**RESEARCH ARTICLE**

**OPEN ACCESS**

# Validation of the Arabic Version of the Orthognathic Quality of Life Questionnaire

Shoroog H. Agou[1] , Fahad F. Alsulaimani[1], Salma H. Ghoneim[1], Mohamed S. Bamashmous[2], Mohammed Y. Alghamdi[3], Lujain H. Mirdad[4], Maisa M. Marghalani[5], Ayman I. Alsanawi[6,7], Abdullah H. Aljoharji[6,8], Abdulrahman M. Idrees[9] and Razan M. Baabdullah[3,*]

[1]Department of Orthodontics, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[2]Department of Dental Public Health, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[3]Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[4]Department of Endodontics, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[5]Department of Periodontics, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[6]Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[7]Primary Dental Health Care Administration, Alahsa Health Cluster, Ministry of Health, Saudi Arabia
[8]Jeddah Second Health Cluster, Rabigh Hospital, Saudi Arabia
[9]Department of Oral Biology, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

**Abstract:**

*Introduction:* The OQLQ is the most widely used orthognathic patient-reported outcome measure, but an officially translated and validated Arabic version has not been available; prior Arabic use has been informal and untested. This study provides the first cross-cultural adaptation and psychometric evaluation (OQLQ-Ar) in Saudi adults.

*Methods:* Following international cross-cultural adaptation guidelines, the OQLQ underwent independent forward translation, back translation, and expert panel reconciliation. In a prospective, single-center cohort, patients self-administered the Arabic OQLQ (n = 74, 28 males, 46 females, aged 30 ± 4.6) at five treatment stages (pre-treatment, post-orthodontics, 1-week post-surgery, debonding, 1-year retention). Test–retest reliability was examined in 15 clinically stable participants who repeated the questionnaire after a 2-week interval. Responsiveness was assessed in a longitudinal subsample of 12 patients evaluated before and after surgery. Construct validity was explored through correlations with a visual-analogue scale (VAS) for overall satisfaction with facial appearance and surgery.

*Results:* Cronbach's α for the total scale was 0.92, indicating excellent internal consistency. Test–retest Spearman correlations (r) ranged from 0.60 (awareness of dentofacial aesthetics) to 0.88 (oral function), meeting the ≥0.70 reliability benchmark in three of four domains. Construct validity was supported by a weak-to-moderate positive correlation between total OQLQ score and VAS satisfaction (r = 0.34; *p* < 0.05). Statistically significant improvements were observed in the total score after surgery and the facial-aesthetics domain (median change = −24 points; *p* < 0.001).

*Discussion:* Psychometric performance of the Arabic OQLQ was strong: internal consistency was excellent, and domain test–retest was acceptable in most scales, though awareness was lower. Expected post-surgical improvement—greatest in facial aesthetics—demonstrates responsiveness, and correlations with VAS satisfaction support construct validity. These findings extend OQLQ evidence to Arabic-speaking adults with dentofacial deformity and justify broader validation work.

*Conclusion:* The cross-culturally adapted Arabic OQLQ shows sufficient preliminary measurement support for use in research and exploratory outcome tracking in orthognathic care. Larger multi-center studies should confirm factor structure, test longitudinal invariance, and evaluate integration with clinical severity indices (*e.g.*, IOFTN) before widespread clinical adoption.

**Keywords:** Arabic, Orthognathic quality of life questionnaire (OQLQ), Orthognathic surgery, Quality of life, Saudi Arabia.

*Address correspondence to this author at the Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, King Abdulaziz University, Al Ehtifalat St, Jeddah, 21589, PO box 80209, Saudi Arabia; E-mail: rmsbaabdullah@kau.edu.sa

## 1. INTRODUCTION

Orthognathic surgery is primarily performed to correct functional and aesthetic issues related to the jaw and dentofacial deformities, making it crucial to measure not only clinical outcomes but also the broader psychosocial effects of the surgery [1]. In both the United States and the United Kingdom, at least 5% of the general population is estimated to possess pronounced dentofacial deformities, often involving complex malocclusions, such that orthognathic surgery is clinically indicated [2]. Orthognathic surgery is not only a corrective procedure for skeletal discrepancies affecting occlusion and facial harmony, but can significantly improve quality of life through functional gains (*e.g.*, chewing, speech, airway optimization) and aesthetic enhancement [3]. These outcomes are best appreciated through comprehensive, patient-reported measures such as validated tools like the Orthognathic Quality of Life Questionnaire (OQLQ).

The OQLQ is a vital tool used to assess the impact of orthognathic surgery on a patient's quality of life. The OQLQ, developed by Cunningham *et al.* [4, 5], is a condition-specific instrument designed to assess the impact of dentofacial deformities and the benefits of orthognathic surgery on patients' quality of life. It consists of 22 questions divided into four clinically meaningful domains: Facial Esthetics, Oral Function, Awareness of Facial Deformities, and Social Aspects of the Deformity. Patients respond on a scale from 0 ("it does not bother me") to 4 ("it bothers me a lot"), with total scores ranging from 0 to 88; lower scores indicate better quality of life. The OQLQ has demonstrated good validity, reliability, and responsiveness, in its original English form and in subsequent translations, making it useful for both clinical trials and quality assurance in orthognathic treatment. It captures aesthetic, functional, psychological, and social dimensions relevant to patients undergoing orthognathic surgery [4, 5].

The OQLQ provides a structured, validated approach to understanding how surgical interventions affect patients' lives across the dimensions of facial aesthetics, oral function, social interaction, and self-confidence [6]. The OQLQ is important as it can also quantify subjective experiences, offering a comprehensive evaluation of treatment outcomes from the patient's perspective [7]. This is particularly significant in the context of orthognathic surgery, where the success of the procedure is often judged not solely by the surgeon's technical achievement but also by the extent to which the patient's quality of life is improved

[8]. By capturing data on factors such as self-confidence, facial aesthetics, and social interactions, the OQLQ helps clinicians and researchers identify the benefits and challenges faced by patients, guiding both clinical practice and future research [1].

