

# A Systematic Review and Meta-Analysis of Rhinoplasty Using the Rhinoplasty Outcome Evaluation Scale

Riccardo Nocini, Salvatore Chirumbolo<sup>1</sup>, Ali Pirayesh<sup>2</sup>, Eqram Rahman<sup>3</sup>, Krishan Mohan Kapoor<sup>4</sup>, Gulser Caliskan<sup>5</sup>, Dario Bertossi<sup>6</sup>

Department of Surgery, Dentistry, Paediatrics and Gynaecology, Unit of Otorhinolaryngology, University of Verona, <sup>1</sup>Department of Neurosciences, Biomedicine and Movement Sciences, Unit of Human Anatomy, University of Verona, <sup>2</sup>Department of Diagnostics and Public Health, Unit of Epidemiology and Medical Statistics, University of Verona, <sup>3</sup>Department of Surgery, Dentistry, Paediatrics and Gynaecology, Unit of Maxillo-Facial Surgery, University of Verona, Verona, Italy, <sup>4</sup>Department of Plastic Surgery, Amsterdam Plastic Surgery, Amsterdam, Netherlands, <sup>5</sup>Department of Plastic and Reconstructive Surgery, Royal Free Hospital, London, UK, <sup>6</sup>Department of Plastic Surgery, Anticlock-Age Reversal Clinic and Medispa, Chandigarh, India

## Abstract

**Background:** The purpose was to evaluate the role and impact of the rhinoplasty outcome evaluation (ROE) in rhinoplasty. **Objectives:** To systematically review the ability of ROE to assess rhinoplasty outcome in the very recent surgical approaches. **Data Sources:** Searching MESH terms “rhinoplasty” and “rhinoplasty and ROE” on PubMed, Scopus, Embase, Google Scholar, and other major literature databases. **Study Eligibility Criteria:** A Preferred Reporting Items for Systematic Reviews and Meta-Analyses model was adopted for eligibility criteria. **Participants and Interventions:** A first large screening on 1,660 papers allowed the final retrieval of 896 reports, of which only 101 were eligible for all the inclusion criteria, which allowed finally to select a group of 12 reports published in the time range 2018–2021. **Study Appraisal and Synthesis Methods:** Meta-regression with Durbin–Watson test, checking of meta-analysis bias with Rosenthal’s estimator, and heterogeneity through the  $I^2$  statistics were performed. The analysis included a Mantel–Haenszel test and associated or meta-analysis. **Results:** Twelve eligible and qualified papers were included in the meta-analysis, out of 1,263 patients, from 1,660 papers. ROE confirmed previous reports about its ability to discriminate preoperative from postoperative outcome (50/100), as from the different rhinoplasty approaches, a value very close to the functional Nasal Obstruction Symptom Evaluation scale. This study strengthens the role of ROE in the evaluation of rhinoplasty outcomes encouraging further research to improve ROE scale on different cultural habits and geographical areas. **Limitations:** The main limitations were due to the weak inclusion and exclusion criteria, methodological flaws with the statistical analysis, and poor data synthesis and evaluation in many papers about ROE.

**Keywords:** Meta-analysis, rhinoplasty, rhinoplasty outcome evaluation, scoring, survey

## INTRODUCTION

Nasal reconstruction, i.e., nasal surgery or rhinoplasty, encompasses a huge deal of practices dating back, surprisingly, to very ancient times.<sup>[1]</sup> Surgical rhinoplasty is currently considered, alongside nonsurgical rhinoplasty, within the widest terminology of “rhinoplasty,” probably because of the huge diffusion of dermal fillers and other nonsurgical devices, rapidly reported as a sound alternative to minor surgical procedures. An explanation of this rapid diffusion may come from its relatively low cost, low-risk profile, professional convenience, and patient’s rapid recovery.<sup>[2–6]</sup> As a matter of fact, rhinoplasty is widely considered to date a very common cosmetic approach, reaching, in the last 5 years, the third most frequently requested aesthetic procedure in the U.S., with an estimated total expenditure exceeding 1 billion dollars.<sup>[7]</sup>

On the other hand, surgical care has more and more shifted to freestanding Ambulatory Centers of Surgery (ASCs). Some authors recently assessed that the proportion of outpatients referring to ASCs, in the years 2010–2017, increased by 1.8%, whereas the proportion located in hospital departments decreased by 6% in the indicated time range. Interestingly,

