

# Development and Validation of Facial Line Distress Scale-Glabellar Lines (FINE-GL)

Danbee Kang<sup>1,2,\*</sup>, Eunjee Kang<sup>2,\*</sup>, Kyeongrok Choi<sup>3</sup>, Sooyeon Kim<sup>1,2</sup>, Woo Shun Lee<sup>4</sup>, Juhee Cho<sup>2</sup>

<sup>1</sup>Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea; <sup>2</sup>Center for Clinical Epidemiology, Samsung Medical Center, Seoul, Republic of Korea; <sup>3</sup>Department of Medical Device Management and Research, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea; <sup>4</sup>Medytox Inc, Seoul, Republic of Korea

\*These authors contributed equally to this work

Correspondence: Juhee Cho, Center for Clinical Epidemiology, Samsung Medical Center, 115 Irwon-ro, Gangnam-gu, Seoul, 06351, Republic of Korea, Tel +82-2-3410-1448, Fax +82-2-3410-6639, Email jcho@skku.edu

**Purpose:** We developed and validated facial line distress scale-glabellar lines (FINE-GL) to evaluate the severity and psychosocial distress associated with GL.

**Patients and Methods:** In Phase I, a preliminary item pool for the FINE-GL was developed through a literature review and expert consultation. This was followed by cognitive interviews to ensure comprehensibility of the items. In Phase II, we conducted a cross-sectional survey at a tertiary hospital and two local clinics in Korea. Exploratory factor analysis (EFA) was conducted to identify the underlying factor structure of the FINE-GL, and internal consistency and test-retest reliability were also examined.

**Results:** We yielded 20 items in four domains. The model fit was good. Coefficient alphas ranged from 0.92 to 0.95 for sub-domains and 0.97 for the total. The FINE-GL was moderately correlated with the appearance appraisal score and body image. In the test-retest, the range of ICC was 0.77–0.90.

**Conclusion:** FINE-GL is a reliable, valid, and comprehensive patient-reported outcome measure for assessing GL severity and distress. This will be helpful to determine a patient's eligibility for inclusion of study and to measure primary or secondary effectiveness endpoints for glabellar line treatment.

**Keywords:** glabellar lines, psychosocial distress, questionnaire, validation, patient-reported outcome

## Introduction

Glabellar lines (GL) constitute one of the most visible signs of aging.<sup>1,2</sup> As the glabellar is located in the center of the face, it is an easy focal point of attention for a person and those around them. GL give the impression of being angry or sad, which can negatively affect social interactions.<sup>3</sup> According to a previous study of 945 patients with a minimum of three consecutive treatments in more than one area of the face, the glabellar was the most frequently treated area (93.9%).<sup>4</sup> It is one of the facial lines the general population prefers to undergo cosmetic procedures.<sup>5</sup>

Recently, minimally invasive aesthetic procedures, such as botulinum toxin, filler injections, and laser abrasion have gained attention in the prevention or treatment of GL. Both objective and subjective outcomes have been used to evaluate the effects of these treatments. Skin imitation<sup>6</sup> and three-dimensional (3D) imaging<sup>7,8</sup> are the most accurate and reliable quantitative methods of assessment. However, they require special equipment and professional manpower. As subjective measurements, photographic scales, such as the Facial Wrinkle Scale, Glabellar Line Scale, and Merz Aesthetic Scale (MAS), are commonly used.<sup>9</sup> However, they are based on a single question and only assess the severity of lines without assessing their consequent psychological distress. Furthermore, physicians' assessments of aesthetic outcomes differed from those of the patients.<sup>10</sup>

Recently, several patient-reported outcome measures (PROM) such as the Facial Line Satisfaction Questionnaire,<sup>11</sup> Facial Line Outcomes,<sup>12</sup> and FACE-Q have been developed to assess both the severity of facial lines and their

psychosocial impact.<sup>13,14</sup> While we developed a specific PROM, existing measures are designed for general facial lines, and they may not adequately capture the unique functional and aesthetic concerns specific to the glabellar area, such as its central role in facial expressions and its high visibility in social interactions. To fill this gap, we provided a condition-specific tool that evaluates both the severity and the psychosocial distress caused by GL. By addressing these unique aspects, the instrument enables clinicians to better identify patients in need of treatment, assess treatment outcomes, and consider the psychological dimensions of care, thereby improving patient-centered outcomes in clinical and research settings. Thus, we developed and validated “Facial line distress scale-glabellar lines (FINE-GL)” to evaluate the severity and psychosocial distress associated with glabellar lines.

## Methods

We developed the FINE-GL according to the CONsensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist<sup>15</sup> and the US FDA criteria for PROM.<sup>16</sup> The development of the facial line distress scale for GL process included two phases. Phase I is the instrument development phase, which includes item development and content validation. Phase II involved item reduction and psychometric validation. To develop this tool, we established an expert group of two clinicians and four behavior scientists.

Initially, we planned to develop a tool to assess the severity of and distress due to 3 upper facial lines including GL, lateral canthal line (LCL), and forehead lines (FL). However, during Phase I, we found that the tool needs to be specific for each type of line to accurately evaluate the concerned severity and distress. Subsequently, the research team decided to develop a separate tool for GL in view of its utility and feasibility.

### Phase I: Instrument Development Item Development

To develop items, we conducted an extensive literature review and semi-structured in-depth interviews with 25 adults (>18 years) who were concerned about upper facial lines. The participants were asked how much they were aware of each upper facial line and how these affected their daily living and social activities. We found that people experienced different levels of distress due to upper facial lines. People were concerned about these lines when they had facial expressions, concentrate, or look in the mirror. Some people with severe lines had a negative self-image and felt less confident about their appearance. Sometimes, people avoid participating in social activities due to concerns of GL or FL. People were less stressed about LCL than about GL or FL.

