Original Article

Measuring patient-reported outcomes in orthognathic surgery: Linguistic and psychometric validation of the Mandarin Chinese version of FACE-Q instrument

Yu-Ying Su [a,1], Rafael Denadai [b,1], Cheng-Ting Ho [a], Bo-Ru Lai [b,c], Lun-Jou Lo [b,c,*]

[a] Division of Orthodontics, Department of Dentistry, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan
[b] Department of Plastic and Reconstructive Surgery and Craniofacial Research Center, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan
[c] College of Medicine, Chang Gung University, Taoyuan, Taiwan

ARTICLE INFO

ABSTRACT

Background: Orthognathic surgery is useful for correction of dental malocclusion and improvement of facial appearance. The FACE-Q is a patient-reported outcome instrument for evaluation of surgical and psychosocial effect. The purposes of this study were to conduct a linguistic validation of all FACE-Q scales to Mandarin Chinese, to test the orthognathic surgery-related scales for reliability and validity, and to evaluate the effect of orthognathic surgery.
Methods: All FACE-Q scales and checklists were translated from English to Mandarin Chinese according to international recommendations: forward translations, backward translation, and cognitive interviews. Psychometric testing of orthognathic surgery-related scales of translated version was administered to patients with facial deformities and history of orthognathic surgery (n = 53; 17 scales) or no history of orthognathic surgery (n = 44; 11 scales), and control subjects (n = 57; 11 scales).
Results: All FACE-Q scales and checklists were linguistically validated into Mandarin Chinese. The contents were confirmed valid among Mandarin Chinese-speaking population. The FACE-Q scales had excellent internal consistency (Cronbach's alpha >0.70) and discriminated (p < 0.05) well between patients before and after orthognathic surgeries and normal subjects.
Conclusions: This study discovered significant benefit of orthognathic surgery on improving facial appearance and psychosocial function, as compared with the non-surgical patients and normal controls.

At a glance of commentary

Scientific background on the subject

Face-Q is a useful patient-reported outcome instrument for evaluation of surgical and psychosocial effect after facial plastic surgery. This study reported the validation of the Mandarin Chinese version of Face-Q tool to be applied to all Chinese patients. Standard procedures were followed.

What this study adds to the field

Patients with malocclusion undergoing orthognathic surgery were tested, and proved significant benefit of the orthognathic surgery on improving the facial appearance and psychosocial function, in addition to correction of malocclusion, as compared with the nonsurgical patients and normal controls. The Face-Q could be applied for all facial cosmetic related treatment.

Orthognathic surgery (OGS) has been adopted to correct several types of congenital or acquired facial deformities associated with malocclusion [1–5]. By altering the position of maxillary and mandibular segments, the facial bony structure and soft-tissue envelope can be harmonized, leading to substantial anatomical and functional correction, as well as improvement of overall facial appearance, symmetry, and esthetics [1–5]. Given that patient satisfaction may be one of the most important determinants for success in OGS [6–10], it is important to elicit and analyze patient-reported outcomes (PRO) for this surgical modality [6–12]. Most of previous PRO studies on OGS were focused on quality of life or functional outcomes [6–10], but few investigations have included facial appearance satisfaction-specific measures in their results [11,12].

FACE-Q instrument, a PRO measure for facial aesthetic plastic surgery procedures created according to rigorous international guidelines [13–17], encompasses a set of more than 40 independently functioning scales and checklists measuring concepts that matter to facial aesthetic patients. Four main constructs were included, appearance appraisal, quality of life, adverse effect, and patient's experience of care [13–17]. The English-language FACE-Q version is a psychometrically validated and effective PRO measure to assess different facial aesthetic interventions, including face lift, rhinoplasty, eyelid surgery, minimally invasive procedures, and OGS [12–17]. However, it cannot be adopted in non-English speaking populations as a careful process of linguistic and psychometric validation in the specific cultural context is required to achieve maximum semantic, idiomatic, experiential, and conceptual equivalence between the source and target instruments [18,19]. The FACE-Q instrument was recently translated for French-speaking, Italian-speaking, and German-speaking populations [20], but it is not currently translated to use in the Mandarin Chinese-speaking population. Although OGS-related scales of FACE-Q instrument have also recently been translated for Hong Kong Chinese patients, it was performed in a Cantonese Chinese-speaking population [11].

Validating multiple language versions of FACE-Q instrument is extremely relevant for future international cooperation [21]. As Mandarin Chinese is the most widely used language with more than 1 billion Mandarin speakers worldwide, in addition to 2 million or more living in the United States speaking the language [22]. The purposes of this study were (1) to conduct a linguistic validation of all FACE-Q scales to Mandarin Chinese, and (2) to assess patient satisfaction after orthognathic surgery using relevant FACE-Q measures.

Material and Methods

This was a cross-sectional study carried out at the Craniofacial Research Center, Chang Gung Memorial Hospital after approval by the Institutional Review Board. Written consent was obtained from all subjects, in accordance with the principles outlined in the Declaration of Helsinki.

Subjects

For linguistic validation phase, facial aesthetic interventions patients were recruited. For psychometric validation phase, 3 groups of subjects were included: consecutive subjects diagnosed with facial deformity, dental malocclusion and prior to the orthognathic surgery (pre-OGS group), subjects at least 6 months after receiving OGS (post-OGS group), and normal subjects with no facial deformity, dental malocclusion, or history of OGS (control group). All subjects had Mandarin Chinese as their mother language. All OGS procedures were performed by one senior surgeon (LJL) following standard planning and surgical principles [1–5]. Exclusion criteria were subjects with abnormal mentality that would impair the questionnaire application, as well as having previous facial trauma, facial surgery, or facial aesthetic procedure.