The OQLQ was developed just over twenty years ago to measure QoL changes in patients with dentofacial deformities undergoing orthognathic surgery [4, 5], and it remains the only questionnaire for this purpose [9]. As a result, it has been used extensively in observational studies evaluating the impact of orthognathic surgery on the QoL of patients with dentofacial deformities, with a recent meta-analysis confirming the generally positive impact of the procedure [10]. The instrument has high validity and reliability [4, 5] and, given its utility in assessing patient-reported QoL outcomes after orthognathic surgery, it has been used worldwide after translation and validation in several languages, including Spanish [7, 11], Serbian [12], German [13], Portuguese [14, 15], Chinese [16], Dutch [17], and Swedish [18]. Validating the Arabic translation of the Orthognathic Quality of Life Questionnaire (Ar-OQLQ) is essential to ensure that the instrument accurately captures the experiences and quality of life of Arabic-speaking patients undergoing orthognathic surgery. Arabic versions of the OQLQ have also been used to assess QoL before or after surgery in patients in Middle Eastern and North African (MENA) countries [19-26]. However, none of these studies performed formal psychometric evaluations of the translated versions, namely assessing the reliability (how consistent the results are), the validity (whether the questionnaire measures what it is supposed to measure), or the responsiveness (ability to detect clinically important changes over time, even if small).

This study aimed to assess the reliability, validity, and responsiveness of an Arabic version of the OQLQ in a cohort of Saudi Arabian patients with malocclusions and dentofacial deformities.

## 2. MATERIALS AND METHODS

### 2.1. Translation Process

The OQLQ instrument comprises 22 items and is divided into four dimensions: social aspect of dentofacial deformity (eight items); facial aesthetics (five items); oral function (five items); and awareness of dentofacial aesthetics (four items) [4, 5]. The items are rated on a four-point Likert scale, and the total score ranges from 0 to 88. A higher score indicates a lower quality of life [4, 5].

The translation process used an established method for adapting health questionnaires [27, 28], where the questionnaire undergoes translation and back-translation by bilinguals, followed by consultation with professionals. The first translation into Arabic was carried out by an independent bilingual professional translator with Arabic as their first language, who was asked to maintain conceptual equivalence. Then, the first version of the Arabic translation was back-translated into English by another independent professional translator. The committee of professionals included the authors (orthodontists, oral and maxillofacial surgeons, a methodologist, and a dental public health professional), two dental interns, and both the forward and backward translators, who reviewed the translations and determined whether the translated and original versions achieved semantic, idiomatic, experiential, and conceptual equivalence. Any discrepancies were resolved, and members of the expert committee reached a consensus on all items used to generate a prefinal version of the translated questionnaire. For pilot testing, five volunteers were asked to respond to each questionnaire item and report any difficulty in interpreting or answering the questions. This was done to assess general comprehension and cultural appropriateness. Following any necessary changes after piloting, the final Arabic translation process was deemed adequate (Supplementary Table **S1**).

## 2.2. Study Design, Participants, and Ethical Approval

This study is reported according to the STROBE statement for cross-sectional studies [29]. All Arabic speaking adults (≥18 y) with moderate–severe malocclusions and/or dentofacial deformities receiving orthognathic care at a specialized private centre in Saudi Arabia were invited to participate. Exclusion criteria included syndromic craniofacial anomalies, prior jaw surgery outside the study treatment sequence, cognitive or literacy barriers preventing self-completion, or incomplete baseline questionnaire data. A clinically stable subsample repeated the questionnaire after 2 weeks for test–retest reliability; a surgical subsample was assessed pre and post surgery to evaluate responsiveness. This was a prospective, cross-sectional, longitudinal study of a non-probabilistic convenience sample of adult patients seeking orthognathic treatment. All Arabic speaking adults (≥18 y) with moderate–severe malocclusions and/or dentofacial deformities receiving orthognathic care at a specialized private center in Saudi Arabia were invited to participate. Exclusion criteria included syndromic craniofacial anomalies, prior jaw surgery outside the study treatment sequence, cognitive or literacy barriers preventing self-completion, or incomplete baseline questionnaire data. A clinically stable subsample repeated the questionnaire after 2 weeks for test–retest reliability; a surgical subsample was assessed pre and post surgery to evaluate responsiveness. All participants were fully informed of the study protocol and provided written informed consent.

Recruitment was carried out between January 2021 and December 2022 in Jeddah, Saudi Arabia. The sample size was determined based on the recommendation that scale validation studies should include at least 5 participants per questionnaire item. Given that the OQLQ contains 22 items, a minimum sample size of 110 participants was targeted.

However, due to recruitment challenges, the final sample included 74 participants, which still exceeded the minimum requirement of 3 participants per item (66 participants) suggested by some scale development experts.

Participants in the longitudinal sample (n=74) attended a private orthodontic practice for orthognathic treatment. Each patient was invited to complete a questionnaire [**Supplementary Table S1**.], that included basic information about the reasons for surgery, the OQLQ itself, four questions asking to what extent pain/discomfort, chewing, appearance, and speaking had been affected by the surgery, and a visual analog scale (VAS) to rate the overall satisfaction with orthognathic treatment. Patients were asked to complete the questionnaire at various times throughout the treatment journey: before starting treatment, after orthodontic treatment but before surgery, after surgery (at follow-up), at debonding, and after one year with retainers. A subgroup of patients (n=12) completed the questionnaire both before surgery (before treatment or after orthodontic treatment) and after surgery (after one year with retainers). To examine test-retest reliability, 15 patients attending King Abdulaziz University Hospital for orthognathic treatment also completed the same questionnaire twice at a two-week interval.

