original article

Psychometric Arabic Sino-Nasal Outcome Test-22: validation and translation in chronic rhinosinusitis patients

Fatma Alanazy,[a] Surayie Al Dousary,[a] Ahmed Albosaily,[b] Turki Aldriweesh,[c] Saad Alsaleh,[a] Turki Aldrees[c]

From the [a]Rhinology Chair Director in King Saud University, Department of Otolaryngology, Head and Neck Surgery, King Saud University College of Medicine, Riyadh, Saudi Arabia; [b]Department of Otolaryngology, Head and Neck Surgery, Prince Mohammed bin Abdulaziz Hospital, Riyadh, Saudi Arabia; [c]Department of Otolaryngology, Head and Neck Surgery, Prince Sattam bin Abdulaziz University College of Medicine, Al Kharj, Saudi Arabia

Correspondence: Turki Mohammed Aldrees · Department of Otolaryngology, Head and Neck Surgery, Prince Sattam bin Abdulaziz University College of Medicine, Al Kharj, 173 Saudi Arabia · T: +966-583953446 · pt.turki@hotmail.com · ORCID: http://orcid.org/0000-0002-8955-1778

Ann Saudi Med 2018; 38(1): 502-507

DOI: 10.5144/0256-4947.2018.502

BACKGROUND: The Sino-Nasal Outcome Test (SNOT)-22 has multiple items that reflect how nasal disease affects quality of life. Currently, no validated Arabic version of the SNOT-22 is available.
OBJECTIVE:. To develop an Arabic-validated version of SNOT-22.
DESIGN: Prospective.
SETTING: Tertiary care center.
PATIENT AND METHODS: This single-center validation study was conducted between 2015 and 2017 at King Abdul-Aziz University Hospital, Riyadh, Saudi Arabia. The SNOT-22 English version was translated into Arabic by the forward and backward method. The test and retest reliability, internal consistency, responsiveness to surgical treatment, discriminant validity, sensitivity and specificity all were tested.
MAIN OUTCOME MEASURES: Validated Arabic version of the SNOT-22.
RESULTS: Of 265 individuals, 171 were healthy volunteers and 94 were chronic rhinosinusitis patients. The Arabic version showed high internal consistency (Cronbach's of 0.94), and the ability to differentiate between diseased and healthy volunteers ($P<.001$). The translated versions demonstrated the ability to detect the change scores significantly in response to intervention ($P<.001$).
CONCLUSION: This is the first validated Arabic version of SNOT-22. The instrument can be used among the Arabic population.
LIMITATIONS: No subjects from other Arab countries.

Chronic rhinosinusitis (CRS) is a common disease worldwide.[1] In the United States, it accounts for millions of visits to the doctor annually. It is defined as mucosal inflammation of the nose and paranasal sinus for more than 12 weeks, and it can occur with and without polyps. Multiple guidelines have been published in the literature on the diagnosis of the disease.[2] One of the widely accepted guidelines worldwide is the European Position Paper on Rhinosinusitis and Nasal Polyposis (EPOS), the most recent version of which was published in 2012.[2] Those guidelines propose criteria including at least two of the following symptoms: nasal obstruction, nasal discharge, reduced sense of smell, and facial pain in conjunction with endoscopic or imaging findings.

CRS has a major impact on patient quality of life (QoL), which may lead the patient to seek medical or surgical treatment.[3] Multiple patient-reported outcome measures have been documented, including the Sino-Nasal Outcome Test (SNOT). The SNOT has multiple items that reflect how nasal disease affects QoL. It was first developed by Anderson et al[4] in 1998. A widely accepted version of the SNOT that is now used, the SNOT-22, contains 22 items reflecting various elements—rhino-logical symptoms, ear and facial symptoms, sleep functioning, and psychological factors—

that may be associated with nasal and paranasal disease.[5] Its validity has been confirmed in multiple studies and in multiple languages, including English, Greek, Lithuanian, Brazilian, Portuguese, Czech, Hebrew and Danish.[5-12] Although Arabic is one of the most widely spoken languages worldwide, limited data have been published about CRS outcome measures in Arabic speaking patients. In the current study, an Arabic version of the SNOT-22 was devised and tested, with the aim of developing a validated tool to measure patient-reported CRS outcomes in Arabic.

