Development and Psychometric Validation of Patient-Reported Outcome Measures for Masseter Muscle Prominence: The Lower Facial Shape Questionnaire

Catherine Foley, MA, MPH[1]; Julia Garcia, PhD[1]; Julie Whyte, MPH[2]; Martha Gauthier, MA[2]; Brandon Foster, PhD[2]; Elisabeth Lee, MBA, MPH[1]; Vaishali Patel, PharmD[1]

[1] AbbVie, Irvine, CA, USA; [2] Lumanity, Boston, MA, USA

OBJECTIVE

To evaluate the:

- Content validity of the Lower Facial Shape Questionnaire (LFSQ)–Treatment Satisfaction Assessment, a 4th module of the masseter muscle prominence (MMP)–specific LFSQ patient-reported outcome (PRO) measure
- Psychometric validation of all 4 LFSQ modules:
  - Impact Assessment (LFSQ-IA)
  - Signs Assessment (LFSQ-SA)
  - Satisfaction Assessment (LFSQ-SAT)
  - Treatment Satisfaction Assessment (LFSQ-TXSAT)

CONCLUSIONS

The content validity of the LFSQ-TXSAT was confirmed through cognitive debriefing interviews with patients with MMP

All psychometric property criteria were met or exceeded for all LFSQ modules

LFSQ assessments can be used to capture the perspectives of patients with MMP as well as guide patient-physician dialogue in clinical practice

For additional information or to obtain a PDF of this poster

Scan QR to download an electronic version of this presentation and other Allergan Aesthetics and AbbVie 2025 ISPOR Montreal scientific presentations:

QR code expiration: 16 April 2026

To submit a medical question, please visit www.abbviemedinfo.com

Allergan Aesthetics, an AbbVie company, and the authors thank all the trial investigators and the patients who participated in these studies.

Allergan Aesthetics, an AbbVie company, funded this study and participated in the study design, research, analysis, data collection, interpretation of data, reviewing, and approval of the publication. All authors had access to relevant data and participated in the drafting, review, and approval of this publication. No honoraria or payments were made for authorship.

Medical writing was provided by Gabrielle Dardis, PhD, and Matt Fitzpatrick, PhD, of RTI Health Solutions, an independent nonprofit research organization that received funding from AbbVie to provide publication support in the form of writing, styling, and submission.

Financial arrangements of the authors with companies whose products may be related to the present report are listed as declared by the authors: Catherine Foley, Julia Garcia, Elisabeth Lee, and Vaishali Patel are full-time employees of AbbVie. Julie Whyte, Martha Gauthier, and Brandon Foster are full-time employees of Lumanity.

References

1. Almukhtar RM, Fabi SG. Plast Reconstr Surg. 2019;143(1):39e-48e.
2. Tabrizi R, et al. J Craniofac Surg. 2010;21(4):1096-7.
3. Crawford B, et al. Value Health. 2018;21:S145.
4. Fabi S, et al. Development of patient-reported outcome measures for masseter prominence [abstract]. Presented at ISPOR; 19-23 May 2018. Baltimore, MD.
5. Food and Drug Administration. April 2023. Patient-focused drug development: incorporating clinical outcome assessments into endpoints for regulatory decision-making.
6. Beatty PC, Willis GB. Public Opin Q. 2007;71(2):287-311.

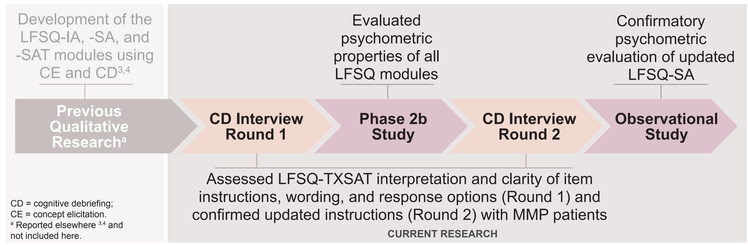
Presented at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 13-16 May 2025

