Validation and clinical application of the Arabic rhinoplasty outcomes evaluation questionnaire

Dhaifallah S. Mulafikh, MD, *Sami E. Alharethy*, MD, PhD, *Almaha A. Alqabbani*, MD, *Tamer A. Mesallam*, MD, PhD.

ABSTRACT

الأهداف: ترجمة وتطبيق واختبار صلاحية استبيان تقييم مظهر الأنف باللغة العربية واستعماله كأداة لتقييم مظهر الأنف بعد عمليات تجميل الأنف.

المنهجية: تمت هذه الدراسة التقدمية في قسم الأنف والأذن والحنجرة بالمدينة الطبية في جامعة الملك سعود، حيث تضمنت إدراج 50 مريض ومقارنتهم ب89شخص من عامة الناس وعمل بعض الاختبارات الإحصائية للتحقق من صلاحية وفاعلية الاستبيان.

النتائج: أظهرت الدراسة القدرة على التفريق بين مجموعة المرضى ومجموعة المقارنة، حيث أن غالبية المرضى حققوا درجات عالية من الرضى عن مظهر الأنف بعد عملية تجميل الأنف.

الخلاصة: النسخة العربية المترجمة من استبيان تقييم مظهر الأنف هي أداة تقييم صالحة وفعالة ويمكن تطبيقها بسهوله لتقييم مستوى الرضى عن مظهر الأنف بعد عملية التجميل في الدول الناطقة باللغة العربية.

Objectives: To carry out translation, cross-cultural adaptation, and validation of the rhinoplasty outcome evaluation (ROE) into Arabic.

Methods: This non-randomized, prospective study included 50 patients from the Otorhinolaryngology Department, King Saud University, Riyadh, Saudi Arabia who underwent primary rhinoplasty from January to October 2020 and a control group of 89 healthy individuals. The reliability of the Arabic (ROE) was tested using Cronbach's alpha. The test-retest reliability was assessed by estimating the intra-class correlation coefficient for the total Arabic (ROE) score and individual items. Discriminant validity was used to examine the validity of the Arabic (ROE) by comparing the scores of the patients and the control group. The Friedman test was used to measure differences in Arabic (ROE) ratings within the study patients' group, including preoperative, and 2 weeks, and 3 months postoperative ratings.

Results: The internal consistency and reliability of the Arabic (ROE) were good. There was a significant difference in ratings between rhinoplasty patients and the control group regarding both the individual questions and total scores. Significant improvement was observed in the patient group ratings at 2 weeks and 3 months postoperatively compared to the preoperative rating ($p<0.0001$).

Conclusions: The Arabic version of the (ROE) showed good reliability and validity and can be used in the assessment of rhinoplasty outcomes in the Arabic population.

Keywords: Arabic, patient satisfaction, patient selection, quality of life, rhinoplasty, rhinoplasty outcomes evaluation

Rhinoplasty is very popular aesthetic surgery performed by otorhinolaryngologists and plastic surgeons.[1] It is a challenging and complex surgery because it is designed according to the unique requirements of each patient. However, the outcome assessment for rhinoplasty has not yet been agreed upon globally. The surgical method, approach, complications, and rates of revision surgeries have all gained a lot of attention; however, little emphasis has been placed on assessing the patient's view of the final rhinoplasty result.[1] The physician and patient may differ on what

they think is a good outcome.[2] To address this concern, Alsarraf developed 4 new outcome instruments to assess the results of aesthetic facial surgeries (because the current instruments were not related to rhinoplasty and evaluated blepharoplasty, face lift, and skin resurfacing). One of these tools is the rhinoplasty outcomes evaluation (ROE). The ROE focuses on 3 key factors that represent patient satisfaction, namely physical, emotional, and social factors, that are evaluated in order to measure outcomes. Each aspect is defined by 2 questions, for a total of 6 questions.[3,4] The ROE is broadly implemented and has been translated from the original English into various different languages, including German, Brazilian-Portuguese, and Turkish.[5-7] The ROE is a tool that remains in use to assess the impaction of rhinoplasty on patients' quality of life,[8-11] and it guides the preoperative selection of appropriate patients for surgery.[12] The main goal of this paper was to do translation, cross-cultural adaptation, and validation of the ROE into Arabic.