FACE-Q instrument

For linguistic validation phase, all included subjects (n = 10) completed all FACE-Q scales and checklists. For psychometric validation phase, subjects of control group and pre-OGS group completed 11 FACE-Q scales (Satisfaction with facial appearance overall, Satisfaction with cheeks, Satisfaction with cheekbones, Satisfaction with lower face and jawline, Satisfaction with lips, Satisfaction with chin, Appraisal of area under chin, Satisfaction with nose, Satisfaction with nostrils, Social function and Psychological well-being). Subjects of post-OGS group also completed these 11 scales and six additional scales (Recovery-early life impact, Satisfaction with outcomes, Satisfaction with decision, Satisfaction with medical team, Satisfaction with office staff, and Satisfaction with surgeon). As recommended in the original FACE-Q studies [12–17], all subjects completed the instrument unaided and independently; they answered all questions based on their own understanding of the items.

Linguistic validation

The translation and cross-cultural adaptation process of FACE-Q instrument was conducted according to international standards and recommendations [19]. In the first step, we obtained the permission from the original developers of the FACE-Q, represented by the Mapi Research Trust (https://eprovide.mapi-trust.org). In the second step, two forward translations were performed by two independent translators using simple terminology rather than literal translations. In the third step, a primary reconciliation version (defined as the first Mandarin Chinese version) was constructed by a local committee (two translators, orthodontics, and plastic surgeons) with experience in management of facial aesthetic patients, knowledge of Mandarin Chinese and English languages, and on the research methodology. In the 4th step, the first Mandarin Chinese version was back-translated into English by a translator which had no access to the original FACE-Q instrument. In the 5th step, this back-translated version was reviewed by the local committee in order to ensure semantic and idiomatic equivalence between the original and translated versions. Items whose meaning was not maintained were re-translated in an iterative manner until an acceptable result was achieved, leading to the approved harmonized English version, and consequently the second Mandarin Chinese version. This phase also had support from authors of the original instrument. In the 6th step, cognitive debriefing interviews and post-cognitive review were performed. In the final step, the final Mandarin Chinese version (linguistically validated translation) of the FACE-Q tool constructed by consensus/proofreading was then ready for validation [Fig. 1].

Cognitive debriefing interviews

A cognitive debriefing interview was performed with a group of 5 patients per interview to ensure that the meaning of the instruction, recall period, items, and response options were the same, and that the wording was appropriate [23]. We conducted a face-to-face interview with each group by reading the translated instrument and also adopting the FACE-Q user manual provided by the Mapi Research Trust. Participants were encouraged to make comments, express their difficulties, and give their opinion on possible changes that they thought would enable comprehension. They were also requested to explain why any words were difficult to understand. Items deemed problematic by two or more participants were revised for clarity [24].

Psychometric validation

Internal-consistency reliability was calculated using data on all subjects according to each specific scale, as established in the FACE-Q instrument subhead; for example, patient experience scales were evaluated based on scores from subjects of OGS group. Discriminant (divergent) validity was evaluated by comparing the scores of pre-OGS, post-OGS, and control groups.

Based on previous FACE-Q studies [11,15,16], additional aspects of construct validity (i.e., further contrast, discriminant, and convergent validity) of Mandarin Chinese FACE-Q version were assessed to examine scale performance. Pre- and post-OGS scores for specific scales (Satisfaction with facial appearance overall, Satisfaction with lower face and jawline, Satisfaction with chin, Satisfaction with nose, Satisfaction with nostrils, Psychological well-being, and Social function) were compared to examine the contrast validity. For this, pairs of pre- and post-OGS subjects were matched based on age, gender, and skeletal class III pattern, as previous validation reports [11,12]. Post-OGS scores were predicted to be significantly higher, as it was assumed that patients after OGS were more satisfied with their appearance and demonstrated a better psychosocial status [11].

Scores for particular scales (Satisfaction with facial appearance overall, Satisfaction with lower face and jawline,
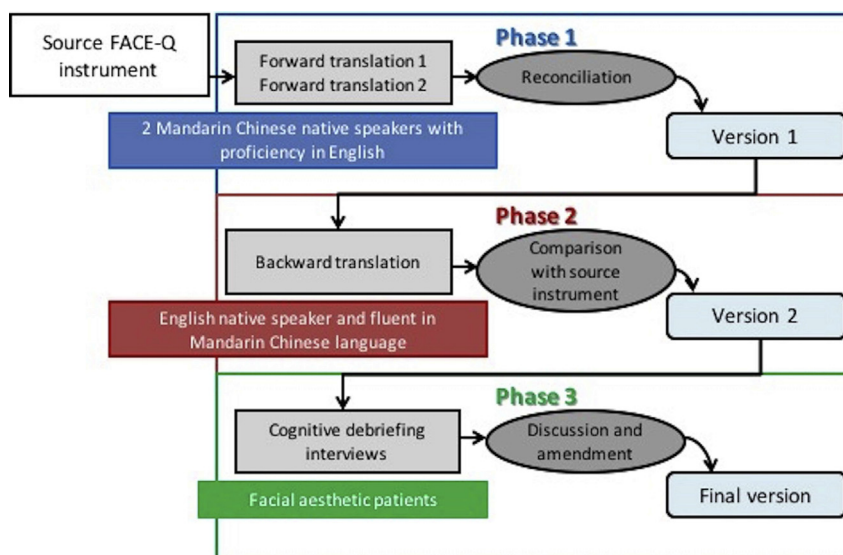

Fig. 1 Linguistic validation steps.