PATIENTS AND METHODS

This single-center prospective validation study was conducted at King Abdul-Aziz University Hospital, Riyadh, Saudi Arabia between 2015 and 2017. The patients were recruited consecutively from the rhinology clinic and inpatients ward prior to functional endoscopic sinus surgery (FESS). We included patients aged ≥18 years who met the EPOS criteria for CRS. We excluded patients who refused to participate, who were missing more than 20% of the data from the Arabic version of the SNOT-22, who could not read or write, and patients with mental conditions that may have impeded their ability to follow the questionnaire instructions. The study was approved by the Research Ethics Committee of King Saud University, and informed consent was obtained from all individual participants included in the study.

The English version of the SNOT-22 was translated into Arabic by two independent translators. These Arabic versions were then checked by two rhinologists, and compared to the original English version. The two Arabic versions were merged into one version, and this version was back-translated into English by another independent translator who was not aware of the original English version of the SNOT-22. There were no major differences between the back-translated version and the original English version. The final Arabic version was used in a pilot study in 10 patients to assess its clarity.

Test-retest reliability was performed in a sample of 10 patients preoperatively within a period of 1 week and not less than 2 days, with stable symptoms and without any de-escalation or escalation of their medications.

Preoperative and postoperative evaluations were conducted in a group of patients who underwent FESS. These patients' symptoms were evaluated via the Arabic version of the SNOT-22 twice, the first time 2 days prior to surgery and a second time within 9 months and not less than 1 month after the surgery. The control group consisted of healthy adult volunteers without CRS symptoms, who were accompanying patients during their visit to the rhinology clinic visit. All controls were subjected to an endonasal scope assessment.

Results are presented as means and standard deviation (SD) for quantitative data, and frequencies and percentages for qualitative data. The normality of distributions was assessed via the Shapiro-Wilk test. The internal consistency of the Arabic version of the SNOT-22 was assessed via Cronbach's α. Test-retest reliability was assessed via Spearman's test. Correlation coefficients were rated as excellent (≥0.91), good (0.71–0.90), moderate (0.51–0.70), acceptable (0.31–0.50), or low (≤0.30). The responsiveness of the Arabic version of the SNOT-22 was assessed by comparing preoperative and postoperative mean scores via the Wilcoxon signed-rank test and assessing the effect size—the mean difference in score divided by baseline SD. An effect size can be small (0.2), moderate (0.5), or large (0.8). In all analyses, $P<.05$ (5%) was considered statistically significant. The minimally important difference (MID) was estimated statistically as a change of more than or equal to half standard deviation (SD) of the baseline SNOT-22 score.

Discriminant validity between controls and CRS patients was assessed via the Mann-Whitney U test and a receiver operating characteristic (ROC) curve. The area under the curve (AUC) was defined as ROC-AUC. The evaluation standard used was 0.90 to 1.00=excellent, 0.80 to 0.89=good, 0.70 to 0.79= fair, and 0.50 to 0.59= fail. All statistical tests were performed using IBM SPSS Version 16.0.

RESULTS

Of 265 individuals, there were of 171 healthy volunteers (control group) and 94 CRS patients. The mean ages were 36 (14.3) years in the control group and 38.9 (13.8) years in the CRS group (**Table 1**). In the control group, 130 (76%) subjects were male, and 41 (24%) were female. In the CRS group, 62 (66%) subjects were male, and 32 (34%) were female. There was no significant difference in sex distribution between the two groups ($P=.207$). The mean preoperative SNOT-22 score in the CRS group was 38.6, the median was 37.5, and the interquartile range (IQR) was 20–55. The mean score in the control group was 9.78 (median 9, IQR 3–31), and the difference in the means was statistically significant ($P<.001$).