Methods. Our institutional review board approved the study protocol, and the author who first described the ROE granted permission. All patients who participated in the study provided written informed consent. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. We forward and backward translated the ROE-Arabic version (ROE-A) based on internationally accepted guidelines (**Figure 1**). The ROE comprises 6 questions. Each answer is scored on a scale from 0-4, where 0 corresponds to the "most negative answer" and 4 corresponds to the "most positive answer." As a result, the total score can range from 0 to 24. To make the results easier to understand, the total score can be divided by 24 and multiplied by 100 to obtain a value between 0% and 100%, where higher values represent greater patient satisfaction.

We applied the ROE-A in a non-randomized, prospective study conducted in the Otorhinolaryngology Department, King Saud University, Riyadh, Saudi Arabia from January to October 2020. We included all patients who were >18 years of age and underwent

a primary rhinoplasty ± septoplasty ± turbinoplasty. Exclusion criteria included patients undergoing revision rhinoplasty and patients who were not willing to participate in the study. The control group comprised healthy individuals with no desire for aesthetic or functional nasal procedures. This group comprised hospital employees and their relatives and friends.

Patients in the rhinoplasty group took the self-administered ROE-A questionnaire 4 times in our clinic. We administered the ROE-A twice preoperatively, with a gap of 2 weeks in between, to check the test-retest reliability, and twice postoperatively, 2 weeks and 3 months after surgery.

Statistical analysis. Normality test was conducted to explore and check the normal distribution of the data. The study data was not normally distributed according to the results of the data exploration. Accordingly, non-parametric tests were used for statistical analysis. The Mann-Whitney and Wilcoxon tests were used to compare the ORE-A results between groups with a level of significance of 95%.

A *p*-value <0.05 was considered statistically significant. Cronbach's alpha was used as a measure of the internal consistency of the ROE-A. A value between 0.7 and 0.8 was considered "satisfactory," a value between 0.8 and 0.9 was considered "good," and a value >0.9 was considered "excellent."[13] The intra-class correlation coefficient (ICC) was used to assess the test-retest reliability of the total ROE-A score and individual items. Discriminant validity was used to examine the validity of the ROE-A by comparing the scores of the

Figure 1 - Arabic version of rhinoplasty outcome evaluation questionnaire.

patients and the control group. Comparing the surgery outcomes over repeated measures after follow up was conducted using Friedman test to evaluate the ROE rating within the study subjects' groups, including preoperative, 2 weeks, and 3 months post-operative ratings. The Statistical Package for Social Sciences, Version 17 (SPSS Inc., Chicago, IL, USA) was used for all statistical analyses.

Results. The rhinoplasty group (n=50) included patients who had rated the ROE-A preoperatively and at 2 weeks and 3 months postoperatively. The control group (n=89) also rated the ROE-A (at the same time points). **Table 1** shows the age and gender distribution among the study and control subjects. There was no significant difference between the patient and control groups regarding age. However, there was a significant difference regarding the gender distribution between the 2 groups. Despite this difference in gender distribution, it had no association with the subjects' ROE ratings in either group, as there was no significant difference between men and women for the mean values of the

ROE-A ratings for either the patient group (*p*=0.87) nor the control group (*p*=0.39).

Reliability of testing. The overall results of the internal consistency of the ROE-A were good, as shown by a Cronbach's alpha value of 0.78. In addition, the Cronbach's alpha value maintained satisfactory levels ranging from 0.71 to 0.8 even if a question was deleted, with significant item-total correlation coefficient values ranging from 0.3 to 0.88 (**Table 2**).