Based on the literature review and qualitative study, experts including 2 clinicians, 2 nurses and 3 behavior scientists reviewed the items for relevance, resulting in a preliminary pool of 54 items in six domains (general appraisal of line,  $n = 8$ ; aging appearance appraisal,  $n = 2$ ; appraisal of line in certain situations  $n = 11$ ; public self-consciousness,  $n = 9$ ; psychological distress,  $n = 8$ ; and social distress,  $n = 16$ ) were included in the initial version of the Facial Line Distress Scale (iFINE). All questions were answered based on the present time point on a 5-point Likert scale (1 = Not at all, 2 = A little, 3 = Somewhat, 4 = Quite a bit, and 5 = Very much).

### Content Validation

To validate the content of the iFINE, we conducted cognitive interviews. We recruited participants until saturation; 15 participants with and without experience with aesthetic procedures for any facial lines participated in the cognitive interviews; 26.7% of the patients were men and 46.7% were over 50 years old. Participants underwent 30–60-minute interviews for evaluation of comprehension, ease of responses, and acceptability of the terminology, phrasing, and response options.

The cognitive debriefing revealed that the participants generally comprehended the iFINE well. However, participants thought of different facial lines and gave different answers to general questions. In other words, unless the question did not specify a type of upper facial line, participants answered the questions based on evaluation of certain facial lines that they perceived to be most stressful. Therefore, we developed a glabellar line-specific PROM, the FINE-GL, based on qualitative interviews and feedback on glabellar line-specific concerns. Items such as “I look angry because of the glabellar lines” reflect common concerns identified during interviews. All participants reported that the questionnaire covered the aspects of

distress due to GL and understood all items clearly except one question related to aging appraisal. After discussing with experts, we decided to exclude the item due to its ambiguity, thus resulting in 36 items in the FINE-GL draft.

## Phase 2: Item Reduction and Psychometric Validation of FINE-GL

### Study Participants

We conducted a cross-sectional survey at a tertiary hospital and two local clinics in Korea between April 18 and June 17, 2022. The eligibility criteria were 1) adults aged 18 years or older; 2) who were concerned or distressed about GL; and 3) who could read and write Korean. We excluded participants who had severe scars on the upper face or mental disorders. The sample size was determined using guidelines for psychometric validation, with a target of 5 participants per item.<sup>17</sup> Since the number of draft items was 38, it might need to have 180 participants, and considering 5% missing values, we need to have 190 participants. The first 80 participants who completed the FINE-GL to re-take it to provide measures to assess its test–retest reliability. The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Samsung Medical Center, Seoul, Republic of Korea (SMC-2021-07-166). Informed consent was obtained from the study participants. We provided the participants with a \$10 as an incentive for their participation.

### Item Reduction

To reduce the number of items and to extract factor structure, we performed exploratory factor analysis (EFA) using varimax rotation. We used the maximum likelihood methods and checked for Tucker Lewis Index (TLW) for factoring reliability  $>0.9$  to confirm the adequacy of the number of factors. For an efficient scale with improved item information, low-discriminatory items were removed. The item response theory (IRT)-graded response model (GRM) was used in the distress domain. Measures of between-item relationships, unidimensionality, and monotonicity were estimated before applying the 2-PL logistic IRT model. To confirm monotonicity, we used the S-X2 statistic with a value  $<0.001$ , indicating a poor fit. To confirm the informative item, we calculated the discriminatory and informative scores. Discriminatory values were computed for item assessment. Information score  $>1.7$  is generally considered informative.<sup>18</sup> Differential item functioning (DIF) analyses were performed to evaluate whether persons from different groups, given similar levels of distress due to GL, have different probabilities of providing a response after conditioning on the level of the state measured.

After item reduction, no follow-up cognitive debriefing was conducted as the retained items had already been validated for comprehension and appropriateness. Confirmatory factor analysis (CFA) was performed with the retained items using the maximum likelihood without missing values method.<sup>19</sup> CFI indicates Comparative fit index and SRMR indicates standardized root mean squared residual.  $CFI > 0.9$  and  $SRMR < 0.08$  indicated a good fit.<sup>20,21</sup>

### Psychometric Validation

#### Measurement

To evaluate the convergent and discriminant validity of the FINE-GL, PROMs for health-related quality of life (HRQoL), body image, and appearance appraisal were included. For HRQoL, the World Health Organization-Quality of Life assessment instrument (WHOQoL)-BREF<sup>22</sup> was included. The body image scale (BIS) was used to measure the body image.<sup>23</sup> The line severity was measured using the MAS.<sup>24</sup> The scores of these PROMs were calculated according to scoring manuals.<sup>22–24</sup> Additional sociodemographic and clinical information, including age, sex, marital status, education, monthly family income, employment, and GL treatment experience was noted.