## 2.3. Statistical Analysis

Data were analyzed using SPSS v22 (IBM Statistics, Chicago, IL). The normality of continuous data was assessed with the Shapiro-Wilk test. Means (standard deviation, SD) and medians (interquartile range, IQR) are presented for quantitative data, while frequencies and percentages are presented for all qualitative data. The chi-square test was used to compare the proportions of two variables. The Mann-Whitney U test was used to compare QoL scores between two groups, and the Wilcoxon signed-rank test was used to compare QoL scores among three or more groups. Internal consistency was assessed in the entire study population using Cronbach's α coefficient, which is considered good if ≥ 0.70 (excellent if ≥ 0.90) [30]. Test-retest reliability was determined using Spearman's rank correlation coefficient (r) [30], as were correlations between satisfaction levels and QoL. A *p*-value <0.05 was considered statistically significant.

## 3. RESULTS

### 3.1. Baseline Characteristics of the Study Population

The baseline characteristics of the study population are shown in Table **1**. Of the 74 participants, 46 (62.2%) were female, and the average age was 30 ± 4.6 years (range 20 – 44). Most patients were seeking treatment due to concerns with both orofacial appearance and function, which necessitated a range of surgical solutions to one or both jaws (83.8% of cases) that most commonly included asymmetric mandibular setback (51.4%) and maxillary advancement (25.7%).

### 3.2. Reliability of the Arabic Version of the OQLQ

Seventy-four patients completed the Arabic version of the OQLQ either before or after surgery. Across the entire study population (n=74), Cronbach's α coefficient for inter-

nal consistency was 0.92, denoting "excellent" internal consistency.

**Table 1. Baseline characteristics of the study population (cross-sectional data, n=74).**

| Variable | - | Frequency | Percent |
|---|---|---|---|
| Sex | Male | 28 | 37.8 |
| | Female | 46 | 62.2 |
| Age (Years) | Mean ± SD (range) | 30 ± 4.6 (20-44) | |
| Angle's classification of deformity | Class I | 5 | 6.8 |
| | Class II | 13 | 17.6 |
| | Class III | 47 | 63.5 |
| Facial type | Mesocephalic | 24 | 32.4 |
| | Brachycephalic | 4 | 5.4 |
| | Dolichocephalic | 25 | 33.8 |
| Vertical | Deep bite | 11 | 14.9 |
| | Anterior overbite | 17 | 23 |
| | Posterior cross bite | 3 | 4.1 |
| | Vertical maxillary excess | 5 | 6.8 |
| | Multiple responses | 12 | 16.2 |
| Transverse | Unilateral | 5 | 6.8 |
| | Bilateral posterior cross bite | 18 | 24.3 |
| Surgical treatment details | Maxillary advancement | 19 | 25.7 |
| | Maxillary segmental setback | 1 | 1.4 |
| | Maxillary impaction | 16 | 21.6 |
| | Advancement and impaction | 18 | 24.3 |
| | Multiple responses | 8 | 5.4 |
| | Mandibular advancement | 20 | 27 |
| | Mandibular segmental setback | 5 | 6.8 |
| | Mandibular asymmetric setback | 38 | 51.4 |
| Jaws operated on | Maxilla | 1 | 1.4 |
| | Mandible | 2 | 2.7 |
| | Both | 62 | 83.8 |
| Reasons for surgery | Appearance | 15 | 20.3 |
| | Function | 17 | 23 |
| | Confidence | 3 | 4.1 |
| | Appearance and function | 16 | 21.6 |
| | Appearance and confidence | 1 | 1.4 |
| | Function and confidence | 2 | 2.7 |
| | All | 17 | 23 |

Fifteen individuals from a separate cohort of patients attending the University Hospital were invited to complete the questionnaire and repeat it two weeks later. Test-retest reliability assessed in these 15 individuals was between 0.601 and 0.879 (Table 2). All questions showed "excellent" test-retest reliability (Spearman's correlation (r) ≥ 0.75 [30]), apart from the awareness of dentofacial deformity, which was "good" (r = 0.601).

### 3.3. Validity of the Arabic Version of the OQLQ

Validity of the OQLQ was measured with respect to two variables that might be expected to be related to QoL in patients with dentofacial deformity undergoing orthognathic surgery: subjective effect on appearance after surgery (measured using a Likert scale) and overall satisfaction with surgery (measured with a VAS). The overall OQOL and

facial esthetics domain scores were weakly but positively correlated with VAS-recorded satisfaction with surgical treatment (Table 3). Further supporting the hypothesis that OQLQ is valid for measuring QoL in orthognathic surgery patients, overall OQLQ scores were significantly lower (better QoL) in patients who felt that their overall appearance was better after surgery (median 9, IQR 7-11) compared with those who felt that their appearance was only a little better (median 18.5, IQR 13.5-24) or a lot worse (median 15, IQR 5-35) after surgery ($p < 0.001$, Wilcoxon signed-rank test).