We also assessed test-retest reliability by comparing 2 preoperative ROE-A ratings given by the patients with a gap of 2 weeks in between. The ICC of the individual questions and the total score of the ROE-A was 0.8, indicating good reliability (**Table 3**).

Validity testing. **Table 4** shows the discriminant validity of the ROE-A. We used the Mann-Whitney test to compare the ROE-A ratings of rhinoplasty

Table 1 - Demographic characteristics of the study group.

Characteristics	Patient group (n=50)	Control group (n=89)	*P*-value
Number	50	89	
Age (years)			
Mean (SD)	26.70 (7.33)	28.33 (7.86)	0.23
Range	17-47	16-44	
Gender, n (%)			
Female	21 (42)	75 (84)	0.001
Male	29 (58)	14 (16)	

Table 2 - Internal consistency results of the arabic rhinoplasty outcomes evaluation questionnaire.

ROE question	Cronbach's Alpha (if item was deleted)	Item-total correlation coefficient
Q1	0.73	0.88
Q2	0.80	0.30
Q3	0.74	0.90
Q4	0.78	0.59
Q5	0.73	0.86
Q6	0.71	0.88
Total	**0.84**	**1.00**

ROE: rhinoplasty outcomes evaluation

Table 3 - Test-retest reliability of the Arabic rhinoplasty outcomes evaluation.

ROE question	Test-retest correlation	Intra-class correlation coefficient	*P*-value
Q1	0.95		
Q2	1.00		
Q3	0.98		
Q4	0.97	0.80	0.0001
Q5	1.00		
Q6	1.00		
Total	**0.98**		

ROE: rhinoplasty outcomes evaluation

Table 4 - Discriminant validity of the rhinoplasty outcomes evaluation questionnaire between the rhinoplasty patient and control groups.

ROE question	Rhinoplasty patient group	Control group	P-value
Q1	0.46 (0.57)	2.74 (1.08)	0.0001
Q2	1.6 (1.22)	3.54 (0.71)	0.0001
Q3	1.24 (1.07)	2.76 (0.96)	0.0001
Q4	3.66 (0.71)	3.91 (0.32)	0.005
Q5	0.32 (0.65)	3.07 (1.18)	0.0001
Q6	0.12 (0.43)	3.78 (0.41)	0.0001
Total	**7.42 (2.28)**	**19.78 (3.26)**	**0.0001**

Values are presented as mean±SD,
ROE: rhinoplasty outcomes evaluation

Table 5 - Comparison between the preoperative and 2 weeks and 3 months postoperative assessment ratings of the rhinoplasty outcomes evaluation questionnaire.

ROE scale	Mean (SD)	*P*-value
Preoperative	7.42 (2.28)	
2 weeks postoperative	15.08 (4.28)	0.0001
3 months postoperative	20.36 (4.42)	0.0001
2 weeks postoperative	15.08 (4.28)	
Preoperative	7.42 (2.28)	0.0001
3 months postoperative	20.36 (4.42)	0.0001
3 months postoperative	20.36 (4.42)	
Preoperative	7.42 (2.28)	0.0001
2 weeks postoperative	15.08 (4.28)	0.0001

ROE: rhinoplasty outcomes evaluation

patients and the control group. There was a significant difference in the ratings of the rhinoplasty patients and those of the control group regarding both the individual questions and the total scores (P value ranging from 0.005 to 0.0001). This indicated that the ROE-A can differentiate between the responses of patients and controls.

We measured the responsiveness rate of the ROE-A by comparing the preoperative rating to the ratings at 2 weeks and 3 months postoperatively. The Friedman test and post-hoc tests were used to compare the mean values of these ratings. As shown in **Table 5,** there was a significant difference between the preoperative and postoperative ratings. We observed significant improvements in the patient group ratings at 2 weeks and 3 months postoperatively compared to the preoperative rating ($p<0.0001$). Similarly, significantly higher ROE-A scores were reported by patients 3 months postoperatively compared to 2 weeks postoperatively ($p=0.0001$).