Discriminant validity assessment suggested that the Arabic SNOT-22 was highly capable of discriminating between healthy subjects and those with CRS. In the Mann-Whitney U test, U=18.36 and Z=-9.8 ($P<.001$). In ROC curve analysis of discriminant validity, the AUC

Table 1. Comparison between cases and control group.

		Cases (n=94)	Control (n=171)	*P* value
Age (years)	38.9 (13.8)	36.0 (14.3)	0.11	
Sex	Female	32.0 (34.0)	4.0 (24.0)	.14
	Male	62.0 (66.0)	130.0 (76.0)	
SNOT-22	Preoperative 1	38.6(23.9)	9.8 (7.5)	<.001
	Preoperative 2 (test 1)	29.0 (21.8)		
	Preoperative 2 (retest 1)	33.4 (28.1)		
	Post-operative	13.0 (11.1)		

Data are mean (standard deviation).

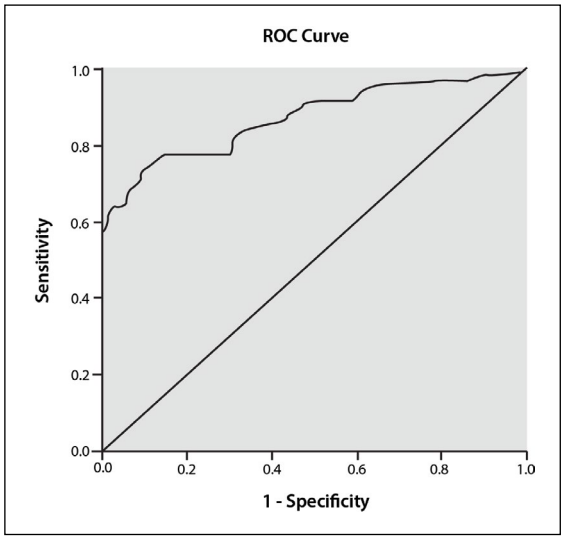

Figure 1. Discriminant validity assessed via ROC curve analysis of the SNOT-22 scores.

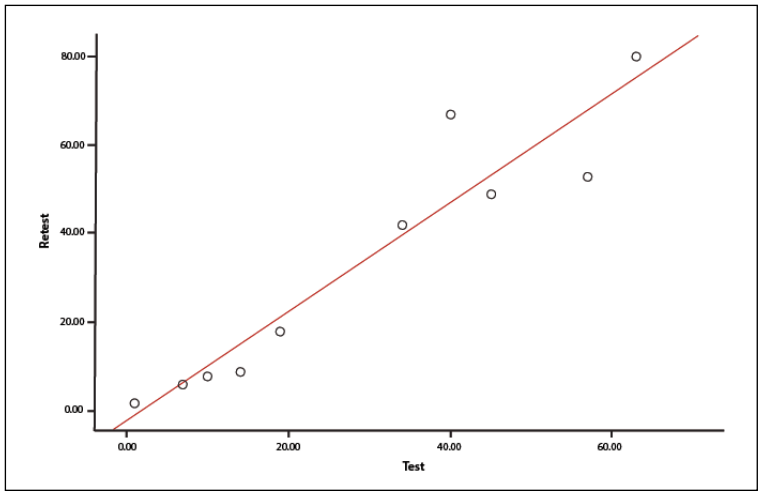

Figure 2. Spearman correlation for test and retest reliability.

was 0.87 at a cut-off threshold of 18.5, sensitivity was 77%, and specificity was 84% (**Figure 1**). It has been established that an AUC of 0.8–0.89 equates to "good".