**Address for correspondence:** Dr. Salvatore Chirumbolo, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Strada Le Grazie 9 37134 Verona, Italy. E-mail: salvatore.chirumbolo@univr.it

Received: 09-10-2021

Last Revised: 12-05-2022

Accepted: 29-06-2022

Published: 16-08-2022

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow\_reprints@wolterskluwer.com

**How to cite this article:** Nocini R, Chirumbolo S, Pirayesh A, Rahman E, Kapoor KM, Caliskan G, *et al.* A systematic review and meta-analysis of rhinoplasty using the rhinoplasty outcome evaluation scale. *Ann Maxillofac Surg* 2022;12:60–8.

### Access this article online

Quick Response Code:



Website:  
<https://journals.lww.com/aoms/>

DOI:  
10.4103/ams.ams\_244\_21

rhinoplasty accounted for the largest absolute increase, i.e., 8.9% absolute and 33.5% relative, in the frequency and use of ASCs, resulting in an estimated cost saving of more than 7.1 million dollars in 2017.<sup>[8]</sup>

In this perspective, rhinoplasty outcomes are a fundamental issue in this surgical practice, not only for functional reasons but also for aesthetic ones, which yet appear to overcome, in their overall impact, the comprehensible need to restore the simplest nose function. This may elucidate why surgeons have very recently shifted in paying major attention toward scoring panels including patient's satisfaction and aesthetic outcome in rhinoplasty.

Open and closed rhinoplasty are two different strategies to surgically access the nose. The Rhinoplasty Outcome Evaluation (ROE) is a very easy-to-use questionnaire, which should allow a comprehensive and thorough assessment of rhinoplasty-related patient satisfaction, in order to improve rhinoplasty and encourage the application of more straightforward and innovative techniques in nose plastic and reconstruction surgery.<sup>[9,10]</sup>

A recent paper by Metin and Avcu evaluated the ROE scale in both open and closed septorhinoplasty performed on 370 patients and concluded that in open septorhinoplasty 194 patients reached an ROE  $\geq 12$  whereas only 23 with ROE  $< 12$  points (ratio 16:1),<sup>[11]</sup> while in closed septorhinoplasty 137 patients with ROE  $\geq 12$  and 16 with ROE  $< 12$  points (ratio 8:1), apparently suggesting that open surgery resulted in a higher outcome performance.<sup>[11]</sup> The conclusions held by the authors, using also the functional scale known as the Nasal Obstruction Symptom Evaluation (NOSE), assessed that the aesthetic component affecting the nasal topographic satisfaction was the nasal tip. However, they did not succeed in confirming the purported correlation between low satisfaction with surgical approach and aesthetic outcome in the few groups of patients with post-operative low ROE.<sup>[11]</sup> Therefore, NOSE alone might not be particularly suitable to assess outpatients' satisfaction, when undergoing rhinoplasty in ASCs.

In nasal reconstructive surgery and rhinoplasty, therefore, a great debate is going ahead about which questionnaire should be used in weighing patient's satisfaction following rhinoplasty. The ROE questionnaire is particularly easy-to-use, allowing a complete assessment of patient's satisfaction following rhinoplasty surgical outcome.<sup>[9,12,13]</sup>

However, concerns exist about the extreme variability in the different medical approaches used for rhinoplasty. To date, the term rhinoplasty accounts on more than 8 million items on the web and more than 11,000 reports in the National Library of Medicine, being a highly widespread terminology in the most common popular knowledge.<sup>[14,15]</sup> Surgeons are particularly interested in retrieving data about which surgical approach is the better available to reach the highest ROE score.

Some systematic reviews have attempted a possible response.

Floyd *et al.*, for example, used the functional scale known as NOSE, to assess changes in the score following surgery, i.e., using the mean difference between preoperative (baseline) and postoperative results, including the standardised mean difference, and calculating heterogeneity with  $I^2$  statistics. Their conclusions were that NOSE was able to measure nasal obstruction (as personal perception) by 43–50 points (out of 100).<sup>[16]</sup> This evidence is fundamental to comprehend the level of patient's satisfaction following rhinoplasty, which cannot be only of aesthetic nature. Actually, ROE includes both functional and aesthetic issues and can be used to trace the reliability and quality of different rhinoplasty approaches used in the very recent years, accounting on both functional and aesthetic outcomes.