Satisfaction with chin, Satisfaction with nose, Satisfaction with nostril, Psychological well-being, Social function, Satisfaction with decision, Satisfaction with outcome, and Recovery early life impact) were correlated to determine the extent to which they measured separate but related constructs. These correlations were hypothesized to be moderate as these scales were developed to measure distinct but related clinical variables [11,15,16]. Scores for these scales were also correlated with patients' sociodemographic characteristics to determine the extent to which a scale may be susceptible to bias due to age and gender. Theses correlations were predicted to be low [11,15,16].

Statistical analysis

For the descriptive analysis, the mean was used for metric variables, and percentages were given for categorical variables. Using Excel (Microsoft Corporation, USA), the total score for each FACE-Q scale was calculated by adding the scores of each item of that specific scale. The sum score was then converted to an equivalent Rash score ranging from 0 to 100 with higher scores indicating a better outcome [13–17]. Standard psychometric tests and criteria were adopted to evaluate the reliability and validity [25–27]. For internal consistency reliability, the Cronbach's alpha of each scale, the Cronbach's of all but the item concerned, the inter-item correlation coefficients, and the item-total correlation coefficients were calculated. A Cronbach's alpha >0.70 was considered acceptable [25]. For inter-item correlation and item-total correlation, ideal coefficients were defined as 0.40–0.50 and $\geq$0.20, respectively [28–31]. For discriminant validity, intergroup comparative analyses (pre-OGS vs. post-OGS vs. control groups) were performed with paired t-test [11,12]. Inter-scale correlation coefficients were interpreted as high ($r > 0.70$), moderate ($r = 0.30$ to 0.70), or low ($r < 0.30$) [11,15,16,27]. Two-sided values of $p < 0.05$ were considered statistically significant. All analyses were performed using SPSS version 20.0 (Chicago, IL).

Results

Forward translations

Reconciliation of the two independent forward translations revealed some discrepancies related to the wording of the items. These discrepancies were properly resolved after a consensus meeting held between the 2 forward translators and the local committee [Table 1]. Instructions, recall period, and response options were translated without discrepancies. In this translation phase, we noticed that there were three items of Satisfaction with cheekbones scale (Item b. How high your cheekbones look? Item f. How sculpted your cheekbones look? Item i. How prominent your cheekbones look?) which presents a culture diversity as Asian population generally do not interpret high, sculpted, or prominent cheekbones as aesthetically pleasing. We followed all guidelines for translation process and maintained this scale for further interpretation in cognitive interviews.

Table 1 Examples of discrepancies and decisions of the forward translations of the FACE-Q into Mandarin Chinese.

| FACE-Q scales (Satisfaction)	Original item	Translator 1 (Mandarin Chinese/English)	Translator 2 (Mandarin Chinese/English)	Solution by consensus (Mandarin Chinese/English)
Facial appearance overall	e. How fresh your face looks?	您的臉看起來具有青春氣息？ Is your face looks young?	你的臉看起來有多清新？ How fresh your face looks?	你的臉看起來有多清新？ How fresh your face looks?
Chin	a. The style of your chin (e.g. masculine or feminine)?	你的下巴的類型(例如:男性或女性)? The Classification of your chin (e.g. male or female)?	您下額的風格（例如:陽剛或陰柔）？ The style of your chin (e.g. masculine or feminine)?	你的下巴的類型(例如:男性化或女性化)? The style of your chin (e.g. masculine or feminine)?
Chin	d. How well your chin suits your face?	你的下巴與你的臉部的適合程度？ How well your chin suits your face?	你的下巴適合你的臉部嗎？ Does your chin suit for your face?	你的下巴與你的臉部的適合程度？ How well your chin suits your face?
Eyes	d. How alert (not tired) your eyes look?	你的眼睛看起來多有神？ How wakeful your eyes look?	你的眼睛看起來有多警性？ How alertness your eyes look?	你的眼睛看起來多有神？ How wakeful your eyes look?

(Cupids bow)

邱比特弓
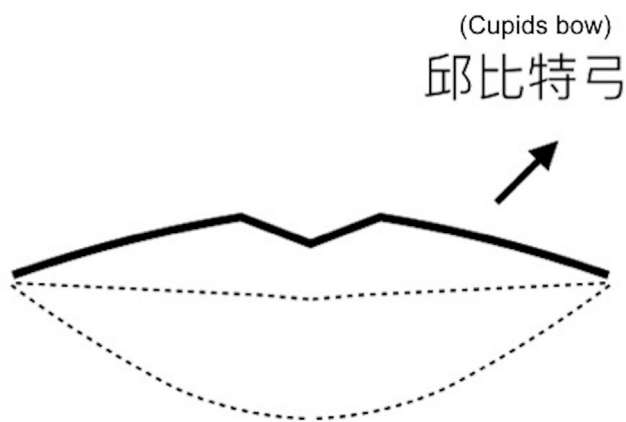
Fig. 2 Stylistic drawing of the cupids bows adopted for Satisfaction with lips scale.
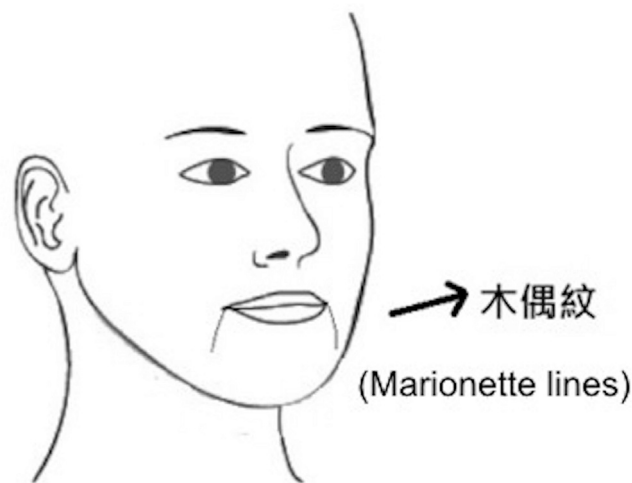
木偶紋

(Marionette lines)

Fig. 3 Stylistic drawing of the marionette lines adopted for Satisfaction with marionette lines scale.