We performed the investigation twice preoperatively in the CRS group, to investigate the reliability and stability of the Arabic SNOT-22 over time. The mean (SD) SNOT-22 score in the first preoperative test (test 1) was 29.0 (21.8), and in the second preoperative test (test 2) it was 33.40 (28.1). The overall mean for both preoperative tests was 31.2 (24.0). The test-retest Spearman correlation coefficient was 0.964 (*P*<.001), indicating the high reliability of repeated measures and high stability over time **(Figure 2**).

Internal consistency was measured via Cronbach's alpha, and the minimal score deemed acceptable was ≥0.7. The Cronbach's alpha for total scores for the Arabic SNOT-22 indicated high internal consistency ranging between 0.94 to 0. 84. Patient responsiveness was assessed via the Wilcoxon signed-rank test. The mean SNOT-22 score was 42.1 (25.4) preoperatively, and postoperatively it was 13.0 (11.1). The reduction in mean score postoperatively was highly significant (*P*<.001). In the Wilcoxon signed-rank test, t=6.28, z=4.11, r=0.41, and the effect size was 1.2. The MID was set as a decrease of 12.7 points.

DISCUSSION

Measuring health-related QoL has become an important part of healthcare. It assists in the assessment of the burden of diseases and the effects of treatment from the patient's own perspective and can therefore be used to improve the quality of care. This can be achieved using general or disease-specific validated QoL instruments. However, disease-specific instruments are preferable because they focus on measuring specific factors that are pertinent to the disease, and therefore they are more relevant and sensitive to

clinical changes. Furthermore, they are less likely to be affected by comorbidities.[13] Morley and Sharp[14] identified 15 CRS-specific instruments. After comparing their reliability, validity, sensitivity, and ease of use, they concluded that the SNOT-22 was the most suitable. Quintanilla-Dieck et al[15] conducted a systematic review of CRS-specific QoL surveys and concluded that the SNOT-22 was among the most commonly used validated instruments.

The SNOT-22 was adapted from the SNOT-20 via the addition of two items (nasal blockage and change in smell/taste) in a trial to improve the content validity (ability to measure all important aspects of the disease) and responsiveness of the instrument. It has since proved to be a reliable, valid, and highly responsive instrument[5] that cover all four cardinal CRS symptoms described in the EPOS diagnostic criteria.[2] Moreover, factor analysis of the English SNOT-22 revealed five subscales measuring disease-specific (rhinological, extra-nasal rhinological, and ear/facial symptoms) and general health (psychological and sleep dysfunction) domains, which adds to the advantages of the SNOT-22.[16]

Therefore, SNOT-22 is a valuable clinical instrument that can be used at the patient's initial presentation as well as following medical or surgical management to help evaluate the severity of CRS, its progression and the effectiveness of given treatments.

Arabic ranks fifth in the world's most spoken languages and is one of six official languages of the United Nations. Arabic is natively spoken by approximately 422 million people in 22 countries (*www.UNESCO.org*). It has two main variants, literary and dialectal. Literary Arabic is the official variant used in education, printed documents, and media across the Arab world, and is understood by all Arabs with a basic education. Dialectal Arabic varies depending on the region. Therefore, each dialect is limited to a specific population or populations. Both literary and dialectal Arabic have been used in the adaptation of health-measuring instruments, with the majority being in literary Arabic.[17]

The English SNOT-22 has been translated into many languages including Greek, Lithuanian, Brazilian and European Portuguese, Czech, Hebrew, Danish, French, Persian, Spanish, Turkish, and Thai.[6-8,10-12,18-23] Adnane et al[24] adapted and validated a Moroccan version, but it is in a dialect that is difficult for other Arab populations to interpret. Marglani et al[25] adopted a modified Arabic version of the SNOT-16, but they did not report testing its reliability or responsiveness. Therefore, we sought to adapt and validate a literary Arabic version of the SNOT-22. Our aim in the cross-cultural adaptation of this instrument, in addition to facilitating its clinical use, was to develop a uniform measuring tool that could be used when comparing studies on CRS and its treatment across different cultures, as well as in multinational multicenter research.