Allergan Aesthetics
an AbbVie company

INTRODUCTION

- MMP is characterized by a unilateral or bilateral enlargement of the masseter muscles that presents as a widened lower face or square jaw[1,2]
- MMP can be of aesthetic concern and is associated with negative psychological and emotional impacts, such as feeling self-conscious and less attractive[2,3]
- PRO measures, including the LFSQ, have been developed to comprehensively understand MMP treatment benefit from the patient perspective
- Key outcomes of voluntary aesthetic treatment are assessed across 4 modules: LFSQ-IA, -SA, -SAT, and -TXSAT

Figure 1. LFSQ Module Development and Psychometric Validation Timeline



METHODS

- LFSQ modules were qualitatively developed and psychometric properties were evaluated across several studies (Figure 1) using established qualitative and quantitative research methods and US Food and Drug Administration guidance on patient-focused drug development[5]
- Across studies, key exclusion criteria included those with facial nerve abnormalities and those with excess lower face fat, loose skin, or parotid gland hypertrophy, as assessed by the investigator or a healthcare provider

The LFSQ-TXSAT and the previously developed LFSQ-IA/-SA/-SAT modules[3,4] were qualitatively developed and evaluated in study populations of adults aged ≥18 years residing in the United States with bilaterally symmetrical MMP

Content validity of LFSQ-TXSAT was assessed across 2 rounds of CD interviews. Eligible LFSQ-TXSAT CD interview participants were required to be "*somewhat*," "*a lot*," or "*extremely*" bothered by MMP

- Participants were encouraged to verbalize their thoughts using a think-aloud process during audio-recorded interviews[6]

Psychometric properties of LFSQ modules were evaluated at 2 timepoints during phase 2b study NCT03861936 and, for LFSQ-SA, confirmed at 2 post-screening visits 14 days apart during an observational study

- Multi-item LFSQ modules were also assessed for structural validity

RESULTS

LFSQ Modules (Table 1)

- LFSQ-IA/-SA/-SAT were previously developed through qualitative interviews with patients with MMP[3,4] to assess psychosocial impacts, signs of MMP, and condition-related satisfaction
- LFSQ-TXSAT was developed to assess treatment expectations and satisfaction with a single item measured at baseline and follow-up, respectively

Table 1. LFSQ Modules

	LFSQ-IA	LFSQ-SA[a]	LFSQ-SAT	LFSQ-TXSAT
General concept measured	Psychosocial impacts	Signs of MMP	Condition-related satisfaction	Treatment satisfaction
Items	6 items	3 items	7 items	1 item each (Baseline and Follow-Up)
Scoring	Summary score range, 0-24; higher scores indicate greater psychosocial impact	Summary score range, 0-12; higher scores indicate more severe signs of MMP	Appearance domain summary score range, -8 to 8 and Psychosocial domain summary score range, -6 to 6; higher scores indicate a higher level of satisfaction[b]	Baseline range, 0 (not at all) to 4 (extremely); Follow-Up range, -2 (very dissatisfied) to +2 (very satisfied)

Note: Updates to LFSQ-IA/-SA/-SAT instructions following LFSQ-TXSAT CD interviews were made prior to inclusion in the Phase 2b study and psychometric evaluation.
[a] Signs Assessment previously called Symptom Assessment.[3,4]
[b] Appearance domain: Items 1-4 (Satisfaction with appearance); Pyschosocial domain: Items 5-7 (Satisfaction with psychosocial impact).

LFSQ-TXSAT CD Interviews

Instructions and Items

- Interpreted as intended by 100.0% (Table 2) and reported as clear by most (85.7%-100.0%) participants in Round 1
- For LFSQ-TXSAT Follow-Up:
  - Item reported as unclear by 2 participants (13.3%); 1 reported considering additional aspects of treatment beyond the scope of the measure, including "*cost, pain, downtime, and results*"
  - Instructions and items were modified to clarify that only the appearance of the lower face should be considered
- In Round 2, all participants with evaluable responses interpreted the updated instructions as intended

Response Options

- Interpreted as intended by most participants (Baseline, 93.3%; Follow-Up, 71.4%-85.7%)
- For LFSQ-TXSAT Follow-Up:
  - 2 participants reported considering factors beyond the appearance of their lower face
  - Revisions made to the instructions addressed these issues

