valid, reliable, and responsive PRO instruments for surgical and nonsurgical facial rejuvenation were lacking.

FACE-Q | Aesthetics© demonstrates several advantages, including the following:

- Covers various anatomic locations independently in order to assess commonly targeted areas by facial aesthetic devices
- Provides reproducible methods that can be used by Sponsors and the Agency to measure the impact of facial aesthetic devices on a patient with respect to facial appearance satisfaction and health-related quality of life
- Allows for assessments of patient perspective

The Agency has had experience with MDDT tools in the past. Sponsors should engage with the Agency to determine if a FACE-Q | Aesthetics© scale is being used within the qualified context of use to support their proposed study design and proposed indications for use.

Assessments of Disadvantages of Using the MDDT

The following disadvantages of using the MDDT were identified:

1) inability to measure all important outcomes relevant to facial aesthetic device use,
2) development in a population that may not reflect all patient or procedure populations, and
3) insufficient evidence to determine a clinically meaningful difference estimate.

The inability to measure all important outcomes relevant to facial aesthetics devices can be mitigated through the MDDT's use as a co-primary, secondary, or ancillary endpoint depending on the proposed indication and the clinical meaningfulness of the scale used.

Although the developer demonstrated evidence of responsiveness to change in populations that differed from those used in the initial validation, additional studies are still needed to understand the functioning of the MDDT in diverse patient and procedure populations. For several of the scales, the population studied was not representative of the United States population in terms of gender, ethnicity, skin-type, or procedure (see

Table 2).

For example, for the Satisfaction with Cheeks and Appraisal of Lines: Nasolabial Folds scales, the study population was 91.9% female, 93.3% Caucasian, and 100.0% facelift. This may question the ability of these scales to function in all of the populations that may seek device-based facial aesthetic procedures, such as men, other ethnicities or skin types, or aesthetic procedures aside from facelift.

Additional Factors for Assessing Advantages and Disadvantages of Using the MDDT

In addition to the demographics used to develop and study the scales, several of the scales have items that may affect the use of the scale. For example, scales with questions assessing for symmetry may not be appropriate for split-face studies. In addition, some of the scales have large ceiling effects, where more than 10 to 15% of respondents achieve the highest possible score. This may detract from the scale's ability to assess change and should be considered in determining the applicability of the scale for a clinical study.

CONCLUSIONS

The submitted qualification materials, including numerous published clinical studies, support the validity and reliability of the qualified FACE-Q | Aesthetics© scales within the specified contexts of use. The proposed scales listed above are generally acceptable. These FACE-Q scales may be used as a co-primary, secondary, or ancillary endpoint depending on the proposed indication and the clinical context.

REFERENCES

1. Klassen AF, Cano SJ, Pusic AL. FACE-Q Satisfaction with Appearance Scores from Close to 1000 Facial Aesthetic Patients. Plastic and Reconstructive Surgery 2016;137:651e-2e.

2. Klassen AF, Cano SJ, Schwitzer J, Scott A, Pusic AL. FACE-Q scales for health-related quality of life, early life impact and satisfaction with outcomes and decision to have treatment: development and validation. Plast Reconstr Surg. 2015; 135(2):375-86.

3. Klassen AF, Cano SJ, Schwitzer JA, Baker SB, Carruthers A, Carruthers J, Chapas A, Pusic AL. Development and Psychometric Validation of the FACE-Q Skin, Lips, and Facial Rhytides Appearance Scales and Adverse Effects Checklists for Cosmetic Procedures. JAMA Dermatol 2016; 152(4): 443-451.

4. Klassen AF, Cano SJ, Schwitzer JA, et al. Development and Psychometric

Validation of the FACE-Q Skin, Lips, and Facial Rhytids Appearance Scales and Adverse Effects Checklists for Cosmetic Procedures. JAMA Dermatology 2016;152:443-51.

5. Klassen AF, Cano SJ, Scott A, Snell L, Pusic AL. Measuring patient-reported outcomes in facial aesthetic patients: development of the FACE-Q. Facial Plastic Surgery 2010;26:303.

6. Klassen AF, Cano SJ, Scott AM, Pusic AL. Measuring Outcomes That Matter to Face-Lift Patients: Development and Validation of FACE-Q Appearance Appraisal Scales and Adverse Effects Checklist for the Lower Face and Neck. Plast Reconstr Surg. 2014; 133(1):21-30.

7. Panchapakesan V, Klassen AF, Cano SJ, Scott AM, Pusic AL. Development and Psychometric Evaluation of the FACE-Q Aging Appraisal Scale and Patient-Perceived Age Visual Analog Scale. Aesth Surg J. 2013; 33(8):1099-1109.

8. Pusic AL, Klassen AF, Scott AM, Cano SJ. Development and psychometric evaluation of the FACE-Q satisfaction with appearance scale: a new patient-reported outcome instrument for facial aesthetics patients. Clinics in plastic surgery 2013;40:249-60.

CONTACT INFORMATION FOR ACCESS TO TOOL

More information about FACE-Q | Aesthetics© is available at www.qportfolio.org. In addition, inquiries can be sent to qportfolioteam@gmail.com

The tool developers may be contacted by the following:

Anne Klassen, DPhil (Oxon)
McMaster University, 3N27, 1280 Main St W, Hamilton, ON, Canada, L8S 4L8, aklass@mcmaster.ca

Andrea L Pusic MD MSc, Brigham and Women's Hospital, 75 Francis S, Boston, MA, USA, 02116, apusic@bwh.harvard.edu

Stefan J Cano PhD, Modus Outcomes LLP, 39 Moffat Road, Newton, MA 02468, USA, stefan.cano@modusoutcomes.com