Back translation

The harmonized backward translation was compared to the original English FACE-Q version, with no discrepancies in the instructions, recall period, and response options. It revealed some discrepancies in specific words. These translated words were considered to have a different meaning than the original version and required revision (re-translation until a satisfactory result was achieved).

Cognitive debriefing interviews

Ten facial aesthetic interventions (surgical and/or minimally invasive procedures) patients (age from 25 to 39; 5 female) participated cognitive debriefing interviews. They were encouraged to make comments, express their difficulties, and give their opinion on possible changes that they thought would enable comprehension. Overall, the participants had a good understanding of the form filling, titles, instructions, recall period, and response items. Only specific words were subject to difficulties of understanding or interpretation. It was addressed and reviewed by the interview and the local committee. In the first round of cognitive interview (n = 5), it was noticed that most of the patients had misunderstanding and/or misinterpretation about two scales, namely Appraisal of lines-marionette and Satisfaction with lips (item g. How turned up your upper lip [Cupid's bow] looks?) scale. Therefore, two stylistic drawings related to marionette lines and Cupid's bow were included [Figs. 2 and 3]. A second round of cognitive interview (n = 5) demonstrated that all subjects understood the meaning without further issues. All participants paraphrased the items with ease and reported no difficulty in comprehending the items. Since no difficulty was expressed for any item/word, no additional changes or round of cognitive interview were implemented.

Final version

Minor changes to punctuation and standardization of format and layout (as in the original instrument) were required upon final proofreading of the Mandarin Chinese FACE-Q version. The final cross-culturally adapted, Mandarin Chinese FACE-Q version preserved the main features of the original English instrument.

Assessment of facial appearance in patients receiving orthognathic surgery

A total of 154 subjects [Table 2] were enrolled for reporting facial outcome using the Mandarin Chinese FACE-Q tool. The scores for each FACE-Q scale were compiled in [Table 3]. All

Table 2 Characteristics of sample included fr psychometric validation of Mandarin Chinese FACE-Q version.			
Characteristics	Pre-OGS group	Post-OGS group	Control group
Participants, n	44	53	57
Age, yr M±SD	21.80 ± 5.11	23.25 ± 4.64	22.74 ± 1.28
Gender, n (%)			
Male ＼ Female	22 (50) ＼ 22 (50)	27 (50.9) ＼ 26 (49.1)	29 (59.9) ＼ 28 (49.1)
Previous procedures, n (%)			
Bimaxillary OGS (Yes＼No)	0 (0) ＼ 44 (100)	53 (100) ＼ 0 (0)	0 (0) ＼ 57 (100)
Facial aesthetic procedures[a] (Yes＼No)	0 (0) ＼ 44 (100)	0 (0) ＼ 53 (100)	0 (0) ＼ 57 (100)
Time from OGS to FACE-Q completion, yr M±SD	−	1.87 ± 0.76	−

Abbreviations: OGS: orthognathic surgery; Pre-OGS group: subjects with no history of OGS; Post-OGS group: subjects with history of OGS; yr: years; M: mean; SD: standard deviation; n: number of subjects; %: percentage of subjects; −: not applicable.
[a] Face lift: eyelid surgery, fillers and＼or toxin botulinum.

Table 3 Scores and discriminant validity for Mandarin Chinese FACE-Q scales among groups of subjects.

FACE-Q scales	Pre-OGS group	Post-OGS group	Control group	Pre-OGS group *vs* Post-OGS group	Post-OGS group *vs* Control group	Pre-OGS group *vs* Control group
				p-value	*p-value*	*p-value*
Appearance appraisal scales						
Facial appearance overall	44.52 ± 19.36 (0−100)	68.06 ± 20.38 (35−100)	56.51 ± 16.01 (19−93)	<0.01	<0.01	<0.01
Cheeks	46.02 ± 19.13 (0−100)	66.11 ± 21.11 (30−100)	60.67 ± 20.84 (0−100)	<0.01	<0.01	<0.01
Cheekbones	68.05 ± 22.23 (11−95)	74.83 ± 21.53 (25−100)	72.00 ± 20.94 (0−100)	<0.01	>0.05	<0.01
Chin	35.57 ± 22.91 (0−87)	67.83 ± 18.56 (57−100)	64.18 ± 24.55 (0−100)	<0.01	<0.01	<0.01
Under chin	69.52 ± 19.75 (30−100)	83.00 ± 15.23 (42−100)	77.77 ± 18.01 (30−100)	<0.01	<0.01	<0.01
Lower face and jawline	35.11 ± 26.72 (0−100)	72.32 ± 21.00 (34−100)	63.74 ± 23.24 (0−100)	<0.01	<0.01	<0.01
Nose	44.16 ± 26.15 (0−100)	59.79 ± 24.54 (0−100)[a]	59.93 ± 19.48 (35−100)	<0.01	>0.05	<0.01
Nostrils	46.57 ± 25.70 (0−100)	59.11 ± 28.52 (0−100)[a]	67.56 ± 24.73 (0−100)	<0.01	<0.01	<0.01
Lips	47.00 ± 21.41 (0−100)	64.79 ± 22.17 (25−100)	66.16 ± 19.24 (32−100)	<0.01	>0.05	<0.01
Quality-of-life scales						
Social function	42.89 ± 21.66 (0−100)	56.72 ± 23.59 (0−100)[a]	54.53 ± 21.65 (9−100)	<0.01	>0.05	<0.01
Psychological well-being	53.16 ± 24.21 (5−100)	66.42 ± 23.09 (36−100)	69.74 ± 21.26 (23−100)	<0.01	<0.01	<0.01
Early life impact	−	76.40 ± 25.81 (26−100)	−	−	−	−
Patient experience scales						
Outcome	−	71.35 ± 21.65 (0−100)[a]	−	−	−	−
Decision	−	76.66 ± 22.24 (33−100)	−	−	−	−
Medical team	−	91.26 ± 15.24 (53−100)	−	−	−	−
Office staff	−	89.62 ± 17.04 (33−100)	−	−	−	−
Surgeon	−	91.92 ± 16.14 (37−100)	−	−	−	−