**Table 2. Test-retest reliability of the OQLQ (n=15, Spearman's rank correlation coefficient).**

| Domain | Correlation Coefficient | Significance |
|---|---|---|
| Overall | 0.859 | <0.001 |
| Social aspects of deformity | 0.849 | <0.001 |
| Facial esthetics | 0.849 | <0.001 |
| Oral function | 0.879 | <0.001 |
| Awareness of dentofacial deformity | 0.601 | 0.018 |

**Table 3. Correlation between VAS responses for satisfaction with surgical treatment and OQOL scores (n = 74, Spearman's rank correlation coefficient).**

| Domain | Correlation Coefficient |
|---|---|
| Overall | 0.34* |
| Social aspects of deformity | 0.35 |
| Facial esthetics | 0.437* |
| Oral function | 0.14 |
| Awareness of dentofacial deformity | -0.05 |

**Note:** *Correlation significant at the 0.05 level.

### 3.4. Responsiveness of the Arabic Version of the OQLQ

The responsiveness of the Arabic version of the OQLQ was examined by comparing scores before and after surgery in both the entire cohort (n=74, completed questionnaire either before or after surgery) and in a subset of patients who completed the questionnaire both before and after surgery (n=12) (Table 4). For both groups, overall OQLQ scores were significantly lower after surgery than before surgery ($p < 0.0001$ and p = 0.09, respectively). In addition, in the paired analysis, scores for the facial esthetics domain were significantly lower after surgery ($p=0.03$).

## 4. DISCUSSION

This study developed and formally validated an Arabic version of the OQLQ. Data collected from this sample of consecutive patients attending a private practice for orthognathic treatment suggested that the translated version was conceptually similar to the English version, offering initial evidence that the Ar-OQLQ was valid, reliable, and responsive.

**Table 4. Responsiveness of the Arabic version of the OQLQ.**

| - | All Data | | | Paired Data | | |
|---|---|---|---|---|---|---|
| | **Before Surgery, Median (IQR) (n=74)** | **After Surgery, Median (IQR) (n=74)** | ***p*-value** | **Before Surgery, Median (IQR) (n=12)** | **After Surgery, Median (IQR) (n=12)** | ***p*-value** |
| Overall | 34 (21 - 42) | 10 (4 - 28) | **<0.0001** | 22 (14.5 - 50.5) | 11.5 (9.5 - 21) | **0.09** |
| Social aspects of deformity | 5 (3 - 9) | 3 (1 - 7.5) | 0.279 | 5.5 (0 - 13) | 1.5 (0 - 6) | 0.236 |
| Facial esthetics | 8 (5 - 15) | 5 (2 - 9) | 0.159 | 9.5 (6 - 14.5) | 6 (2.5 - 8.5) | **0.03** |
| Oral function | 6 (2 - 11) | 4 (2 - 7) | 0.343 | 4.5 (3 - 10.5) | 2.5 (0.5 - 7.5) | 0.327 |
| Awareness of dentofacial deformity | 6 (2 - 11) | 6 (2 - 11) | 0.952 | 6 (1.5 - 9) | 3 (0.5 - 6) | 0.265 |

Although the included patients represented a range of clinical presentations and motivations for orthognathic surgery, generalization of these psychometric properties was constrained by the absence of factor-analytic confirmation and convenience sampling. Future investigations that remedy these methodological limitations will markedly enhance confidence in the instrument's measurement properties and cultural suitability.

As seen with the original English version and other translations [4, 5, 7, 11, 12, 14-18], the Arabic OQLQ showed strong reliability, both in terms of internal consistency (Cronbach α > 0.9) and test-retest reliability, which was excellent (r >0.75) overall and for three out of four domains. The "awareness of dentofacial deformity" domain showed a slightly lower reliability, likely due to the small sample size (n=15) of pretreatment patients who completed the questionnaire twice.

The construct validity was tested using two variables related to QoL in patients with dentofacial deformity undergoing orthognathic surgery: subjective effect on appearance after surgery and overall satisfaction with surgery. While the exact relationship between global measures of satisfaction with surgery and the instrument is difficult to predict, the constructs are sufficiently related to expect a correlation; indeed, this was confirmed. Both metrics were associated with the overall OQLQ, suggesting adequate construct validity of the Arabic OQLQ. The facial esthetics subdomain was positively and significantly correlated with satisfaction with surgical treatment, which might be expected since facial esthetics has a profound impact on quality of life, particularly self-esteem [31]. However, correlations between the "social aspects of deformity," "oral function", and "awareness of dentofacial deformity" subdomains and satisfaction with surgical treatment were not significant, mirroring the results of the original validity testing, which showed the lowest correlations between VAS scores and the "oral function", and "awareness of dentofacial deformity" subdomains [5]. Also, as with previous studies, the treatment showed an overall positive effect on QoL, with an average improvement of 24 OQLQ points, similar to the mean difference of a 20-point improvement reported in a previous meta-analysis, and therefore supporting the construct validity of the Ar-OQOL [10].

Psychometric performance is generally shaped by two related factors: cultural context, which governs how patients interpret and prioritize questionnaire items, and methodological design choices—sample size, follow-up

interval, and assessment schedule that determine the statistical behavior of those items. Compared with the Spanish validation by Duarte *et al*. [11], our Saudi-based Arabic version exhibited even stronger internal consistency (Cronbach's α = 0.92 *vs* 0.87), while both translations comfortably exceeded the ≥0.70 benchmark for test–retest reliability—though the Arabic confidence intervals were necessarily wider because of the smaller retest cohort. Each version recorded a significant overall gain in OQLQ scores after orthognathic surgery; however, only the Spanish study detected additional improvements in the social aspects and oral function domains. These discrepancies may reflect cultural expectations or the shorter 3-month postoperative interval used in the Spanish protocol. Notably, our multistage assessment schedule was most sensitive in capturing change within the facial aesthetics domain, underscoring subtle but important differences in responsiveness between the two cultural adaptations. This side-by-side comparison therefore highlights both the shared psychometric strengths and the context-specific nuances of the Arabic and Spanish OQLQ versions.