In the process of translation, we followed guidelines for cross-cultural adaptation found in the literature to ensure that the Arabic version incorporated the concepts and semantic structure of its parent English SNOT-22.[26] No modifications to the Arabic version were required after pilot testing, indicating that it was easy for the respondents to interpret.

Internal consistency is a measure of the correlation between items that are intended to measure a certain construct and is measured via Cronbach's alpha. The Arabic SNOT-22 exhibited alpha scores ranging from good to excellent (0.83–0.94) for the total score in both preoperative and postoperative, similar to the English (0.91) and other versions (**Table 2**).

Test-retest reliability reflects an instrument's ability to produce stable results when it is re-administered after a period of time. In the current study, the interval prior to repeating testing ranged from 2 to 7 days. We believe this was long enough to limit recall bias and short enough to minimize the potential for clinical changes that could result in wrongly concluding that the tool lacked reliability. The Arabic SNOT-22 exhibited excellent stability, with a test-retest correlation coefficient of 0.964, which is similar to that of the original English version of the SNOT-22 (**Table 2**).

In the current study, the Arabic SNOT-22 yielded an effect size for surgery of 1.2, which is considered large. This effect size is smaller than those reported in studies using Lithuanian (1.48),[7] Brazilian (1.55),[8] and Moroccan (1.27)[24] adaptations of the SNOT-22. However, it is larger than that reported in a study that validated the English version (0.81).[5] This discrepancy may be because that study had a much larger number of subjects (n=2077 vs. 94) and because the mean postoperative score was a lot lower in the current study (13.04 vs. 25.5), while their mean baseline score was comparable to ours (41.7 vs. 38.5). **Table 2** compares data from validation studies investigating different versions of the SNOT-22.

The MID pertaining to the Arabic SNOT-22 was 12.69. The MID is defined as the smallest change in scores that a group of patients can detect as a clinical change. MID signifies a score change that is clinically significant, rather than statistically significant. In other words, a change in score that is less than the MID is not considered a clinical improvement or decline, even if it is statistically significant. MIDs have been reported for the original English SNOT-22 and some translated ver-

Table 2. Psychometric analysis across different validation studies.

	English[5]	Arabic	Brazilian[8]	Morrocan[24]	Luithianian[7]	Spanish[21]	Greek[6]	French[19c]
Score (control)	9.3	9.78 (7.5)	11.4 (9.5)	14.5 (5.1)	16.8 (16.1)	4.5 (7.3)	13.0 (11.7)	8.3 (8.7)
Score (preoperative)	41.7 (19.8)	38.5 (23.9)	62.4 (25.3)	50.4 (21.1)	52.4 (20.2)	47.0 (20.9)	44.3 (12.6)	41.0 (20.8)
Score (postoperative)	25.5 (20.8)	13.0 (11.1)	23.1 (18.8)	23.4 (19.1)	22.5 (20.9)	16.5 (13.4)	11.2(11.4)	NA
Internal consistency	0.91	0.949	0.927	0.968	0.93	0.91	0.84-.89	0.93
Test-retest correlation	0.93	0.964	0.72-0.81	0.993	0.72	0.87 ICC	0.91 r	0.78
Discriminant validity	*P*<.0001, *t*	*P*<.001,*U*	*P*<.0001, *t*	*P*<.0001, *t*	*P*<.0001, *U*	*P*<.0001, *U*	*P*<.0001, *U*	*P*<.0001), *U*
Sensitivity	NA	77%[a]	NA	NA	91.7[b]	NA	NA	NA
Specificity	NA	84%[a]	NA	NA	82.6[b]	NA	NA	NA
Responsiveness	*P*<.0001, *t*	*P*<.001, *W*	*P*<.0001, *t*	*P*<.0001, *W*	*P*<.0001, *t*	*P*<.0001, *W*	*P*<.0001, *t*	*P*<.0001, *W*
Effect size	0.81	1.2	1.55	1.27	1.48	1.03	0.81	NA
MID	8.9	12.69	13.87	10.57	13.3	NA	NA	9.6
ROC-AUC	NA	0.87	NA	0.994	0.92	NA	NA	NA