Data are presented as mean ± standard deviation (range); Abbreviations: OGS: orthognathic surgery; Pre-OGS group: subjects with no history of OGS; Post-OGS group: subjects with history of OGS; −: not applicable.

[a] One (social function and outcome scales) and three (nose and nostril scales) patients had zero values.

Table 4 Reliability for Mandarin Chinese FACE-Q scales.

FACE-Q Scales	Cronbach's Alpha	Cronbach's Alpha if item deleted[a]	Inter-item correlation coefficients	Item-total correlation coefficients
Appearance appraisal scales				
Facial appearance overall	0.94	0.91—0.96	0.42—0.49	0.73—0.82
Cheeks	0.90	0.88—0.93	0.43—0.47	0.67—0.75
Cheekbones	0.88	0.84—0.91	0.40—0.43	0.59—0.67
Chin	0.93	0.90—0.94	0.43—0.46	0.72—0.79
Under chin	0.91	0.87—0.93	0.41—0.47	0.68—0.73
Lower face and jawline	0.93	0.86—0.95	0.42—0.45	0.70—0.81
Nose	0.93	0.89—0.95	0.44—0.48	0.67—0.77
Nostrils	0.91	0.90—0.94	0.42—0.46	0.69—0.80
Lips	0.92	0.87—0.93	0.40—0.46	0.54—0.68
Quality-of-life scales				
Social function	0.91	0.87—0.93	0.45—0.48	0.63—0.75
Psychological well-being	0.93	0.90—0.95	0.42—0.47	0.69—0.77
Early life impact	0.86	0.82—0.89	0.40—0.44	0.56—0.63
Patient experience scales				
Outcome	0.90	0.88—0.91	0.41—0.46	0.64—0.79
Decision	0.92	0.90—0.94	0.43—0.45	0.58—0.70
Medical team	0.89	0.86—0.92	0.40—0.46	0.59—0.72
Office staff	0.90	0.88—0.93	0.44—0.47	0.60—0.71
Surgeon	0.93	0.89—0.95	0.42—0.46	0.62—0.75

[a] Indicates internal consistency of remaining items if the elected item was deleted from the total score.

Table 5 Correlation of Mandarin Chinese FACE-Q scales (satisfaction with facial appearance) and sociodemographic characteristics.

Parameters	Facial appearance overall		Lower face and jawline		Satisfaction with chin		Satisfaction with nose		Satisfaction with nostril	
	r	p-value	r	p-value	r	p-value	r	p-value	r	p-value
Facial appearance overall										
Lower face and jawline	0.62	<0.01								
Satisfaction with chin	0.46	<0.01	0.49	<0.01						
Satisfaction with nose	0.39	<0.01	0.33	<0.01	0.36	<0.01				
Satisfaction with nostril	0.43	<0.01	0.40	<0.01	0.37	<0.01	0.35	<0.01		
Age	0.14	>0.05	0.27	>0.05	0.16	>0.05	0.02	>0.05	0.09	>0.05
Gender	0.08	>0.05	−0.05	>0.05	0.22	>0.05	0.11	>0.05	0.21	>0.05

Table 6 Correlation of Mandarin Chinese FACE-Q scales (quality of life and patient experience) and sociodemographic characteristics.

Parameters	Psychological well-being		Social function		Satisfaction decision		Satisfaction outcome		Recovery early Life Impact	
	r	p-value	r	p-value	r	p-value	r	p-value	r	p-value
Psychological well-being										
Social function	0.57	<0.01								
Satisfaction with decision	0.51	<0.01	0.48	<0.01						
Satisfaction with outcome	0.44	<0.01	0.63	<0.01	0.37	<0.01				
Early Life Impact	0.26	<0.01	0.21	<0.01	0.17	<0.01	0.24	<0.01		
Age	0.14	>0.05	0.05	>0.05	0.09	>0.05	0.15	>0.05	0.18	>0.05
Gender	0.17	>0.05	0.08	>0.05	0.13	>0.05	0.20	>0.05	0.11	>0.05

Cronbach's alpha values, item-total correlation coefficients, and inter-item correlation coefficients satisfied the criterion of >0.70, ≥0.20, and 0.40−0.50, respectively [Table 4].

Each of the FACE-Q scales had discriminant validity [Table 3]. Comparative analyses performed on matched pre- and post-OGS patients revealed significantly (all $p < 0.01$) higher post-OGS scores compared with pre-OGS scores for all tested scales. All scores were significantly (all $p < 0.01$) lower in pre-OGS group than the control group. All scores were significantly (all $p < 0.01$) higher in post-OGS group than the control group, except for *Satisfaction with cheekbones*, *Satisfaction with nose*, *Satisfaction with lips*, and *Social function* scales with similar (all p > 0.05) scores between post-OGS and control groups, and for *Satisfaction with nostrils* scale with significantly ($p < 0.01$) higher scores for control group than post-OGS group.

Scale validity was also supported by the interscale correlations [Tables 5 and 6]. All interscale correlations were moderate (all *r-values* between 0.30 and 0.70), except for "*Early life impact*" scale (all r = 0.30). The correlations between the tested scales and sociodemographic characteristics were low (all $r < 0.30$).