In addition, Saudi culture places a strong emphasis on harmonious facial aesthetics, with public perception studies indicating that visible dentofacial disharmony can negatively impact self-esteem, marriage prospects, and employment opportunities [19, 23, 26, 32-34]. This cultural premium on facial harmony likely explains why, in studies of Saudi patients undergoing orthognathic surgery, the facial-aesthetics domain of quality of life showed the greatest baseline impairment and the most significant post-operative improvement, while functional complaints were less prominent. Such findings highlight the psychosocial pressures related to appearance in this cultural context. Future research, such as through cognitive interviewing, could refine Arabic questionnaire items to better capture culturally specific constructs like "social acceptance", which are particularly salient in Saudi society. This nuanced understanding aligns with evidence from quality-of-life studies using the Arabic version of the OQLQ, which demonstrated reliable measurement of these culturally relevant impacts [19, 23, 26, 32-34].

Previous translations and validations of the OQLQ have failed to measure responsiveness (the ability to detect clinically important changes over time), which is the most important metric to measure if a patient-reported outcome instrument is to have clinical utility. Here, responsiveness was tested through group comparisons of patients completing the questionnaire either before or after surgery and

in a subset of patients who completed the questionnaire both before and after surgery (n=12). For both groups, overall OQLQ scores were significantly lower after surgery than before surgery, similar in magnitude to OQLQ decreases observed in meta-analysis (24 *vs.* 20 points) [10].

For the entire cohort, there were no significant decreases in OQLQ scores for any of the four subdomains, which may be due to the grouping of questionnaire responses at two time points before surgery (on assessment and after orthodontic treatment) and three time points after surgery (after surgery (at follow-up), on debonding, and one year with retainers). In the paired analysis, mirroring the validity analysis, the facial esthetics subdomain was significantly improved after surgery; given that patients are likely to focus on esthetics after undergoing a procedure to alter a facial deformity, this is perhaps unsurprising. Indeed, the responsiveness testing of the original questionnaire similarly found that not all subdomains showed favorable changes after surgery [5].

While the overall OQLQ scores improved markedly after surgery, oral function and awareness of dentofacial aesthetics did not show statistically significant changes, and several inter-related factors may explain this pattern. As we know, the relationship between the oral function determined by IOFTN and the dentofacial deformity discrepancy is not straight forward (3). The baseline responses in these domains clustered toward the favorable end of the scale, suggesting a ceiling effect that left little head-room for detectable postoperative gain. Patients often compensate functionally before surgery and, within the Saudi cultural context—where facial harmony is highly prized—may assign lower salience to chewing- or speech-related items; thus, functional improvements do not automatically translate into perceived quality-of-life gains unless accompanied by aesthetic benefits. Framed through the Wilson and Cleary HRQoL model [35], changes at the "functional status" tier (oral function) may not propagate to the higher-order "general health perception" tier unless they interact with culturally salient appearance outcomes. Methodological factors further compound this issue. Grouping responses across multiple postoperative stages, from one week after surgery to one-year retention, may have diluted time-specific effects, masking subtle functional or psychosocial changes. Future research should therefore (i) collect more granular, stage-specific data, (ii) extend the test–retest window, and (iii) employ qualitative cognitive interview studies to probe item relevance and comprehension. Such work could lead to recalibration or replacement of low-impact items and strengthen the oral function and awareness domains' ability to capture nuanced postoperative change. This inclusion of formal qualitative methods in future studies is likely to enhance the cultural adaptations for a more robust semantic validation.

The current results may also have been limited by the small number of patients in the paired analysis. Based on recommendations for psychometric validation studies requiring 5-10 subjects per item, with the 22-item OQLQ requiring a minimum of 110-220 participants, our achieved sample of 74 participants represents a limitation of the study that we acknowledge may affect our findings. As for

the longitudinal cohort (n=12), the small sample size is another limitation for assessing responsiveness. A post-hoc power calculation revealed this sample size would only detect large effects (Cohen's d > 0.8) with 80% power at α=0.05. Therefore, future studies should aim for at least 27 paired samples to detect medium effects. Regarding the test-retest cohort (n=15), our sample size was calculated based on detecting an ICC of 0.8 with a 95% CI of 0.2, requiring 15 participants. While meeting minimum requirements, a larger sample would provide more robust reliability estimates.

This validated version of the OQLQ is expected to improve future QoL studies in Arabic-speaking nations. Our responsiveness testing adds to a body of evidence from the MENA region showing that orthognathic treatment improves QoL. Previous studies from Saudi Arabia [19], Kuwait [20], Morocco [21], Egypt [25], and Jordan [22] all found that OQLQ total and sub-domain scores improved after orthognathic surgery. Other comparisons using the OQLQ in the MENA region included a report of higher OQLQ scores in Jordanian patients with post-surgical temporomandibular disorders [24]. There has also been a cross-cultural comparison of OQLQ scores between Jordanian and British patients [23], which revealed a significant difference for oral function (poorer QoL in the Jordanian cohort). Although the authors suggested that this difference could be attributed to cultural or systemic healthcare differences, it might also stem from a misunderstanding of the questionnaire, underscoring the value of the current study. Altogether, these studies indicate that, when used alongside the Index of Orthognathic Functional Treatment Need (IOFTN) [36], the Arabic OQLQ offers a complementary patient-centered metric for assessing dentofacial-deformity cases.