### Statistical Analysis

We calculated the internal consistency of FINE-GL using Cronbach's  $\alpha$ . Cronbach's  $\alpha$ , a widely used statistical measure that evaluates how well a set of items in a questionnaire measure a single underlying construct. Higher values of Cronbach's  $\alpha$  indicate better internal consistency, with an  $\alpha$ -value  $\geq 0.8$  commonly interpreted as very good reliability.<sup>25</sup>

The test–retest reliability of the FINE-GL was measured using the intra-class correlation coefficient (ICC) using a two-way mixed model. A questionnaire is considered reliable at ICC values  $>0.70$ .<sup>26</sup> Using the repeated measure data

set, the standard error of measurement (SEM) was also calculated by first creating a variable for the difference between the score obtained during the first and second administrations.<sup>27</sup> Subsequently, we calculated the smallest detectable change (SDC). The SDC can be calculated as  $SEM \times 1.96 \times \sqrt{2}$ .<sup>28</sup>

Convergent and discriminant validity were tested using Pearson’s correlations between the FINE-GL and the WHOQoL-BREF, BIS, and MAS. Our hypothesis was that FINE-GL to demonstrate negative correlations with WHOQoL-BREF ( $-0.70 \leq r \leq -0.30$ ) and positive correlations with BIS and MAS in convergent validity ( $0.30 \leq r \leq 0.70$ ). To adjust for multiple comparisons, we estimated the adjusted p-values using the Holm–Bonferroni method.

To test criterion validity, we calculated the area under the curve (AUC) of the FINE-GL for severe GL according to physicians (severe glabellar line and “very severe glabellar line”). We also used the Youden index to determine the cut-off values.<sup>29</sup>

The significance level was set at  $p < 0.05$  (two-sided). Statistical analyses were performed using STATA v16 (StataCorp LP, College Station, TX, USA) and R 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

# Results

## Study Participants

Of 198 patients, we excluded six with missing data on at least one item of FINE-GL. The mean age (SD) of participants was 47.68 (13.71) years; 81.2% were women, and 30.7% had experience with facial aesthetic treatments (Table 1). While more than 22% of the participants were aged >60 years, 97% completed all the questions without problems.

## Item Reduction

All 36 semi-final items (Table 2) were satisfied. EFA indicated a five-factor solution with an eigenvalue >1.0, while it was initially designed for a six-factor solution (Table 2). For individual items, one item loaded significantly on the interpretable factor solution. Three items (#19, #34, #36) that loaded on  $\geq$  two-factor solutions with low loading values ( $r < 0.5$ ) were excluded. #14 was loaded on the “public self-consciousness” domain with a relatively low factor-loading; however,

**Table 1** Characteristics of Study Participants (N = 192)

| Characteristics  |               |
|--|---------------|
| <b>Age (years)</b>                                     | 47.68 ± 13.71 |
| <b>Sex, female</b>                                     | 156 (81.2)    |
| <b>Marital status</b>                                  |               |
| Single   | 49 (25.5)     |
| Married  | 129 (67.2)    |
| Separated/Divorced/Widowed                             | 13 (6.8)      |
| Unknown  | 1 (0.5)       |
| <b>Education</b>                                       |               |
| < High school  | 61 (31.8)     |
| $\geq$ University                                      | 130 (67.7)    |
| <b>Monthly family income, <math>\geq</math> \$6000</b> | 39 (20.3)     |
| <b>Working status at diagnosis</b>                     |               |
| No work/Housewife                                      | 24 (12.5)     |
| White collar   | 68 (35.4)     |
| Blue collar  | 23 (12.0)     |
| Service  | 43 (22.4)     |
| Others   | 34 (17.7)     |
| <b>Facial aesthetic treatments, yes</b>                | 59 (30.7)     |
| Botulinum toxin, yes (N = 59)                          | 47 (79.7)     |
| Glabellar lines, yes (N = 59)                          | 38 (64.4)     |

**Notes:** Values are presented as n (%) or mean ± standard deviation. Monthly family income values is US Dollars (USD).

**Table 2** Factor Loadings from the Exploratory Factor Analysis and Reliability of FINE-GL Dimensions (N = 192)

| No  | Items (Theme)   | Factor |      |      |      |      | Decision              |
|-----|---|--------|------|------|------|------|-----------------------|
|     |   | 1      | 2    | 3    | 4    | 5    |                       |
|     | <b>Appraisal of GL</b>  |        |      |      |      |      |                       |
|     | <b>General</b>  |        |      |      |      |      |                       |
| #1  | I have a lot of glabellar lines   |        | 0.43 |      |      | 0.67 | General               |
| #2  | Glabellar lines stand out   |        |      |      |      | 0.67 | General               |
| #3  | My glabellar lines are deep.  |        | 0.41 |      |      | 0.75 | General               |
| #4  | The length of the glabellar lines is long.  |        | 0.43 |      |      | 0.69 | General               |
|     | <b>Facial expression</b>  |        |      |      |      |      |                       |
| #5  | I have a lot of glabellar lines   |        | 0.79 |      |      |      | Facial expression     |
| #6  | Glabellar lines stand out   |        | 0.84 |      |      |      | Facial expression     |
| #7  | My glabellar lines are deep.  |        | 0.84 |      |      |      | Facial expression     |
| #8  | The length of the glabellar lines is long.  |        | 0.70 |      |      | 0.41 | Facial expression     |
|     | <b>Certain situation</b>  |        |      |      |      |      |                       |
| #9  | I am concerned about the glabellar lines when talking.                              |        |      | 0.61 |      |      | Certain situation     |
| #10 | I am concerned about the glabellar lines under bright light.                        | 0.42   |      | 0.62 |      |      | Certain situation     |
| #11 | I am concerned about the glabellar lines when looking the mirror.                   |        |      | 0.63 |      |      | Certain situation     |
| #12 | I am concerned about the glabellar lines when concentrating.                        |        |      | 0.78 |      |      | Certain situation     |
| #13 | I am concerned about the glabellar lines when frowning my face.                     |        |      | 0.88 |      |      | Deleted <sup>†</sup>  |
|     | <b>Distress due to GL</b>   |        |      |      |      |      |                       |
|     | <b>Public self-consciousness</b>  |        |      |      |      |      |                       |
| #14 | I look older than my peers because of the glabellar lines                           | 0.49   |      |      | 0.48 |      | Self-consciousness    |
| #15 | I look angry because of the glabellar lines   |        |      |      | 0.63 |      | Self-consciousness    |
| #16 | I look depressed because of the glabellar lines                                     |        |      |      | 0.69 |      | Deleted <sup>†</sup>  |
| #17 | I look worried because of the glabellar lines                                       |        |      |      | 0.71 |      | Self-consciousness    |
| #18 | I look like having a strong impression or character because of the glabellar lines. |        |      |      | 0.58 |      | Self-consciousness    |
| #19 | I look like I went through difficulties in the past due to the glabellar lines.     | 0.52   |      |      | 0.58 |      | Deleted               |
|     | <b>Psychological distress</b>   |        |      |      |      |      |                       |
| #20 | I have low self-confidence because of the glabellar lines                           | 0.70   |      |      |      |      | Psychosocial distress |
| #21 | I think negatively about myself because of the glabellar lines                      | 0.67   |      |      |      | 0.47 | Deleted               |
| #22 | I feel unattractive because of the glabellar lines                                  | 0.73   |      |      |      | 0.43 | Deleted*              |
| #23 | I am unsatisfied with myself because of the glabellar lines                         | 0.66   |      |      |      |      | Deleted*              |
| #24 | I feel bad about myself because of the glabellar lines                              | 0.71   |      |      |      |      | Deleted*              |