How rhinoplasty evolved in the latest 3 years of surgical expertise? New innovative techniques emerged, for example, those regarding nasal hump deformity, such as let-down technique,<sup>[17-19]</sup> or a combination of push-down and let-down techniques,<sup>[18]</sup> or even the debate about the role and the concept of septolateral cartilage in septoplasty.<sup>[20]</sup> Understanding how ROE can be used as a surgical outcome marker to follow the feasibility and reliability of the very recent reports on rhinoplasty, aside from different approaches, is the major goal of this contribution.

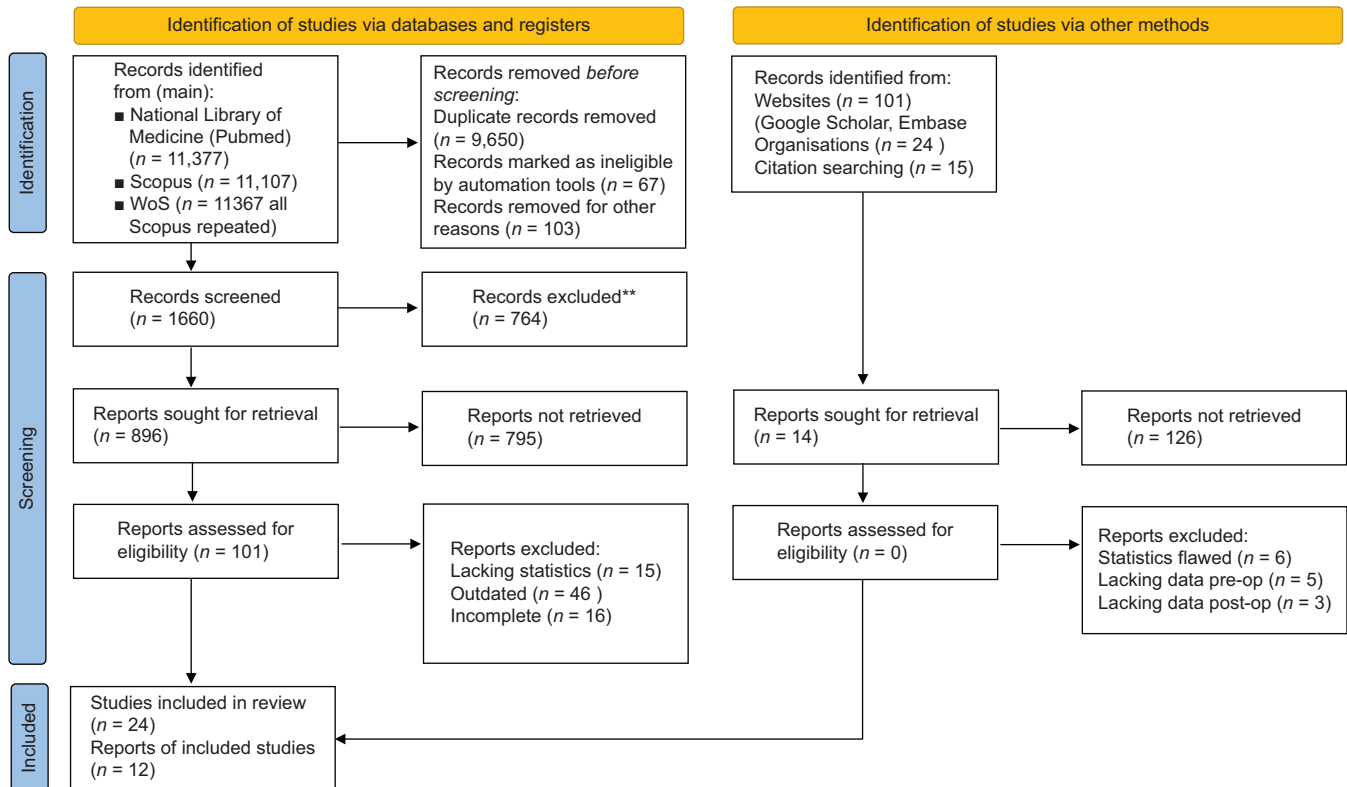
## MATERIALS AND METHODS

### Bibliographic search

Figure 1 shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses performed for the bibliographic search reported in this manuscript. Two independent people (see acknowledgments) conducted a thorough research on the National Library of Medicine (Pubmed), Google Scholar, Embase, Scopus, Web of Science, besides to further collective databases from the European Scientific Societies in Rhinoplasty, Otorhinolaryngology and Maxillo-Facial Surgery. A trusting number of eligible papers not exceeding 12,000 was retrieved, using the major MESH term "rhinoplasty." They provided the selection of data, extraction, and elaboration, by a Prometa 3.0 software with a Meta-Mar graphical software, using the MESH terms and keywords "rhinoplasty," "ROE," "patient's satisfaction," "outcome," "scoring," and meta-analysis was approached according to the Cochrane Handbook and the guidelines of meta-analyses of observational studies.<sup>[21]</sup>

### Inclusion and exclusion criteria

The inclusion criteria considered rhinoplasty clinical studies with ROE scores expressed as mean  $\pm$  standard deviation (SD) encompassing reports published within at least the latest 3 years, 2018–2021. Exclusion criteria were: (a) number of patients  $\leq 10$ ; (b) repeated papers; (c) papers with ROE expressed as graphs, without values; (d) papers in languages other than English; (e) paper lacking either inclusion or exclusion criteria, number of patients, surgical approach, type of study; and (f) papers lacking a DOI number.



**Figure 1:** PRISMA model for paper recruitment and selection. \*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers). \*\*If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools. From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>

**Data extraction**

Articles were retrieved and selected by two independent researchers from our clinics (DB, GC), following the aforementioned inclusion and exclusion criteria, with the aid of our co-authored