## 4.1. Study Limitations and Future Directions

Reliance on a single-center convenience sample limits generalizability across other Arabic-speaking populations. This study provides preliminary validation evidence for the Arabic version of the OQLQ, demonstrating good reliability and initial indications of responsiveness. However, several limitations should be acknowledged. As discussed above, the test-retest validation cohort (n=15) and the pre- and post-surgery paired cohorts (n=12) were both small, so the analyses may have been underpowered. Recruitment was constrained by the relatively low volume of orthognathic surgery candidates, attending a single private clinic, who met all inclusion criteria during the study window, an issue compounded by extended treatment timelines. Because of this limited convenience sampling, the potential influence of demographic variables such as age or sex or type of surgery on QoL outcomes could not be investigated, leading to limited external validity. Also, the recall interval time was relatively limited, potentially leading to recall bias. Accordingly, a multi-centre follow-up study to pool cases across institutions, recruit larger and more diverse samples (≥ 27 for paired analyses and 110–220 overall for 22-item validation), conduct both exploratory and confirmatory factor analyses, integrate convergent instruments such as the OHIP-14 [37] and Orofacial Esthetic Scale [39], extend the test–retest interval to 4–6 weeks, and stratify outcomes

by IOFTN grade and socio-economic status will help to overcome the present methodological constraints.

No comparator patient-reported outcomes (*e.g.*, OHIP-14, JFLS, OES) were used in this study to assess convergent or discriminant validity. Although construct validity was supported through correlations with surgical satisfaction and self-reported appearance changes, and instrument performance before and after surgery, further validation using instruments such as the OHIP-14 [37], Jaw Functional Limitation Scale (JFLS) [38], or Orofacial Esthetic Scale (OES) [39] is warranted to assess convergent and discriminant validity. Additionally, potential response bias, including social desirability effects and clinician-induced demand characteristics, may have influenced participant reporting given the clinical context of administration. To mitigate these sources of bias and strengthen methodological rigor, future research should employ anonymized, self-administered data collection protocols.

It is also worth noting that this study did not directly test comprehension of the questionnaire, instead inferring cultural appropriateness and understandability through analysis of the reliability and clinically correlated responsiveness data, which suggests that this Arabic questionnaire is fit for purpose. Similarly, the internal structure of the Arabic OQLQ was not tested using factor analysis, which limits our understanding of its dimensional integrity. Future research should include exploratory and confirmatory factor analysis to better establish structural validity. Similarly, the sub-threshold correlations we observed (r < 0.50), which fall below the range considered clinically meaningful in orthodontic outcomes research [40], underscore the need for future multi-center studies with larger, IOFTN-stratified samples to verify whether these associations persist in a more diverse population. Present findings are preliminary, and there is a need for stronger, clinically significant correlations. The validated tool can be integrated into routine surgical assessment protocols in Arabic-speaking regions.

Questionnaire administered in clinical settings may introduce social desirability or clinician-induced bias and an anonymous self-administration protocol is recommended. While cultural adaptation was assessed through pilot testing, formal qualitative methods, such as cognitive interviews or focus groups, could further strengthen evidence of semantic and experiential equivalence. Additionally, the use of convenience sampling from a single private clinic may limit generalizability. Employing probability or stratified sampling across diverse clinical settings would enhance external validity. Addressing these aspects will support the broader applicability and psychometric robustness of the Arabic OQLQ in both clinical and research settings.

Future studies should address the limitations of this study by recruiting larger and more diverse samples, having a longer (≥4-week) retest interval, conducting factor analysis and measurement invariance testing, and using validated instruments for a more thorough construct validation. Employing probability sampling methods, when possible, will also enhance the representativeness of future findings. These steps are necessary to ensure that the Arabic OQLQ is a reliable and valid tool for measuring QoL across Arabic-speaking populations.

## CONCLUSION

This study provides preliminary evidence supporting the reliability, validity, and responsiveness of an Arabic version of the OQLQ in a sample of patients from Saudi Arabia. The results suggest that the translated instrument may be conceptually comparable to the original English version. Further research involving larger, more diverse populations is needed to confirm these observations and to strengthen the construct validation of the Arabic OQLQ.

## AUTHORS' CONTRIBUTIONS

The authors confirm contribution to the paper as follows: S.A.: Conceptualization and Methodology; L.M., M.M.: Translation (Forward and Backward); F.S., A.H.A., A.I.A., A.I.: Data Collection; M.B.: Data Analysis; S.A., S.G.: Writing – Original Draft Preparation; M.G., S.G., R.B.: Writing – Review and Editing. All authors have read and agreed to the published version of the manuscript.

## LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| Arabic OQLQ | = | Orthognathic Quality of Life Questionnaire into Arabic |
| MENA | = | Middle Eastern and North African |
| VAS | = | visual analog scale |
| SD | = | Standard Deviation |
| IQR | = | Interquartile Range |
| JFLS | = | Jaw Functional Limitation Scale |
| OES | = | Orofacial Esthetic Scale |

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study protocol was reviewed and approved by the Institutional Review Board of King Abdulaziz University Dental Hospital, Jeddah, Saudi Arabia (reference number: 016-12).

## HUMAN AND ANIMAL RIGHTS

All human research procedures followed were in accordance with the ethical standards of the committee responsible for human experimentation (institutional and national), and with the Helsinki Declaration of 1975, as revised in 2013.

## CONSENT FOR PUBLICATION

All participants were fully informed of the study procedures and provided written informed consent prior to participation.

## STANDARDS OF REPORTING

STROBE guidelines were followed.

## AVAILABILITY OF DATA AND MATERIALS

All raw data are available from the corresponding author on reasonable request.

## CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

## SUPPLEMENTARY MATERIALS

Supplementary material is available on the Publisher's website.

## REFERENCES

[1]   Sun H, Shang HT, He LS, Ding MC, Su ZP, Shi YL. Assessing the quality of life in patients with dentofacial deformities before and after orthognathic surgery. J Oral Maxillofac Surg 2018; 76(10): 2192-201.
http://dx.doi.org/10.1016/j.joms.2018.03.026 PMID: 29684310

[2]   Posnick JC. Definition and prevalence of dentofacial deformities Orthognathic surgery: Principles and practice St Louis. Elsevier 2014; pp. 61-8.