(Continued)

**Table 2** (Continued).

| No  | Items (Theme)   | Factor      |   |      |   |   | Decision                     |
|-----|---|-------------|---|------|---|---|------------------------------|
|     |   | 1           | 2 | 3    | 4 | 5 |                              |
|     | <b>Social distress</b>  |             |   |      |   |   |                              |
| #25 | I am concerned about meeting new people because of the glabellar lines                      | <b>0.77</b> |   |      |   |   | <b>Psychosocial distress</b> |
| #26 | I am concerned about people around me talking about the glabellar lines.                    | <b>0.73</b> |   |      |   |   | Deleted*                     |
| #27 | I am concerned that glabellar lines will affect my social life.                             | <b>0.84</b> |   |      |   |   | <b>Psychosocial distress</b> |
| #28 | I am concerned about going to meetings and important events because of the glabellar lines. | <b>0.83</b> |   |      |   |   | Deleted*                     |
| #29 | I am concerned about going out without makeup because of the glabellar lines.               | <b>0.79</b> |   |      |   |   | Deleted†                     |
| #30 | I am concerned that my social skills will be underrated because of the glabellar lines.     | <b>0.83</b> |   |      |   |   | Deleted*                     |
| #31 | I avoid sitting close to others because of the glabellar lines.                             | <b>0.84</b> |   |      |   |   | <b>Psychosocial distress</b> |
| #32 | I avoid being photographed by others because of the glabellar lines.                        | <b>0.77</b> |   |      |   |   | Deleted*                     |
| #33 | I spend a lot of time preparing to go out because of the glabellar lines.                   | <b>0.73</b> |   |      |   |   | Deleted*                     |
| #34 | I use cosmetics to improve glabellar lines.   | 0.49        |   |      |   |   | Deleted                      |
| #35 | I had a procedure to improve glabellar lines.   |             |   |      |   |   | Deleted                      |
| #36 | I do not make facial expressions that could make the glabellar lines.                       | 0.41        |   | 0.41 |   |   | Deleted                      |

**Notes:** Values > 0.4 are presented, and the highest loading value is highlighted in bold; †Deleted based on expert review. \*Deleted based on Item Response Theory.

**Abbreviation:** NA, Not applicable.

participants considered it important in the qualitative interview, and therefore, we decided to retain it (Table 2). Five items (#20–#24) that were initially designed for the “psychological distress” domain loaded on the “social distress” domain; therefore, we combined the “psychological distress” and “social distress” domains into the “psychosocial distress” domain.

## IRT for Distress Domain

Following IRT in the psychosocial domain, #33 which was a non-monotonic item, was excluded. In IRT-GRM, the probability values for the S-X2 statistics ranged from 0.007 to 0.268. Consequently, four items (#22, #23, #29, #30) with a poor fit (RMSEA > 0.6 and pS-X2 < 0.05) were excluded (Table 2 and Figure 1). We excluded two more items: #28 (McFadden’s pseudo R2 for sex = 0.038) and #24 (McFadden’s pseudo R2 for age = 0.03) based on DIF by sex and age, respectively.

The factor structure of the final 20 items in FINE-GL was evaluated using CFA, which revealed high loading (0.77–0.94). The fit indices for this model were good (CFI = 0.918, SRMR = 0.05) (Figure 1). The final version of FINE-GL included 20 items in two domains and five subscales (1. appearance appraisal: general, facial expression, and certain situations; and 2. distress due to GL: self-consciousness and psychosocial distress). FINE-GL score (0–4) was calculated by summing the responses for the items in each domain. Higher scores indicated higher levels of distress due to GL.

## Psychometric Validation

### Internal Consistency

The possible range of scores on FINE-GL was 0–100. The mean (SD) total score on FINE-GL was 47.88 (21.74). Floor and ceiling effects were 1.0% and 0.5%, respectively. Cronbach’s alpha coefficients of the five subscales ranged from 0.92 to 0.95 indicating satisfactory internal consistency (Table 3). Cronbach’s alpha coefficient of the total score was 0.97 (Table 3).