Data are mean (standard deviation) unless noted otherwise. Cronbach's alpha; ICC: intraclass correlation coefficient; MID: minimally important difference; *t*: t-test. *U*: Mann-whitney U test. *W*: Wilcoxon signed-rank test. NA: Data not available

[a]At cut off 18.5; [b]At the off 29; [c]French study included all sinonasal disease patients and not limited to CRS patients.

sions, and the reported values vary between 8.9 and 13.85,[7,8,19] (**Table 2**).

The present study had some limitations. The sample size was relatively small, but comparable with average numbers of other international SNOT-22 validation studies.[6-8,10,11,18,21] Furthermore, validation of the SNOT-22 in the current study was in a Saudi Arabian population. We did not enroll subjects from other Arab countries. Nevertheless, we used literary Arabic to allow for broader use across Arab nations. We recommend cross-validation in future studies. In conclusion, the first Arabic version of the SNOT-22 demonstrated good reliability and validity for the assessment the QoL in CRS patients in an Arabic population.

Acknowledgment

The authors would like to thank the Rhinology Research Chair at Otolaryngology Head and Neck Surgery, College of Medicine, King Saud University, University for funding this research.

Conflict of interest

Authors declare that no conflict of interest in this study.

REFERENCES

1. Hopkins C, Browne JP, Slack R, Lund V, Topham J, Reeves B, et al. The national comparative audit of surgery for nasal polyposis and chronic rhinosinusitis. Clin Otolaryngol. 2006;31(5):390-398.
2. Fokkens WJ, Lund VJ, Mullol J, Bachert C, Alobid I, Baroody F, et al. EPOS 2012: European position paper on rhinosinusitis and nasal polyps 2012. A summary for otorhinolaryngologists. Rhinology. 2012;50(1):1-12.
3. DeConde AS, Soler ZM. Chronic rhinosinusitis: Epidemiology and burden of disease. Am J Rhinol Allergy. 2016;30(2):134-139.
4. Anderson ER, Murphy MP, Weymuller EA. Clinimetric evaluation of the Sinonasal Outcome Test-16. Student Research Award 1998. Otolaryngol Head Neck Surg. 1999;121(6):702-707.
5. Hopkins C, Gillett S, Slack R, Lund VJ, Browne JP. Psychometric validity of the 22-item Sinonasal Outcome Test. Clin Otolaryngol. 2009;34(5):447-454.
6. Lachanas VA, Tsea M, Tsiouvaka S, Hajiioannou JK, Skoulakis CE, Bizakis JG. The sino-nasal outcome test (SNOT)-22: validation for Greek patients. Eur Arch Otorhinolaryngol. 2014;271(10):2723-2728.
7. Vaitkus S, Padervinskis E, Balsevicius T, Siupsinskiene N, Staikuniene J, Ryskiene S, et al. Translation, cross-cultural adaptation, and validation of the sino-nasal outcome test (SNOT)-22 for Lithuanian patients. Eur Arch Otorhinolaryngol. 2013;270(6):1843-1848.
8. Kosugi EM, Chen VG, Fonseca VM, Cursino MM, Mendes Neto JA, Gregório LC. Translation, cross-cultural adaptation and validation of SinoNasal Outcome Test (SNOT): 22 to Brazilian Portuguese. Braz J Otorhinolaryngol. 2011;77(5):663-669.
9. Caminha GP, Melo Junior JT, Hopkins C, Pizzichini E, Pizzichini MM. SNOT-22: psychometric properties and cross-cultural adaptation into the Portuguese language spoken in Brazil. Braz J Otorhinolaryngol. 2012;78(6):34-39.