Discussion

Although FACE-Q instrument has been applied to measure facelift and double-eyelid blepharoplasty outcomes in Asian patients, there is no explicit information about the language of the used scales and/or data regarding the translation and linguistic validation process for use in a probable Chinese-speaking population [32,33]. It is inappropriate to simply translate and adopt an existing PRO instrument in another linguistic or cultural context as the validity of the resulting instrument may be dubious, hamper research outcomes, and lead to erroneous conclusions [18,19]. PRO instruments require rigorous validation process in the target population [18,19].

We adopted the recommendations provide by the Mapi Research Trust [34], as linguistic validation process should follow a scientifically valid guideline [35]. We noticed a potential cultural diversity, as aesthetic perception of cheekbones may influence three items of *Satisfaction with cheekbones* scale. Overall, these items transmitted the central idea that if the cheekbones are more high, sculpted, or prominent, the more satisfaction with the cheekbones the subject present. Prominent cheekbones are considered a classic beauty trait in Western countries [36,37], but many Asians find them undesirable [38−40], as a soft facial appearance (namely, a small, oval-shaped and slender face with smooth zygomatic bones) seems to be more aesthetically pleasing and attractive [38−40]. Reduction zygoma surgery has been commonly performed among Korean, Japanese, and Taiwanese populations [38−40], whereas aesthetic zygoma interventions mostly involves augmentation in Western cultures [36,37]. Our group [38] previously demonstrated that all patients claimed that aesthetic facial bone contouring for their square faces had a positive influence on their self-confidence, on helping them choose their occupations and on their performance at work, highlighting the psychosocial impact of facial contour appearance in Asians.

This cultural difference regarding cheekbones appearance re-emphasizes the importance of linguistic validation process [18,19,35]. We reported this cultural diversity for Mapi Research Trust as it may potentially generate misinterpretation of items and response options. We fully followed the received recommendations and we maintained the conceptual framework of *Satisfaction with cheekbones* scale, making it possible to compare our findings with those of other FACE-Q studies [11−17].

For cognitive debriefing interviews, subjects were selected to represent, as well as possible, the population that the FACE-Q instrument will target, namely facial aesthetic patients [13−17]. As the participants had misunderstanding and misinterpretation about *Appraisal of lines-marionette* and *Satisfaction with lips* (item related to cupids bow) scales, two stylistic drawings were included to clarify exactly which part of the face the scales referred to. The FACE-Q also have adopted anatomical pictures with similar purposes [13−17].

As defined by the developers of FACE-Q, only those scales and/or checklists relevant to a particular patient or procedure need to be completed [13−17]. Therefore, the Cantonese Chinese and English FACE-Q versions studied 8 and 5 OGS-related scales, respectively [11,12]. We tested these same scales [11,12] and 9 supplementary scales, because we hypothesized that these additional scales have the potential to expand the understanding, recognition, and appraisal of patients presenting for OGS and also the changes and experiences after OGS.

As in English and Cantonese Chinese FACE-Q instruments [11,12], the Mandarin Chinese version meet internal consistency reliability, construct validity, and discriminant validity [25−31] for all tested scales. All inter-item correlation coefficients were within the ideal range, and each item exhibited good individual correlation with the rest of the scale, which indicated an absence of redundancy or unnecessary items [28−31]. Interscale correlations were moderate (except for "*Early life impact*" scale), demonstrating that these scales measured distinct but related clinical variables. The "*Early life impact*" scale correlated less than 0.30 with the other scales, indicating that the construct measured by this scale was distinct. The correlations between the tested scales and sociodemographic characteristics were also low (all $r < 0.30$), suggesting that little bias was present.

For intergroup comparisons, the included groups were expected to be different and thus distinctions in FACE-Q scales were likewise expected. This hypothesis was supported, as the post-OGS group reported higher scores on all tested scales than the pre-OGS group. Not surprisingly, pre-OGS group showed lower scores in all tested scales than the control group, stressing the importance of using a PRO instrument developed to address what matter to specific categories of patients or procedures [11−17]. Moreover, post-OGS group presented higher or similar scores (6 scales and 4 scales, respectively) than control subjects.

The improvement of facial harmony, symmetry, and aesthetic after OGS [1−5] may justify, at least partially, these higher scores reported by post-OGS group when compared to the pre-OGS and control groups. However, although the zygomatic bone itself was not repositioned during OGS, subjects of post-OGS group presented higher and similar scores in

Satisfaction with cheekbones scale when compared with pre-OGS and control groups, respectively. Other group [41] showed that all interviewed patients had a statistically significant increase in their subject perceived malar projection score following Le Fort I advancement with or without simultaneous malar augmentation, demonstrating a change of perception about this region independently of the direct local augmentation. We hypothesize that the high-grade enhancement of general facial appearance and dental function and normalization of the profile (e.g., concave to flat or convex) added to specific enhancements of regions directly addressed by OGS (e.g., lips, lower face, and jawline) [1–5] acted to attenuate the criticality threshold related to facial regions not directly addressed by OGS. From a practical point of view, this aspect underlines the relevance to evaluate the face completely, and not only compartmentalized portions [1–5,38–40,42].

Regarding the nose and nostril regions related scales, it has been demonstrated that OGS procedures (i.e., LeFort I osteotomy and movement of the maxilla) affect the position and shape of the nose, including significant changes of the nasolabial angle, columella inclination, nasal tip projection, width of alar base, and nostril show [43,44]. As the increased nostril show has been considered a negative facial appearance in Asian population [45,46], it may explain why subjects of post-OGS group presented lower scores for *Satisfaction with nostrils* scale than controls. Further investigation is necessary to correlate FACE-Q outcomes and nose changes before and after OGS.