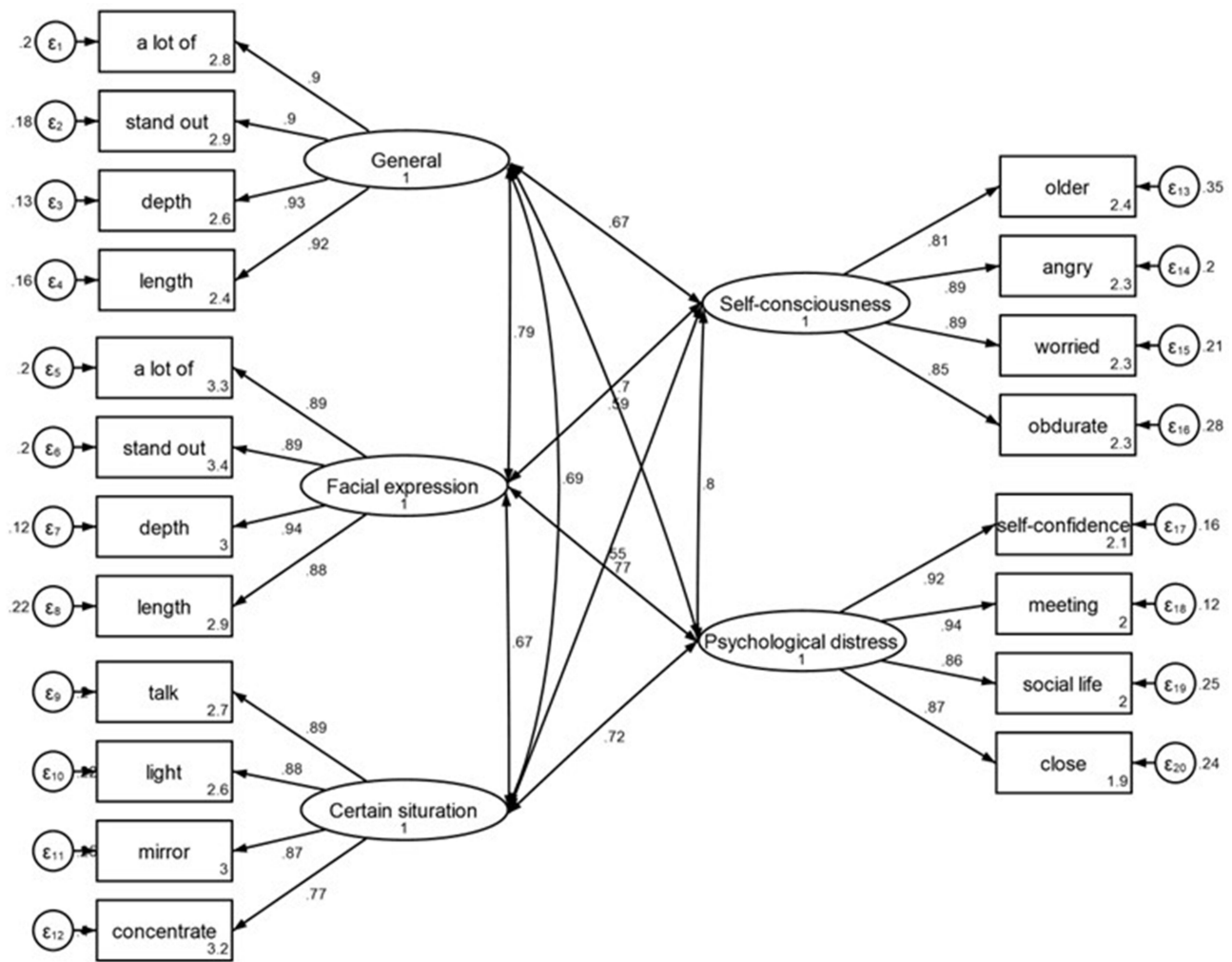


Figure 1 Confirmatory factor analysis.

### Test-Retest

The range of ICC was 0.77–0.90, which indicated satisfactory consistency. The SEM of the total score was 6.66, and the SDC score was 13 (Table 4).

### Hypothesis Testing for Construct Validity

The self-evaluation of GL at rest and dynamic on FINE-GL was moderately correlated with the MAS score at rest (rest,  $r = 0.79$ ; dynamic,  $r = 0.62$ ) and at dynamic (rest,  $r = 0.67$ ; dynamic,  $r = 0.73$ ). FINE-GL subscales of the distress domain were moderately correlated with body image. While WHOQoL was moderately correlated with the psychosocial distress subscale on FINE-GL, a weak correlation was noted with other subscales (Table 5).

### Criterion Validity

When the physician evaluated using MAS, 35.9% of the participants had severe GL. A difference of 16.6 points on FINE-GL was noted between participants with severe GL and those without, as evaluated by the physician (58.5 vs 41.9;  $P < 0.01$ ) (data not shown). The accuracy of FINE-GL score in predicting severe GL was characterized by an AUC of 0.73 (95% confidence interval = 0.66–0.80; Figure 2). The mean (SD) FINE-GL was 47.9 (21.7), and the cut-off value for severe GL was 50.6.