Discussion. Rhinoplasty is very popular aesthetic surgery worldwide.[1] Researchers have developed various instruments to evaluate pre- and post-surgery outcomes.[14,15] The ROE is popular worldwide because it is simple, short, and easy to understand.[4] Furthermore, the primary indication for rhinoplasty, the surgical approach, and differences in gender and age had no influence on the postoperative satisfaction score using ROE.[8,15-17] Based on reports from the literature, rhinoplasty is one of the lowest rated surgeries for patient satisfaction among other aesthetic surgeries.[18] Therefore, appropriate preoperative patient selection is extremely important to achieve a good outcome. Izu et al[12] developed the normality parameter and a cut-off

point of 12 for the ROE to signify a good candidate for surgery and to help predict the outcome.

The ROE-A showed good reliability in internal consistency, as demonstrated with a Cronbach's alpha value of 0.78. In addition, test-retest reliability showed an ICC value of 0.8 for the individual questions and total score, indicating good reliability. As demonstrated by discriminant validity, the ROE-A differentiated between rhinoplasty patients and control subjects. The results of the Mann-Whitney test indicated good validity. The validity and reliability of the ROE-A were comparable to those published previously in related studies.[5-7] There was a significant difference between preoperative (7.42 ± 2.28) and postoperative ratings (20.36 ± 4.42), with a responsive rate of 55%, which is comparable to that of previously published studies,[5-7] but higher than that reported for the original ROE (44%).[4]

Question 2 (**Table 2**) had a low item-total correlation coefficient because it was the only question addressing function, and most of our patients had purely aesthetic concerns. However, all questions had good postoperative scores, indicating that even if the primary aim was purely aesthetic, concentration should be given to obstructive elements during rhinoplasty because this would also lead to a better quality of life.[6]

The patient ratings significantly improved at 2 weeks and 3 months postoperatively compared to the preoperative ratings ($p<0.0001$). Similarly, the patients reported significantly higher ROE-A scores 3 months postoperatively compared to 2 weeks postoperatively ($p<0.0001$), which could be explained by the resolution of edema and swelling of the nose post-surgery, compared with the scores reported by Izu et al.[6] Therefore, these results indicate the sensitivity of the ROE-A in detecting any changes in a patient's response even as early as 2 weeks after surgery.

Many studies have used the ROE to demonstrate rhinoplasty outcomes and its tremendous effect on quality of life.[8-11] For that reason, follow-up is crucial to help assess the outcomes and long-term patient satisfaction rate. Follow-up periods vary in the literature.[10,11,19] However, Arima et al[1] had a follow-up time period ranging from 6 months to 10 years and showed no significant statistical changes in the quality of life correlated to the length of the follow-up period, and they stated that less change was observed after 6 months. In addition, a systematic review conducted by Yang et al[9] concluded that patient satisfaction, as assessed by the ROE scale, is greatly enhanced after aesthetic-functional rhinoplasty for 12 months and beyond. However, the main aim of this study was to

test the validity and reliability of the Arabic version of the ROE in Arabic communities. Therefore, we did not focus on evaluating post-rhinoplasty quality of life.

Study limitations. The present study had some limitations, including the relatively small sample size and the short follow-up period. However, the main aim of the study was to test the validity and reliability of the ROE-A and not to measure patients' satisfaction following surgery.

Developing an Arabic version of the ROE provides a valid tool for surgeons and patients in Arabic-speaking countries to measure the outcomes of rhinoplasty procedures. The tool is easy to be administered and can be used in daily clinical practice.

In conclusion, our ROE-A showed good reliability and validity, and its results were comparable to those of previously translated, adapted, and published ROE questionnaires in the literature as well as the original ROE. In addition, as the ROE has been proven to be the most useful tool in evaluating rhinoplasty outcomes, the ROE-A can be used in the assessment of rhinoplasty outcomes in the Arabic population.