statistician (GC). A total amount of 11,377 papers were considered for a selection process and 9,650 rejected because of reiterated papers. The meta-analysis, limited to the years 2018–2021 considered a total amount of 1,263 patients in 12 studies [Table 1]. This time range was selected to highlight the impact of the most straightforward and homogeneously described surgical techniques on the ROE scale, reducing the effect of more outdated or disused surgical approaches on the patient’s outcome and excluding the introduction of not standardised, highly empirical approaches. A first large screening on 1,660 papers allowed the retrieval of 896 reports, of which only 101 eligible for all the inclusion criteria, which allowed finally to select a group of 12 reports published in the time range 2018–2021. Papers were thoroughly read, discussed, and considered for the meta-analysis and major data reported [Table 1].<sup>[14,16,22-34]</sup>

**Statistics**

Each ROE was expressed as mean ± SD and statistically evaluated through a Wilcoxon rank test with  $P < 0.05$ . For outliers, a Grubb’s test for outliers was performed. Normality assessment of the data distribution was performed with a

Jarque–Bera test. The Jarque–Bera test is a statistical test for verifying the hypothesis of normality and is used very often in meta-analysis. It is based on the measurement of the asymmetry and kurtosis of a distribution.

The analysis of effect size and meta-analysis bias accounted on the Hedge’s g evaluation and the Fail-N-Safe (FNS) test,<sup>[35]</sup> using either Rosenthal’s or Rosenberg’s estimators, a test used to assess the stability of a meta-analysis study, through the demonstration of how many “null” articles should be requested to revise the statistically significant results to not significant ones.<sup>[36,37]</sup> To evaluate how much the results retrieved from the different papers differed in terms of regression, respect to a standard homogeneous mean, we used a type of linear least squares approach known as ordinary least squares (OLS), with which we are enabled to estimate any unknown parameter in a linear regression model and fit the regression test to its best linear function. The model uses the Akaike’s information criterion (AIC), which evaluates how much information is lost, and the Bayesian information criterion (BIC), which evaluates different models with each other. The lowest values of AIC and BIC are preferable.

Meta-regression calculation was performed with a Durbin–Watson *post hoc* test ( $DW \geq 1.0$ ).

When more than one covariate occurs, the Q statistic is used as an omnibus test (Omnibus) of the hypothesis that all the