**Table 3** Description and Statistics for FINE-GL (N = 192)

| Subscales                 | Score Description |               |         |           | Cronbach Alpha |
|---------------------------|-------------------|---------------|---------|-----------|----------------|
|                           | Possible Range    | Mean (SD)     | % Floor | % Ceiling |                |
| Appraisal of GL           | 0–60              | 32.37 (13.36) | 1.0%    | 2.1%      | 0.98           |
| General                   | 0–20              | 9.69 (5.13)   | 8.3%    | 4.7%      | 0.95           |
| Facial expression         | 0–20              | 11.74 (4.96)  | 3.6%    | 9.4%      | 0.94           |
| Certain situation         | 0–20              | 10.94 (5.01)  | 4.2%    | 5.2%      | 0.92           |
| Distress due to GL        | 0–40              | 15.51 (9.70)  | 8.3%    | 1.0%      | 0.98           |
| Public self-consciousness | 0–20              | 9.06 (5.29)   | 9.9%    | 3.1%      | 0.92           |
| Psychosocial distress     | 0–20              | 6.46 (5.06)   | 18.2%   | 1.0%      | 0.93           |
| Total FINE-GL             | 0–100             | 47.88 (21.74) | 1.0%    | 0.5%      | 0.97           |

**Table 4** Reliability of FINE-GL (N = 80)

| Subscales                 | ICC (95% CI)      | SEM  | SDC  |
|---------------------------|-------------------|------|------|
| Appraisal of GL           | 0.88 (0.82, 0.91) | 4.34 | 8.5  |
| General                   | 0.88 (0.83, 0.91) | 1.76 | 3.45 |
| Facial expression         | 0.79 (0.71, 0.85) | 1.89 | 3.71 |
| Certain situation         | 0.77 (0.69, 0.84) | 2.19 | 4.29 |
| Distress due to GL        | 0.87 (0.81, 0.91) | 3.42 | 6.71 |
| Public self-consciousness | 0.83 (0.77, 0.88) | 2.14 | 4.2  |
| Psychosocial distress     | 0.84 (0.78, 0.89) | 1.86 | 3.64 |
| Total FINE-GL             | 0.90 (0.85, 0.93) | 6.66 | 13   |

Abbreviations: FINE-GL, facial line distress scale-glabellar lines; ICC, intra-class coefficient; SDC, smallest detectable change; SEM, standard error of measurement.

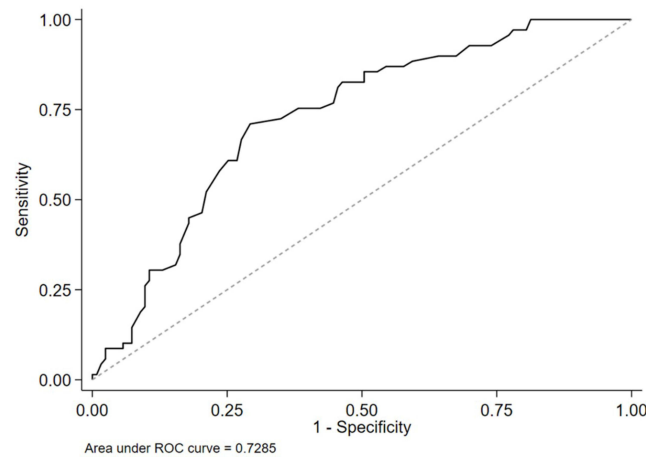
**Table 5** Correlation of FINE-GL with WHOQoL (N = 192)

| Subscales                 | Legacy Measures |         |             |                |
|---------------------------|-----------------|---------|-------------|----------------|
|                           | WHOQoL-BREF     | BIS     | MAS at Rest | MAS at Dynamic |
| Appraisal of GL           | −0.205*         | 0.333** | 0.728**     | 0.722**        |
| General                   | −0.233**        | 0.310** | 0.788**     | 0.671**        |
| Facial expression         | −0.168*         | 0.245** | 0.625**     | 0.726**        |
| Certain situation         | −0.147*         | 0.335** | 0.530**     | 0.534**        |
| Distress due to GL        | −0.305**        | 0.452** | 0.572**     | 0.590**        |
| Public self-consciousness | −0.230**        | 0.413** | 0.494**     | 0.572**        |
| Psychosocial distress     | −0.344**        | 0.435** | 0.581**     | 0.532**        |
| Total FINE-GL             | −0.262**        | 0.406** | 0.702**     | 0.706**        |

Notes: \*\*P < 0.01; \*P < 0.05.

Abbreviations: BIS, Body Image Scale; MAS, Merz Aesthetics Scale; WHOQoL, World Health Organization Quality of Life assessment instrument; FINE-GL, facial line distress scale-glabellar line.





**Figure 2** Criterion validity: receiver-operating characteristic (ROC) curves of FINE-GL for severe glabellar lines evaluated by physicians [69/192 (35.9%)].

## Discussion

In this study, we developed a condition-specific PROM in evaluating the severity and distress due to GL following the COSMIN and US FDA guidelines for PROM development and reported that FINE-GL is a reliable and valid PROM.

Using qualitative interviews, patients expressed specific concerns about GL, such as feeling they appeared angry or sad, which impacted their social interactions. Some avoided certain social settings due to the belief that GL negatively affected how others perceived them.<sup>2</sup> Additionally, we found that we cannot measure GL-specific distress if we use a general question for upper facial lines; thus, we developed FINE-GL to evaluate the severity and distress due to GL. FINE-GL includes two domains—appraisal of GL and distress due to GL. Appraisal of GL includes items to assess the severity of GL not only generally but also in situations where people experience concerns or distress due to GL, such as during facial expressions or in certain situations, such as talking, looking in the mirror, and concentrating. This reflects a real situation in which people are concerned about the GL associated with the high content validity of FINE-GL. Unlike general PROMs, the FINE-GL captures psychosocial distress specific to GL by including items that reflect GL-related concerns in real-life scenarios. This specificity makes it particularly useful for evaluating the impact of treatments targeting GL.