[3]   Borzabadi-Farahani A, Olkun HK, Eslamian L, Eslamipour F. Functional needs in orthognathic patients with different sagittal skeletal discrepancies. Oral Surg Oral Med Oral Pathol Oral Radiol 2024; 138(2): 255-60.
http://dx.doi.org/10.1016/j.oooo.2024.04.006 PMID: 38749877

[4]   Cunningham SJ, Garratt AM, Hunt NP. Development of a condition-specific quality of life measure for patients with dentofacial deformity: I. Reliability of the instrument. Community Dent Oral Epidemiol 2000; 28(3): 195-201.
http://dx.doi.org/10.1034/j.1600-0528.2000.280305.x PMID: 10830646

[5]   Cunningham SJ, Garratt AM, Hunt NP. Development of a condition-specific quality of life measure for patients with dentofacial deformity: II. Validity and responsiveness testing. Community Dent Oral Epidemiol 2002; 30(2): 81-90.
http://dx.doi.org/10.1034/j.1600-0528.2002.300201.x PMID: 12000348

[6]   Ni J, Song S, Zhou N. Impact of surgical orthodontic treatment on quality of life in Chinese young adults with class III malocclusion: A longitudinal study. BMC Oral Health 2019; 19(1): 109.
http://dx.doi.org/10.1186/s12903-019-0782-9 PMID: 31196054

[7]   Sánchez-Burgos R, Martínez-Gimeno C, Arribas-García I, *et al.* Transcultural adaptation and validation of the Spanish language version of the questionnaire OQLQ for the assessment of quality of life in orthognathic patients. J Clin Exp Dent 2018; 10(12): e1184-91.
http://dx.doi.org/10.4317/jced.55366 PMID: 30697377

[8]   Bortoluzzi M-C, Manfro R, Soares I-C, Presta A-A. Cross-cultural adaptation of the orthognathic quality of life questionnaire (OQLQ) in a Brazilian sample of patients with dentofacial deformities. Med Oral Patol Oral Cir Bucal 2011; 16(5): e694-9.
http://dx.doi.org/10.4317/medoral.16938 PMID: 20711138

[9]   Kanatas AN, Rogers SN. A systematic review of patient self-completed questionnaires suitable for oral and maxillofacial surgery. Br J Oral Maxillofac Surg 2010; 48(8): 579-90.
http://dx.doi.org/10.1016/j.bjoms.2009.12.004 PMID: 20045231

[10]  Meger MN, Fatturi AL, Gerber JT, *et al.* Impact of orthognathic surgery on quality of life of patients with dentofacial deformity: A systematic review and meta-analysis. Br J Oral Maxillofac Surg 2021; 59(3): 265-71.
http://dx.doi.org/10.1016/j.bjoms.2020.08.014 PMID: 33546846

[11]  Duarte V, Zaror C, Villanueva J, *et al.* Cross-cultural adaptation and validation of the Spanish version of the Orthognathic Quality of Life Questionnaire for patients with dentofacial deformities. J Craniomaxillofac Surg 2020; 48(12): 1112-8.
http://dx.doi.org/10.1016/j.jcms.2020.10.004 PMID: 33221152

[12]  Vucic L, Glisic B, Kisic-Tepavcevic D, Vucic U, Drulovic J, Pekmezovic T. Cross-cultural adaptation and validation of the disease specific questionnaire OQLQ in Serbian patients with malocclusions. Zdr Varst 2016; 55(3): 166-73.
http://dx.doi.org/10.1515/sjph-2016-0021 PMID: 27703535

[13]  Bock JJ, Odemar F, Fuhrmann RA. Assessment of quality of life in patients undergoing orthognathic surgery. J Orofac Orthop 2009; 70(5): 407-19.
http://dx.doi.org/10.1007/s00056-009-9903-4 PMID: 19997999

[14]  de Araújo AM, Miguel JA, Gava EC, de Oliveira BH. Translation and cross-cultural adaptation of an instrument designed for the assessment of quality of life in orthognatic patients. Dental Press J Orthod 2013; 18(5): 99-106.
http://dx.doi.org/10.1590/s2176-94512013000500017 PMID: 24352395

[15]  Laureano Filho J, Ferreira P, Barros K, Dias E. Translation and validation for Portuguese (Brazil) specific instrument for assessing quality of life of patients undergoing orthognathic surgery: OQLQ. Int J Oral Maxillofac Surg 2011; 40(10): 1198.

[16]  Xue Z, Ma X, Liu X, Wang X, Li Z. The Chinese version of Orthognathic Quality of Life Questionnaire (OQLQ-C): Translation, reliability, and validity. Clin Oral Investig 2021; 25(3): 1497-503.
http://dx.doi.org/10.1007/s00784-020-03457-4 PMID: 32827277

[17]  Ploumen RLM, Duininck JM, Jonkman REG, Nolte JW, Becking AG. The orthognathic quality of life questionnaire: Translation and validation into Dutch. J Craniofac Surg 2021; 32(4): 1448-51.
http://dx.doi.org/10.1097/SCS.0000000000007281 PMID: 33252530

[18]  Ahl M, Marcusson A, Ulander M, Magnusson A, Cardemil C, Larsson P. Translation and validation of the english-language instrument orthognathic quality of life questionair into Swedish. Acta Odontol Scand 2021; 79(1): 19-24.
http://dx.doi.org/10.1080/00016357.2020.1768284 PMID: 32432962

[19]  Abdullah WA. Changes in quality of life after orthognathic surgery in Saudi patients. Saudi Dent J 2015; 27(3): 161-4.
http://dx.doi.org/10.1016/j.sdentj.2014.12.001 PMID: 26236131