References

1. Arima LM, Velasco LC, Tiago RSL. Crooked nose: outcome evaluations in rhinoplasty. *Braz J Otorhinolaryngol* 2011; 77: 510-515.
2. Moolenburgh SE, Mureau MAM, Hofer SOP. Aesthetic outcome after nasal reconstruction: patient versus panel perception. *J Plast Reconstr Aesthet Surg* 2008; 61: 1459-1464.
3. Alsarraf R. Outcomes research in facial plastic surgery: A review and new directions. *Aesthetic Plast Surg* 2000; 23: 192-197.
4. Alsarraf R, Larrabee Jr WF, Anderson S, Murakami CS, Johnson Jr CM. Measuring cosmetic facial plastic surgery outcomes: a pilot study. *Arch Facial Plast Surg* 2001; 3: 198-201.
5. Bulut OC, Plinkert PK, Wallner F, Baumann I. Quality of life in functional rhinoplasty: rhinoplasty outcomes evaluation German version (ROE-D). *Eur Arch Otorhinolaryngol* 2016; 273: 2569-2573.
6. Izu SC, Kosugi EM, Lopes AS, Brandão KV, Sousa LBG, Suguri VM, et al. Validation of the Rhinoplasty Outcomes Evaluation (ROE) questionnaire adapted to Brazilian Portuguese. *Qual Life Res* 2014; 23: 953-958.
7. Çelik M, Altıntaş A. The Turkish version of the rhinoplasty outcomes evaluation questionnaire: validation and clinical application. *Balkan Med J* 2019; 36: 129-133.
8. Esteves SS, Ferreira MG, Almeida JC, Abrunhosa J, Sousa CAE. Evaluation of aesthetic and functional outcomes in rhinoplasty surgery: a prospective study. *Braz J Otorhinolaryngol* 2017; 83: 552-557.
9. Yang F, Liu Y, Xiao H, Li Y, Cun H, Zhao Y. Evaluation of preoperative and postoperative patient satisfaction and quality of life in patients undergoing rhinoplasty: a systematic review and meta-analysis. *Plast Reconstr Surg* 2018; 141: 603-611.
10. AlHarethy S, Al-Angari SS, Syouri F, Islam T, Jang YJ. Assessment of satisfaction based on age and gender in functional and aesthetic rhinoplasty. *Eur Arch Otorhinolaryngol* 2017; 274: 2809-2812.
11. Hellings PW, Trenité GJN. Long-term patient satisfaction after revision rhinoplasty. *Laryngoscope* 2007; 117: 985-989.
12. Izu SC, Kosugi EM, Brandão KV, Lopes AS, Garcia LBS, Suguri VM, et al. Normal values for the Rhinoplasty Outcome Evaluation (ROE) questionnaire. *Braz J Otorhinolaryngol* 2012; 78: 76-79.
13. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53-55.
14. Stewart MG, Witsell DL, Smith TL, Weaver EM, Yueh B, Hannley MT. Development and validation of the Nasal Obstruction Symptom Evaluation (NOSE) scale. *Otolaryngol Head Neck Surg* 2004; 130: 157-163.
15. Rhee JS. Measuring outcomes in nasal surgery: realities and possibilities. *Arch Facial Plast Surg* 2009; 11: 416-419.
16. Meningaud JP, Lantieri L, Bertrand JC. Rhinoplasty: an outcome research. *Plast Reconstr Surg* 2008; 121: 251-257.
17. Kütük SG, Arıkan OK. Evaluation of the effects of open and closed rhinoplasty on the psychosocial stress level and quality of life of rhinoplasty patients. *Plast Reconstr Aesthet Surg* 2019; 72: 1347-1354.
18. Saleh AM, Younes A, Friedman O. Cosmetics and function: quality-of-life changes after rhinoplasty surgery. *Laryngoscope* 2012; 122: 254-259.
19. Biggs TC, Fraser LR, Ward MJ, Sunkaraneni VS, Harries PG, Salib RJ. Patient reported outcome measures in septorhinoplasty surgery. *Ann R Coll Surg Engl* 2015; 97: 63-65.