**Table 1: Papers for eligibility in meta-analysis**

| Exclusion criteria  | Inclusion criteria  | Type                                 | Patients<br><i>n</i> | Preoperative<br>Roe | Postoperative<br>Roe | Type of surgery   | Reference |
|---|---|--------------------------------------|----------------------|---------------------|----------------------|---|-----------|
| Acute nasal trauma or nasal bone fracture up to 3 months before surgery   | Alar notching and nostril exposure existing prior to or following rhinoplasty. Furthermore dorsal humps, lack of projection, crooked nose and tip drooping  | Retrospective review                 | 20                   | 41.13±19.71         | 80.00±14.38          | Alar retraction with LLC graft  | 22        |
| Additional concomitant procedures, (functional endoscopic sinus surgery, blepharoplasty, or also previous diagnosis of BDD) | Patients with nasal obstruction associated with aesthetic complaints  | Cohort study                         | 131                  | 37.65±16.68         | 79.56±2.12           | Septorhinoplasty  | 23        |
| Chronic illnesses incompatible with surgery   | Outpatients incoming for rhinoplasty  | RCT                                  | 74                   | 37.86±10.64         | 84.96±9.95           | Endoscopic surgery  | 24        |
| Chronic illnesses incompatible with surgery. Use of incompatible drugs  | Outpatients incoming for aesthetic rhinoplasty  | Retrospective                        | 100                  | 52.02±1.03          | 65.37±0.59           | Nonsurgical rhinoplasty   | 25        |
| Patients younger than 18 year old or suffering major nasal sad-dling  | Patients selected as surgical candidates  | Retrospective                        | 21                   | 20.23±7.37          | 80.75±6.24           | Revision rhinoplasty with rainbow graft                                 | 26        |
| Patients with congenital nasal deformities  | Patients from both genders, between 16 and 60 years of age  | Retrospective                        | 118<br>rec/90 q      | 30.5±4.65           | 72.5±2.7             | Open rhinoplasty  | 14        |
| Patients younger than 18 year old or suffering major nasal sad-dling  | Patients selected as surgical candidates  | Prospective case-series study        | 36                   | 20.94±8.67          | 79.56±10.65          | Fascia lata graft surgery   | 27        |
| Chronic illnesses incompatible with surgery. Use of incompatible drugs  | Outpatients incoming for aesthetic rhinoplasty  | Prospective, controlled cohort study | 25                   | 19.78±5.74          | 81.22±10.31          | External approach through an inverted V-shaped transcolumellar incision | 28        |
| Chronic illnesses incompatible with surgery. Use of incompatible drugs  | Outpatients incoming for aesthetic rhinoplasty  | Prospective                          | 105 rec,<br>51 q     | 41.10±16.0          | 74.30±24.1           | Functional septorhinoplasty   | 29        |
| Chronic illnesses incompatible with surgery. Use of incompatible drugs  | Outpatients incoming for aesthetic rhinoplasty  | Prospective                          | 102                  | 39.80±15.20         | 75.10±24.0           | Septorhinoplasty  | 30        |
| Incomplete data collection, primary rhinoplasty, and patient age under 18 years at the time of surgery                      | Candidates for aesthetic or functional-aesthetic revision rhinoplasty with one previous aesthetic or functional-aesthetic rhinoplasty in another department | Prospective                          | 64                   | 28.2±12.8           | 61.9±24.2            | Septorhinoplasty  | 31        |
| Incomplete data collection, primary rhinoplasty, and patient age under 18 years at the time of surgery                      | Candidates for aesthetic or functional-aesthetic revision rhinoplasty with one previous aesthetic or functional-aesthetic rhinoplasty in another department | Prospective                          | 60                   | 51.27±10.54         | 79.6±9.67            | Open rhinoplasmy  | 32        |

LLC=Lower lateral cartilage; BDD=Body dysmorphic disorder

covariates are zero. In this case, the z-test may be used to test any coefficient, holding the others constant.

The calculation of heterogeneity in meta-analysis is adopted to evaluate the variation in the different study outcomes

between reports. Besides the Hedges'  $g$  SMD test, we used the statistics of  $I^2$ , which can describe the percentage of variation across different eligible studies due to heterogeneity rather than chance. When the heterogeneity test was not significant and the  $I^2$  statistics was lower than 30%, a fixed-effects model was adopted, while a random-effect model was chosen when the test of heterogeneity was significant or the  $I^2$  statistics was larger than 30%. Moreover, this value should be close to 100%, as  $I^2 = 100\% \times (Q-df)/Q$ , where  $Q$  is the Cochran's  $Q$ .<sup>[38]</sup> Data were elaborated with SPSS v 24 and STATA v. 10. (IBM SPSS Statistics v24, Ivrea (TO), Italy, licensed by University of Verona, Italy and STATA v10 from StataCorp LLC 4905 Lakeway Drive College Station, Texas, USA).