For scales related to jawline, lower face, chin and lips regions, the FACE-Q results were within expected as the bone mobilization and its repercussion on these soft-tissue regions have been reported as positive in OGS patients [1–5]. Furthermore, although post-OGS group presented higher scores for *Psychological well-being* scale than pre-OGS group, results remained lower when compared to the control group. Motivation to underwent OGS procedures may partially explain this finding, as we [47] previously demonstrated that patients sought OGS primarily to become more beautiful and to treat a perceived abnormal facial appearance (i.e., concave profile with malocclusion), but what they are actually seeking is treatment for other facial disfigurements (e.g., a flat nose or a square face) that might be associated with, or confused for, a malocclusion [47]. Therefore, this may impact on the maintenance of certain psychological concerns post-surgery. Consequently, our data highlight that it is essential that patients have realistic expectations about OGS and that the surgeons be honest with his or her ability to fulfill them. Surgeons may use our and previous FACE-Q data [11,12] to alert their patients prior to OGS regarding facial appearance, quality of life, and patient experience of care obtained after OGS interventions.

Potential limitations of this study should be addressed. As our and previous [12] FACE-Q validation results reflect a single senior surgeon's experience, the generalizability of our findings cannot be assumed. Although we have a similar sample of male and female subjects, we and others [11] primarily included young adults with skeletal class III pattern. This reflects typical clinical profile of patient referred for OGS in our center [1–5,47]. We did not include a control group composed by aesthetic facial surgery patients, including surgical interventions (facelift and facial skeletal contouring) and minimal invasive procedures (fillers and toxin botulinum) interventions. Although we included a control group with no history of facial aesthetic surgery in order to reflect the reality of the local population, we did not directly investigate the intention of facial surgery within them. As facial contouring surgery and double eyelid surgery are commonly performed in Asians [33,38–40], our results were likely influenced by the desire for facial changes. Finally, although the criteria were satisfied for all psychometric properties evaluated [25–31], we did not include all existing validation tests [25–27]. As in the Cantonese Chinese FACE-Q study [11], we had different patient samples for the pre- and post-OGS groups and responsiveness should be the target of future investigations.

In conclusion, the linguistic validation process of the FACE-Q instrument ensued in the development of a Mandarin Chinese version for use in Mandarin Chinese speaking population. The tested scales of translated version proved to be reliable and valid for assessing patients with facial deformity and malocclusion before and after OGS. There were significant benefits of OGS on improving facial appearance and psychosocial function, as compared with the non-surgical patients and normal controls.

Funding

Conflicts of interest

The authors declare that they have no competing interests.

Acknowledgement

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bj.2019.05.011.

REFERENCES

[1] Lo LJ, Weng JL, Ho CT, Lin HH. Three-dimensional region-based study on the relationship between soft and hard tissue changes after orthognathic surgery in patients with prognathism. PLoS One 2018;13:e0200589.

[2] Ho CT, Lin HH, Liou EJ, Lo LJ. Three-dimensional surgical simulation improves the planning for correction of facial

prognathism and asymmetry: a qualitative and quantitative study. Sci Rep 2017;7:40423.

[3] Wu TY, Lin HH, Lo LJ, Ho CT. Postoperative outcomes of two- and three-dimensional planning in orthognathic surgery: a comparative study. J Plast Reconstr Aesthet Surg 2017;70:1101−11.

[4] Lin HH, Chang HW, Wang CH, Kim SG, Lo LJ. Three-dimensional computer-assisted orthognathic surgery: experience of 37 patients. Ann Plast Surg 2015;74(Suppl 2):S118−26.

[5] Lonic D, Pai BC, Yamaguchi K, Chortrakarnkij P, Lin HH, Lo LJ. Computer-Assisted orthognathic surgery for patients with cleft lip/palate: from traditional planning to three-dimensional surgical simulation. PLoS One 2016;11:e0152014.

[6] Pachêco-Pereira C, Abreu LG, Dick BD, De Luca Canto G, Paiva SM, Flores-Mir C. Patient satisfaction after orthodontic treatment combined with orthognathic surgery: a systematic review. Angle Orthod 2016;86:495−508.

[7] Liddle MJ, Baker SR, Smith KG, Thompson AR. Psychosocial outcomes in orthognathic surgery: a review of the literature. Cleft Palate Craniofac J 2015;52:458−70.

[8] Broers DLM, van der Heijden GJMG, Rozema FR, de Jongh A. Do patients benefit from orthognathic surgery? A systematic review on the effects of elective orthognathic surgery on psychosocial functioning and patient satisfaction. Eur J Oral Sci 2017;125:411−8.

[9] Sonego CL, Bobrowski ÂN, Chagas Jr OL, Torriani MA. Aesthetic and functional implications following rotation of the maxillomandibular complex in orthognathic surgery: a systematic review. Int J Oral Maxillofac Surg 2014;43:40−5.

[10] Soh CL, Narayanan V. Quality of life assessment in patients with dentofacial deformity undergoing orthognathic surgery−a systematic review. Int J Oral Maxillofac Surg 2013;42:974−80.

[11] Tan SK, Leung WK, Tang ATH, Tse ECM, Zwahlen RA. Orthognathic relevant scales of FACE-Q: translation and validation for Hong Kong Chinese patients. Plast Reconstr Surg Glob Open 2017;5:e1608.

[12] Schwitzer JA, Albino FP, Mathis RK, Scott AM, Gamble L, Baker SB. Assessing patient-reported outcomes following orthognathic surgery and osseous genioplasty. J Craniofac Surg 2015;26:2293−8.