Distress due to GL includes a subscale called public self-consciousness. It includes items related to how a person with GL perceives themselves, which reflects the person's perspectives on their physical selves. According to the objectification theory, people are typically acculturated to internalizing an observer's perspective as the primary view of their physical selves, which can lead to habitual body monitoring and shame and anxiety.<sup>30</sup> The theory could explain the association between psychological and social distress caused by GL. Initially, psychological and social distress were separate subscales; however, after EFA, they were combined into one domain. According to previous studies, GL gives the impression of being angry or sad, which can negatively affect social interactions.<sup>3</sup> Participants in our study might also have considered the social and psychological domains as a single factor in relation to distress due to GL.<sup>31</sup> Furthermore, the internal consistency reliability of the measure was high, and some of the items were loaded in a domain different from our original hypothesis. "I look older than my peers because of the glabellar lines" was originally included in the self-consciousness domain, but it had similar loading values in self-consciousness and psychosocial distress. Since social norms and cues indicate that people's social value decreases with age, perceived aging could be associated with social interactions.<sup>32</sup> Thus, perceived aging might be associated with the psychosocial domain as well as self-consciousness.

With test–retest, we confirmed that FINE-GL is a reliable PROM. When assessing a participant, a change in score >13 points on FINE-GL reflects a true change at the 95% confidence level. The minimal clinically important difference (MCID) is defined as the smallest measured change in score that patients perceive to be important. If SDC < MCID, it is possible to distinguish a clinically important change from a measurement error with high certainty.<sup>33</sup> However, it is important to note that the MCIDs derived in this study were based on a shorter time interval without intervention. Further research is needed to determine if the interval between the study assessments is associated with the magnitude of MCID.

Furthermore, concurrent validity was demonstrated by its degree of correlation with the BIS and MAS. As expected, FINE-GL was highly correlated with the MAS. In contrast, body image and WHOQoL had moderate and low correlations with FINE-GL as we expected, respectively. This is because FINE-GL is a specific tool to measure appraisal and distress due to GL and not for general body image and general QoL.

This study has several limitations. First, to confirm convergent validity, we did not include an existing questionnaire that exactly evaluated distress due to GL. However, there are limited PROMs that specifically assess GL-related severity and distress. Second, as this study was cross-sectional, it could not evaluate the longitudinal validity or responsiveness of the FINE-GL to changes in distress levels over time. Longitudinal studies are needed to confirm its sensitivity to clinical improvements or deterioration. Third, we recruited from only 2 local hospitals in Korea, which limits the generalizability of these results. However, we recruited people who were concerned about GL from different groups from multiple sites, which would capture a wider range of potential subjects for FINE-GL.

In conclusion, FINE-GL is a reliable, valid, and comprehensive PROM to measure GL severity and distress. While previous instruments have only examined wrinkle severity, the FINE-GL included specific questions for GL. Thus, the FINE-GL is useful for capturing the patient's perspective and identifying patient-reported psychosocial effects, which is a current clinical emphasis.<sup>34</sup> In addition, other tools measure the effect of treatment, while the FINE-GL can be used to measure distress due to GL itself. Thus, the FINE-GL could be used to measure pre-treatment patient status as well as efficacy endpoints.

## Ethics Approval and Consent to Participate

The study was approved by the Institutional Review Board of the Samsung Medical Center, Seoul, Republic of Korea (SMC-2021-07-166). Informed consent was obtained from the study participants. A statement to confirm that all methods were carried out in accordance with relevant FDA PRO guidelines and regulations.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was supported by Medytox Inc.

## Disclosure

The authors declare that they have no competing interests in this work.

## References

- Rossi AM, Eviatar J, Green JB, et al. Signs of Facial aging in men in a diverse, multinational study: timing and preventive behaviors. *Dermatol Surg*. 2017;43(2):S210–S220. doi:10.1097/DSS.0000000000001293
- Dayan S, Yoelin SG, De Boulle K, Garcia JK. The psychological impacts of upper facial lines: a qualitative, patient-centered study. *Aesthet Surg J Open Forum*. 2019;1(2):ojz015. doi:10.1093/asjof/ojz015
- Macdonald MR, Spiegel JH, Raven RB, Kabaker SS, Maas CS. An anatomical approach to glabellar rhytids. *Arch Otolaryngol Head Neck Surg*. 1998;124(12):1315–1320. doi:10.1001/archotol.124.12.1315
- Rzany B, Dill-Müller DO, Grablowitz D, Heckmann M, Caird D; German-Austrian Retrospective Study Group. Repeated botulinum toxin A injections for the treatment of lines in the upper face: a retrospective study of 4,103 treatments in 945 patients. *Dermatologic Surg*. 2007;33(s1):S18–S25. doi:10.1111/j.1524-4725.2006.32327.x
- The Korean Association for Laser dat. Botulinum toxin treatment status survey. Available from: <https://www.medicaltimes.com/Main/view.html?ID=1130370>. Accessed February 13, 2025.
- Hatzis J. The wrinkle and its measurement—: a skin surface Profilometric method. *Micron*. 2004;35(3):201–219. doi:10.1016/j.micron.2003.11.007
- Luebberding S, Krueger N, Kerscher M. Comparison of validated assessment scales and 3D digital fringe projection method to assess lifetime development of wrinkles in men. *Skin Res Technol*. 2014;20(1):30–36. doi:10.1111/srt.12079
- Messaraa C, Metois A, Walsh M, et al. Wrinkle and roughness measurement by the Antera 3D and its application for evaluation of cosmetic products. *Skin Res Technol*. 2018;24(3):359–366. doi:10.1111/srt.12436