### Mantel–Haenszel test and odds ratio meta-analysis

Considering the post-intervention (postoperative) and the preintervention (preoperative) as a kind of case–control study, to assess if in the postoperative evaluation, the research can be warranted in its outcome using the ROE scale, a Mantel–Haenszel (MH) test was applied on the eligible papers. The MH method is generally applied to estimate the pooled odds ratio for all strata when one assumes a fixed-effects model:

$$\widehat{OR}_{MH} = \frac{\sum_{i=1}^k \left( \frac{a_i d_i}{n_i} \right)}{\sum_{i=1}^k \left( \frac{b_i c_i}{n_i} \right)}$$

Where  $n_i = a_i + b_i + c_i + d_i$

In general, alternative methods such as inverse variance and the Woolf test could be used, yet the MH test is much more robust.

### Effect size and other statistics

The evaluation of heterogeneity, using either a fixed or random model, used ancillary tests to evaluate the reliability of the model itself.

The Hedge's  $g$  is a measure of effect size and is given by:

$$\text{Hedge's } g = \frac{M_1 - M_2}{SD^*_{\text{pooled}}}$$

where  $M_1 - M_2$  represents the difference in means, and  $SD^*_{\text{pooled}}$  represents the pooled and weighed SD (or standard mean deviation, SMD). The evaluation of the effect size provides us with knowing how much one group (e.g., test group) differs from the other (e.g., the control group). The evaluation includes variability as standard error of  $g$  (SE $g$ ), 95% confidence intervals and statistics as  $z$  score and  $P$  values.

## RESULTS

In the 12 included papers, ranging from 2018 to date, ROE increased from  $35.04 \pm 11.20$  SD in the preoperative period to  $76.80 \pm 5.54$  SD in the postoperative period, and the increase was significant as expected ( $P = 0.000036442$ , i.e.,

$P < 0.0001$ ). The  $\Delta_{\text{pre/post}}$  was + 54.37%, i.e., improvement was higher than 50%. This would mean that ROE, despite the variability in the rhinoplasty techniques adopted, is a highly affordable, easy-to-use, and reliable tool to evaluate rhinoplasty performance. Variability in the different papers may be considered a possible bias. However, outliers were reported for preoperative ( $52.02$ ,  $z = 1.5152$ ,  $P > 0.05$ ) and postoperative period ( $65.37$ ,  $z = 2.18102$ ,  $P > 0.05$ ), as well. If excluding data positive at the Grubb's test for outliers, the  $\Delta_{\text{pre/post}}$  was + 57.12%, including two distinct populations ( $P = 0.0001865$ ), a preoperative ROE of  $33.50 \pm 10.36$  and a postoperative ROE of  $78.13 \pm 4.09$ . Therefore, by adjusting the data distribution with a sound statistical method, ROE application improved from 50% to almost 60%. Furthermore, these values are perfectly comparable to NOSE scores, increased by 50 points out of 100, confirmed also in some of the eligible, selected papers ( $\Delta_{\text{pre/post}} = 62.42\%$ ).<sup>[16]</sup>

Table 2 shows both the fixed and random model estimations, with Forest plots, respectively, i.e., A (fixed) and B (random). Table 3 shows that the robustness of meta-analysis was evaluated by the Rosenthal's FNS estimator, which if higher than (5000 + 10), indicates that the likelihood to meet bias in the publications for meta-analysis is minimal. Actually, FNS for Rosenthal's method was 7451.98, widely overwhelming  $5K + 10$ . The Durbin–Watson test used to detect any autocorrelation in the residuals (used as prediction errors) in a regression analysis, reported an inconclusive test ( $DW = 1.136$  for 12 reports, i.e.,  $DW > 1.0$ , no cause of alarm). The autocorrelation defines the degree of dependence between the values assumed by a sampled function in its domain on the abscissa. If the autocorrelation between two values is demonstrated, as the peculiarities of one of them change, the other will also vary. Therefore, our Durbin–Watson test established the complete independence of the retrieved 12 papers, providing us with the possibility to be thoroughly objective about our investigation.