10. Schalek P, Otruba L, Hahn A. Quality of life in patients with chronic rhinosinusitis: a validation of the Czech version of SNOT-22 questionnaire. Eur Arch Otorhinolaryngol. 2010;267(3):473-475.
11. Shapira Galitz Y, Halperin D, Bavnik Y, Warman M. Sino-Nasal Outcome Test-22: Translation, Cross-cultural Adaptation, and Validation in Hebrew-Speaking Patients. Otolaryngol Head Neck Surg. 2016;154(5):951-956.
12. Lange B, Thilsing T, Al-kalemji A, Baelum J, Martinussen T, Kjeldsen A. The Sino-Nasal Outcome Test 22 validated for Danish patients. Dan Med Bull. 2011;58(2):A4235.
13. Hopkins C. Patient reported outcome measures in rhinology. Rhinology. 2009;47(1):10-17.
14. Morley AD, Sharp HR. A review of sinonasal outcome scoring systems - which is best? Clin Otolaryngol. 2006;31(2):103-109.
15. Quintanilla?Dieck L, Litvack JR, Mace JC, Smith TL. Comparison of disease?specific quality?of?life instruments in the assessment of chronic rhinosinusitis. Paper presented at: International forum of allergy & rhinology 2012.
16. DeConde AS, Bodner TE, Mace JC, Smith TL. Response shift in quality of life after endoscopic sinus surgery for chronic rhinosinusitis. JAMA Otolaryngol Head Neck Surg. 2014;140(8):712-719.
17. Al Sayah F, Ishaque S, Lau D, Johnson JA. Health related quality of life measures in Arabic speaking populations: a systematic review on cross-cultural adaptation and measurement properties. Qual Life Res. 2013;22(1):213-229.
18. de Vilhena D, Duarte D, Lopes G. Sino-Nasal Outcome Test-22: translation, cultural adaptation and validation in Portugal. Clin Otolaryngol. 2016;41(1):21-24.
19. de Dorlodot C, Horoi M, Lefebvre P, Collet S, Bertrand B, Eloy P, et al. French adaptation and validation of the sino-nasal outcome test-22: a prospective cohort study on quality of life among 422 subjects. Clin Otolaryngol. 2015;40(1):29-35.
20. Jalessi M, Farhadi M, Kamrava SK, Amintehran E, Asghari A, Hemami MR, et al. The reliability and validity of the persian version of sinonasal outcome test 22 (snot 22) questionnaires. Iran Red Crescent Med J. 2013;15(5):404-408.
21. de los Santos G, Reyes P, del Castillo R, Fragola C, Royuela A. Cross-cultural adaptation and validation of the sino-nasal outcome test (SNOT-22) for Spanish-speaking patients. Eur Arch Otorhinolaryngol. 2015;272(11):3335-3340.
22. Hanci D, Altun H, Sahin E, Altintoprak N, Cingi C. Turkish translation, cross-cultural adaptation and validation of the SinoNasal Outcome Test (SNOT)-22. ENT Updates. 2015;5(2):51.
23. Numthavaj P, Bhongmakapat T, Roongpuwabaht B, Ingsathit A, Thakkinstian A. The validity and reliability of Thai Sinonasal Outcome Test-22. Eur Arch Otorhinolaryngol. 2017;274(1):289-295.
24. Adnane C, Adouly T, Oubahmane T, Zouak A, Rouadi S, Lah Abada R, et al. Psychometric Validation of a Moroccan Version of the 22-Item Sino-Nasal Outcome Test. Otolaryngol Head Neck Surg. 2016;155(4):681-687.
25. Marglani O, Alherabi A, Abdul Fattah T, Al-Zalabani, A. The introduction and validation of a modified Arabic Sino-Nasal outcome test (MA-SNOT). PAN Arab J Rhinol. 2011;1:10-13.
26. 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(2):94-104.