[20]  Al-Asfour A, Waheedi M, Koshy S. Survey of patient experiences of orthognathic surgery: Health-related quality of life and satisfaction. Int J Oral Maxillofac Implants 2018; 47(6): 726-31.
http://dx.doi.org/10.1016/j.ijom.2017.12.010 PMID: 29373200

[21]  Elmouden L, Ousehal L. Assessment of the quality of life in moroccan patients undergoing orthognathic surgery. Turk J Orthod 2018; 31(3): 79-85.
http://dx.doi.org/10.5152/TurkJOrthod.2018.17040 PMID: 30206566

[22]  Al-Ahmad HT, Al-Sa'di WS, Al-Omari IK, Al-Bitar ZB. Condition-specific quality of life in Jordanian patients with dentofacial deformities: A comparison of generic and disease-specific measures. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009; 107(1): 49-55.
http://dx.doi.org/10.1016/j.tripleo.2008.05.040 PMID: 18718795

[23]  Al-Bitar ZB, Al-Omari IK, Al-Ahmad HT, El Maaytah MA, Cunningham SJ. A comparison of health-related quality of life between Jordanian and British orthognathic patients. Eur J Orthod 2009; 31(5): 485-9.
http://dx.doi.org/10.1093/ejo/cjp034 PMID: 19477973

[24]  Al-Ahmad HT, Al-Bitar ZB. The effect of temporomandibular disorders on condition-specific quality of life in patients with dentofacial deformities. Oral Surg Oral Med Oral Pathol Oral Radiol 2014; 117(3): 293-301.
http://dx.doi.org/10.1016/j.oooo.2013.11.493 PMID: 24528790

[25] Hanafy M, Abou-Elfetouh A, Mounir RM. Quality of life after different approaches of orthognathic surgery: A randomized controlled study. Minerva Stomatol 2019; 68(3): 112-7.
http://dx.doi.org/10.23736/S0026-4970.19.04227-4 PMID: 31014060

[26] Agou S. Changes in oral health-related quality of life before and after orthognathic surgery and predictors of outcome. International Journal Of Orthodontic Rehabilitation 2024; 14(3): 1-11.
http://dx.doi.org/10.56501/intjorthodrehabil.v15i3.1038

[27] Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 2017; 11(Suppl. 1): S80-9.
http://dx.doi.org/10.4103/sja.SJA_203_17 PMID: 28616007

[28] Ferrer M, Alonso J, Prieto L, *et al*. Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: The Spanish example. Eur Respir J 1996; 9(6): 1160-6.
http://dx.doi.org/10.1183/09031936.96.09061160 PMID: 8804932

[29] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. Lancet 2007; 370(9596): 1453-7.
http://dx.doi.org/10.1016/S0140-6736(07)61602-X PMID: 18064739

[30] Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60(1): 34-42.
http://dx.doi.org/10.1016/j.jclinepi.2006.03.012 PMID: 17161752

[31] Gavric A, Mirceta D, Jakobovic M, Pavlic A, Zrinski MT, Spalj S. Craniodentofacial characteristics, dental esthetics-related quality of life, and self-esteem. Am J Orthod Dentofacial Orthop 2015; 147(6): 711-8.
http://dx.doi.org/10.1016/j.ajodo.2015.01.027 PMID: 26038075

[32] Alsaeed S, Afrashtehfar KI, Alharbi MH, *et al*. Impact of orthodontic appliances on hiring prospects in Saudi Arabia: A cross-sectional study. Cureus 2023; 15(6): 40173.
http://dx.doi.org/10.7759/cureus.40173 PMID: 37431341

[33] Bamashmous M, Zahran M, Bushnag A, *et al*. Social anxiety and orthognathic surgery effect on oral health-related quality of life. Cureus 2023; 15(9): 45434.
http://dx.doi.org/10.7759/cureus.45434 PMID: 37859885

[34] Ghoneim S, Alahmadi N, Alsaggaf D, Alsulaimani F, Alansari R, Adel S. Quality of Life (QoL) changes after orthognathic surgery: Do they correlate with the quantum of hard and soft tissue change? Original Research. Int J Orthod Rehabil 2024; 15(2): 1-12.
http://dx.doi.org/10.56501/intjorthodrehabil.v15i2.1012

[35] Sousa KH, Kwok OM. Putting Wilson and Cleary to the test: Analysis of a HRQOL conceptual model using structural equation modeling. Qual Life Res 2006; 15(4): 725-37.
http://dx.doi.org/10.1007/s11136-005-3975-4 PMID: 16688505

[36] Borzabadi-Farahani A. Systematic review and meta-analysis of the index of orthognathic functional treatment need for detecting subjects with great need for orthognathic surgery. Cleft Palate Craniofac J 2025; 62(3): 519-26.
http://dx.doi.org/10.1177/10556656231216833 PMID: 38037271

[37] Slade GD. Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol 1997; 25(4): 284-90.
http://dx.doi.org/10.1111/j.1600-0528.1997.tb00941.x PMID: 9332805

[38] Ohrbach R, Larsson P, List T. The jaw functional limitation scale: Development, reliability, and validity of 8-item and 20-item versions. J Orofac Pain 2008; 22(3): 219-30.
PMID: 18780535

[39] Larsson P, John MT, Nilner K, Bondemark L, List T. Development of an Orofacial Esthetic Scale in prosthodontic patients. Int J Prosthodont 2010; 23(3): 249-56.
PMID: 20552092

[40] Borzabadi-Farahani A, Borzabadi-Farahani A, Eslampour F. The relationship between the ICON index and the dental and aesthetic components of the IOTN index. World J Orthod 2010; 11(1): 43-8.
PMID: 20209176