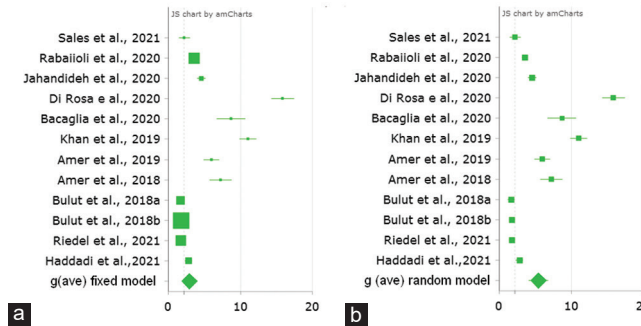
Heterogeneity ( $I^2$ ) was 98.4%, with  $X^2 = 692.99$ ,  $df = 11$ ,  $\text{Tau}^2 = 5.438$ , suggesting that <2% of the variability across studies could be attributed to chance. Tau is the variance in the random-effects meta-analysis, whereas in the fixed model is used a Chi-square test on a defined number of degrees of freedom ( $df$ ). Papers have heterogeneity due to differences in rhinoplasty approaches. Hence, also a random-effects model was adopted, which yielded a Hedges'  $g$  (SMD) of 5.41 ( $IC_{95} = [4.062-6.765]$ ,  $P < 0.0001$ ).

The MH test provided a  $\chi^2$  statistics for preoperative = 16.38 ( $P < 0.001$ ) and for postoperative = 51.15 ( $P < 0.001$ ) (MH  $\chi^2 = 51.04$ ,  $P < 0.001$ ). The use of the MH Chi-square method allows to “adjust” the comparison between the two groups (preoperative and postoperative) for a third variable (heterogeneity). To assess this point, an odd ratio (OR) statistics was also performed. The OR for preoperative was 6.59 ( $IC_{95} = [2.40-18.16]$ ) and for postoperative was 7.21

**Table 2: Statistical evaluation**

| Fixed and random effect model |                 |       |             |         |          |  |
|-------------------------------|-----------------|-------|-------------|---------|----------|--|
| Model                         | Hedge's g (SMD) | SEg   | CI 95       | Z score | P        | Heterogeneity                            |
| Fixed                         | 2.88            | 0.082 | 2.721-3.043 | 35.09   | <0.0001  | $I^2=98.4\%$ , $\chi^2=692.99$ , $df=11$ |
| Random                        | 5.41            | 0.689 | 4.062-6.765 | 7.851   | < 0.0001 | 98.4%, $\text{Tau}^2=5.438$              |

**Forest plots**



SEg=Standard error of g; CI=Confidence interval; SMD=Standard mean deviation

( $IC_{95} = [3.97-13.09]$ ), with an OR in the MH sum = 6.89 ( $IC_{95} = [3.77-12.57]$ ). In this case, the Wolff test, assuming OR homogeneity and 3-way association, gave  $\chi^2 = 0.02$ , OR = 0.882, assessing therefore the previously reported MH test.

**DISCUSSION**

The systematic review here reported showed that ROE can be considered a useful performance tool for rhinoplasty, despite heterogeneity associated with different rhinoplasty approaches. The different statistical tests used in this meta-analysis assessed that ROE can be optimally used, replacing the simplest NOSE questionnaire, despite the slight difference in surgical and practical approaches used in rhinoplasty, so representing a possible forerunner in a standardised panel of rhinoplasty performance. Moreover, the final evaluation of ROE, in assessing the different performances of various rhinoplasty approaches between preoperative and postoperative conditions, complied significantly the NOSE scoring, indicating that ROE can perfectly measure both functional and aesthetic outcomes in different rhinoplasty techniques.

The whole bulk of reports dealing with rhinoplasty and ROE dates back to 2008 when Taylor and Rigby used ROE in a prospective evaluation on six patients undergoing the new technique called Taylor saddle effacement, which used autologous grafts (with cartilages from the lower lateral sides) for correcting nasal saddle deformity.<sup>[39]</sup> Since then, fundamental progress was rapidly earned in rhinoplasty, reaching, as highlighted in this meta-analysis, an increase in the ROE scoring comparable to that of NOSE, i.e., at least 50–60 points out of 100. In 2011, Tastan *et al.* used ROE to assess outcomes following a novel approach for the internal reconstruction of the nasal valve in 19 patients with internal valve dysfunction.<sup>[40]</sup> Tastan *et al.* used ROE not exactly as a satisfaction scale but to assess the good outcome obtained

with NOSE. The integrated assessment between functional and aesthetic outcome following rhinoplasty is a very recent matter of debate, and in this sense ROE, as recently conceived, may provide fundamental and reliable insights into how rhinoplasty reaches its success.<sup>[34,40,41]</sup> ROE is therefore a useful tool to test rhinoplasty performance in the most exhaustive way.