Table 2. LFSQ-TXSAT Instructions and Items Interpretation

	Baseline (n = 15)	Follow-Up (n = 15)
Instructions	✓	✓
Item	✓	✓

✓: No issue with interpretation

Participants (Table 3)

- Across studies, mean participant age range was 39.3-51.4 years and most were female, White, and had at least *Moderate* investigator-rated MMP severity

Table 3. Baseline Characteristics of Study Participants

	CD Interviews		Psychometric Evaluation	
Characteristic	Round 1 (n = 15)	Round 2 (n = 20)	Phase 2b study[a] (n = 145)	Observational study[a] (n = 120)
Age, mean (SD), years	42.4 (13.3)	51.4 (13.9)	39.3 (11.1)	41.2 (10.5)
Female, n (%)	12 (80.0)	16 (80.0)	130 (89.7)	89 (74.2)
White, n (%)	6 (40.0)	15 (75.0)	110 (75.9)	60 (50.0)
Investigator-rated MMP severity, n (%)				
Minimal	0 (0.0)	0 (0.0)	0 (0.0)	14 (11.7)
Mild	0 (0.0)	4 (20.0)	0 (0.0)	27 (22.5)
Moderate	4 (26.7)	6 (30.0)	0 (0.0)	35 (29.2)
Marked	8 (53.3)	7 (35.0)	89 (61.4)	26 (21.7)
Very marked	3 (20.0)	3 (15.0)	56 (38.6)	18 (15.0)

SD = standard deviation.
[a] Phase 2b and observational studies reported demographic characteristics for modified intent to treat populations at baseline; population characteristics may differ at follow-up timepoints or visits.

LFSQ Psychometric Property Evaluation

- All modules met or exceeded the acceptable criteria for all psychometric analyses (Table 4)
  - After phase 2b evaluation, 1 LFSQ-SA item (*looking uneven*) relating to self-perceived lower face symmetry was removed due to poor correlation with other items
  - In the observational study, the updated LFSQ-SA showed strong psychometric properties and improved test-retest reliability and inter-item correlations

Table 4. LFSQ Psychometric Validation Results Against Pre-established Criteria

	Phase 2b Psychometric Validation				Observational Study Evaluation
Property	LFSQ-IA	LFSQ-SA	LFSQ-SAT	LFSQ-TXSAT (Follow-Up)	LFSQ-SA
Response distributions	✓	✓	✓[a]	✓	✓
Internal consistency reliability	✓	✓	✓	n/a	✓
Inter-item correlations	✓	✓[b]	✓[c]	n/a	✓
Test-retest reliability	✓	✓	✓	✓	✓
Convergent validity	✓	✓	✓	✓	✓
Known-groups validity	✓	✓	✓	✓	n/a
Responsiveness	✓	✓	✓	n/a	n/a
Structural validity	✓	✓	✓	n/a	✓

✓: adequate demonstration of psychometric properties ✓: criteria are met but statistical considerations remain

ICC, intraclass correlation coefficient; n/a = not applicable.
Note: Pre-established acceptability criteria of psychometric properties: response distribution (no floor/ceiling effects); internal consistency reliability ($\alpha/\omega \geq 0.70$); inter-item correlations ($0.15 \leq r \leq 0.85$); test-retest reliability (ICC [A,1] ≥ 0.70, $R_s \geq 0.70$ [LFSQ-TXSAT Follow-Up only]); convergent validity ($|r_s| \geq 0.40$); known-groups validity (expected monotonic ordering of summary scores across strata of participants' responses); responsiveness ($|r_s| \geq 0.35$).
[a] Floor effects of *very dissatisfied* at Baseline (expected due to peak efficacy timepoint at Day 90).
[b] The correlation between item 3 (*looking uneven*) and the other 3 items was lower than the other correlations observed between items.
[c] Item 5 (*social*) showed a pattern of weaker correlations with the remaining items, which was different than the pattern observed for the other items.