[13] Klassen AF, Cano SJ, Scott A, Snell L, Pusic AL. Measuring patient-reported outcomes in facial aesthetic patients: development of the FACE-Q. Facial Plast Surg 2010;26:303−9.

[14] Pusic AL, Klassen AF, Scott AM, Cano SJ. Development and psychometric evaluation of the FACE-Q satisfaction with appearance scale. Clin Plast Surg 2013;40:249−60.

[15] 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;133:21−30.

[16] Klassen AF, Cano SJ, Schwitzer JA, Scott AM, Pusic AL. FACE-Q scales for health-related quality of life, early life impact, satisfaction with outcomes, and decision to have treatment: development and validation. Plast Reconstr Surg 2015;135:375−86.

[17] Klassen AF, Cano SJ, Schwitzer JA, Baker SB, Carruthers A, Carruthers J, 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;152:443−51.

[18] Guillemin F, Bonbardier C, Beaton D. Cross-cultural adaptation of health related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417−32.

[19] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health 2005;8:94−104.

[20] Cogliandro A, Barone M, Persichetti P. Italian linguistic validation of the FACE-Q instrument. JAMA Facial Plast Surg 2017;19:336−7.

[21] Chang CH. Patient-reported outcomes measurement and management with innovative methodologies and technologies. Qual Life Res 2007;16(Suppl 1):157−66.

[22] Lim VP, Lincoln M, Chan YH, Onslow M. Stuttering in English-Mandarin bilingual speakers: the influence of language dominance on stuttering severity. J Speech Lang Hear Res 2008;51:1522−37.

[23] Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, 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;14:978−88.

[24] Leidy NK, Vernon M. Perspectives on patient-reported outcomes: content validity and qualitative research in a changing clinical trial environment. Pharmacoeconomics 2008;26:363−70.

[25] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34−42.

[26] Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013;22:1889−905.

[27] Lamping DL, Schroter S, Kurz X, Kahn SR, Abenhaim L. Evaluation of outcomes in chronic venous disorders of the leg: development of a scientifically rigorous, patient-reported measure of symptoms and quality of life. J Vasc Surg 2003;37:410−9.

[28] Clark LA, Watson D. Constructing validity: basic issues in objective scale development. Psychol Assess 1995;7:309−19.

[29] Streiner DL, Norman GR, Fulton C. Health measurement scales. 2nd ed. Oxford: Oxford University Press; 1995.

[30] Cohen RJ, Swerdlik ME. Psychological testing and assessment: an introduction to tests and measurement. 6th ed. New York: McGraw-Hill; 2005.

[31] Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.

[32] Wang R, Yang J, Guo K, Zhong A, Tong J, Xiong L, et al. Asian facelift technique refinement with high patient satisfaction: FACE-Q report. Ann Plast Surg 2018 May 24. https://doi.org/10.1097/SAP.0000000000001496 [Epub ahead of print].

[33] Chen B, Song H, Gao Q, Xu M, Wang J, Wang F, et al. Measuring satisfaction with appearance: validation of the FACE-Q scales for double-eyelid blepharoplasty with minor incision in young Asians- retrospective study of 200 cases. J Plast Reconstr Aesthet Surg 2017;70:1129−35.

[34] Emery MP, Perrier LL, Acquadro C. Patient-reported outcome and quality of life instruments database (PROQOLID): frequently asked questions. Health Qual Life Outcomes 2005;3:12.

[35] Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol 2015;68:435−41.

[36] Yaremchuk MJ. Facial skeletal reconstruction using porous polyethylene implants. Plast Reconstr Surg 2003;111:1818—27.

[37] Rojas YA, Sinnott C, Colasante C, Samas J, Reish RG. Facial implants: controversies and criticism. A comprehensive review of the current literature. Plast Reconstr Surg 2018;142:991—9.

[38] Choi BK, Goh RC, Moaveni Z, Lo LJ. Patient satisfaction after zygoma and mandible reduction surgery: an outcome assessment. J Plast Reconstr Aesthet Surg 2010;63:1260—4.

[39] Morris DE, Moaveni Z, Lo LJ. Aesthetic facial skeletal contouring in the Asian patient. Clin Plast Surg 2007;34:547—56.

[40] Kim YH, Cho BC, Lo LJ. Facial contouring surgery for asians. Semin Plast Surg 2009;23:22—31.

[41] Petersen C, Markiewicz MR, Miloro M. Is augmentation required to correct malar deficiency with maxillary advancement? J Oral Maxillofac Surg 2018;76:1283—90.

[42] Morris DE, Lo LJ, Margulis A. Pitfalls in orthognathic surgery: avoidance and management of complications. Clin Plast Surg 2007;34:e17—29.

[43] Worasakwutiphong S, Chuang YF, Chang HW, Lin HH, Lin PJ, Lo LJ. Nasal changes after orthognathic surgery for patients with prognathism and Class III malocclusion: analysis using three-dimensional photogrammetry. J Formos Med Assoc 2015;114:112—23.

[44] Jung J, Lee CH, Lee JW, Choi BJ. Three dimensional evaluation of soft tissue after orthognathic surgery. Head Face Med 2018;14:21.

[45] Kim SK, Kim HS. Secondary Asian rhinoplasty: lengthening the short nose. Aesthet Surg J 2013;33:353—62.

[46] Lee SH, Koo MG, Kang ET. Septal cartilage/ethmoid bone composite graft: a new and improved method for the correction underdeveloped nasal septum in patients with short noses. Aesthet Plast Surg 2017;41:388—94.

[47] Lee LW, Chen SH, Yu CC, Lo LJ, Lee SR, Chen YR. Stigma, body image, and quality of life in women seeking orthognathic surgery. Plast Reconstr Surg 2007;120:225—31.