9. Hersant B, Abbou R, SidAhmed-Mezi M, Meningaud JP. Assessment tools for facial rejuvenation treatment: a review. *Aesthetic Plast Surg.* 2016;40(4):556–565. doi:10.1007/s00266-016-0640-y
10. Young VL, Hutchison J. Insights into patient and clinician concerns about scar appearance: semiquantitative structured surveys. *Plast Reconstr Surg.* 2009;124(1):256–265. doi:10.1097/PRS.0b013e3181a80747
11. Pompilus F, Burgess S, Hudgens S, Banderas B, Daniels S. Development and validation of a novel patient-reported treatment satisfaction measure for hyperfunctional facial lines: facial line satisfaction questionnaire. *J Cosmet Dermatol.* 2015;14(4):274–285. doi:10.1111/jocd.12166
12. Yaworsky A, Daniels S, Tully S, et al. The impact of upper facial lines and psychological impact of crow's feet lines: content validation of the Facial Line Outcomes (FLO-11) Questionnaire. *J Cosmet Dermatol.* 2014;13(4):297–306. doi:10.1111/jocd.12117
13. 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. *Clin Plast Surg.* 2013;40(2):249–260. doi:10.1016/j.cps.2012.12.001
14. Kosowski TR, McCarthy C, Reavey PL, et al. A systematic review of patient-reported outcome measures after facial cosmetic surgery and/or nonsurgical facial rejuvenation. *Plast Reconstr Surg.* 2009;123(6):1819–1827. doi:10.1097/PRS.0b013e3181a3f361
15. Morkink LB, Prinsen CA, Patrick DL, et al. COSMIN study design checklist for patient-reported outcome measurement instruments. 2019.
16. Services USDoHaH, Administration FaD, (CDER) CfDEaR, (CBER) CfBEaR, (CDRH) CfDaRH. Guidance for industry: patient-reported outcome measures—Use in medical product development to support labeling claims. 2009.
17. DeVellis RF, Thorpe CT. *Scale Development: Theory and Applications.* Sage publications; 2021.
18. Baker FB. The basics of item response theory. 2001.
19. Bryant FB, Yarnold PR. Principal-components analysis and exploratory and confirmatory factor analysis. *Reading and understanding multivariate statistics.* American Psychological Association; 1995:99–136.
20. Rubinshtein R, Halon DA, Gaspar T, et al. Usefulness of 64-slice cardiac computed tomographic angiography for diagnosing acute coronary syndromes and predicting clinical outcome in emergency department patients with chest pain of uncertain origin. *Circulation.* 2007;115(13):1762–1768. doi:10.1161/CIRCULATIONAHA.106.618389
21. Ximénez C. Recovery of weak factor loadings in confirmatory factor analysis under conditions of model misspecification. *Behav Res Methods.* 2009;41(4):1038–1052. doi:10.3758/BRM.41.4.1038
22. The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med.* 1998;28(3):551–558. doi:10.1017/s0033291798006667
23. Hopwood P, Fletcher I, Lee A, Al Ghazal S. A body image scale for use with cancer patients. *Eur J Cancer.* 2001;37(2):189–197. doi:10.1016/s0959-8049(00)00353-1
24. Flynn TC, Carruthers A, Carruthers J, et al. Validated assessment scales for the upper face. *Dermatol Surg.* 2012;38(2 Spec No.):309–319. doi:10.1111/j.1524-4725.2011.02248.x
25. Ursachi G, Horodnic IA, Zait A. How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Econ Finance.* 2015;20:679–686. doi:10.1016/S2212-5671(15)00123-9
26. de Souza JA, Yap BJ, Wroblewski K, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: the validation of the COMprehensive Score for financial Toxicity (COST). *Cancer.* 2017;123(3):476–484. doi:10.1002/cncr.30369
27. Kenaszchuk C, MacMillan K, van Soeren M, Reeves S. Interprofessional simulated learning: short-term associations between simulation and interprofessional collaboration. *BMC Med.* 2011;9(1):29. doi:10.1186/1741-7015-9-29
28. Ries JD, Echternach JL, Nof L, Gagnon Blodgett M. Test-retest reliability and minimal detectable change scores for the timed “Up & Go” test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys Ther.* 2009;89(6):569–579. doi:10.2522/ptj.20080258
29. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biometrical J.* 2008;50(3):419–430. doi:10.1002/bimj.200710415
30. Harper B, Tiggemann M. The effect of thin ideal media images on women's self-objectification, mood, and body image. *Sex Roles.* 2008;58(9):649–657. doi:10.1007/s11199-007-9379-x
31. Sabik NJ. Is social engagement linked to body image and depression among aging women? *J Women Aging.* 2017;29(5):405–416. doi:10.1080/08952841.2016.1213106
32. Saucier MG. Midlife and beyond: issues for aging women. *J Couns Dev.* 2004;82(4):420–425. doi:10.1002/j.1556-6678.2004.tb00329.x
33. 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 dermatol.* 2016;152(4):443–451. doi:10.1001/jamadermatol.2016.0018
34. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes.* 2006;4:79. doi:10.1186/1477-7525-4-79

## Patient Preference and Adherence

### Publish your work in this journal

Patient Preference and Adherence is an international, peer-reviewed, open access journal that focusing on the growing importance of patient preference and adherence throughout the therapeutic continuum. Patient satisfaction, acceptability, quality of life, compliance, persistence and their role in developing new therapeutic modalities and compounds to optimize clinical outcomes for existing disease states are major areas of interest for the journal. This journal has been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/patient-preference-and-adherence-journal>

**Dovepress**  
Taylor & Francis Group