However, a possible limitation of ROE is that the approach and scale probably should be adapted to cultural habits and anthropometric hallmarks. In this perspective, Izu *et al.* tried to validate ROE approach to the Brazilian–Portuguese population,<sup>[12]</sup> even establishing normal values for ROE, i.e., the best cut-off of 50% with 95.16% sensitivity and 95% specificity.<sup>[9]</sup> In Europe, this people adaptation even suggested a German ROE, known as (ROE-D).<sup>[42]</sup> Anyway, it is particularly difficult to standardize a patient’s satisfaction scale merging functional performance and aesthetic perception if this latter closely depends on cultural beliefs and different attitudes. In this regard, further meta-analyses are particularly welcome to shed light on this concern.<sup>[43]</sup>

In conclusion, in this systematic review and meta-analysis, we assessed that the more recent rhinoplasty approaches have an ROE closely joining the 50 points out of 100 of improved outcome, as recommended by ROE validation, aside from the different kinds of rhinoplasty.

**Limitations of the study**

This research study was not registered in PROSPERO as it has to be considered a systematic review without an outcome of clear relevance to the health of humans (controls vs. treated subjects), given from a literature reviews that use a systematic search and assessing only the quality of reporting of ROE. The nature of the research does not required a mandatory registration to PROSPERO.

Due to paucity in ROE evaluation in rhinoplasty, the included articles had one randomised controlled trial and eleven nonrandomised

**Table 3: Meta-analytic data**

| Fail N safe                               |               |                |             |                |                     |   |                    |           |          |           |           |                      |                       |
|---|---------------|----------------|-------------|----------------|---------------------|---|--------------------|-----------|----------|-----------|-----------|----------------------|-----------------------|
| Fail N safe                               |               |                |             |                |                     | Fail-N-Safe                               |                    |           |          |           |           |                      |                       |
| Rosenthal's method (1979) <sup>[36]</sup> |               |                |             |                |                     | Rosenberg's method (2005) <sup>[37]</sup> |                    |           |          |           |           |                      |                       |
| $t_g (\alpha=0.05, df=12)=1.782$          |               |                |             |                |                     | $Z_c (\alpha=0.05)=1.645$                 |                    |           |          |           |           |                      |                       |
| 7451.98                                   |               |                |             |                |                     | 4639.42                                   |                    |           |          |           |           |                      |                       |
| OLS regression results                    |               |                |             |                |                     |   |                    |           |          |           |           |                      |                       |
| Dependent variable                        | $R^2$         | Adjusted $R^2$ | F-statistic | Log-likelihood | Number observations | AIC                                       | BIC                |           |          |           |           |                      |                       |
| g   | -0.0001       | -0.0001        | NA          | NA             | 12                  | 70.80                                     | 71.29              |           |          |           |           |                      |                       |
| Meta regression                           |               |                |             |                |                     |   |                    |           |          |           |           |                      |                       |
| Omnibus                                   | Durbin-Watson | P Omnibus      | Jarque-Bera | Skewness       | P Jarque-Bera       | Kurtosis                                  | Cond. No           |           |          |           |           |                      |                       |
| 5.135                                     | 1.136         | 0.077          | 2.547       | 1.118          | 0.280               | 3.304                                     | Inf                |           |          |           |           |                      |                       |
| Data experiments                          |               |                |             |                |                     |   |                    |           |          |           |           |                      |                       |
| Study name                                | n1            | Mean1          | SD1         | n2             | Mean2               | SD2                                       | Moderator variable | g         | SEg      | g-lower   | g-upper   | Weight % Fixed model | Weight % Random model |
| Sales et al, 2021                         | 20            | 41.13          | 19.71       | 20             | 80.00               | 14.38                                     | 0                  | 2.208299  | 0.396262 | 1.431626  | 2.984972  | 4.297004             | 8.496889              |
| Raballotti et al, 2020                    | 131           | 37.65          | 16.68       | 131            | 79.56               | 2.12                                      | 0                  | 3.514804  | 0.196863 | 3.128951  | 3.900656  | 17.410049            | 8.680373              |
| Jahandideh et al, 2020                    | 74            | 37.86          | 10.64       | 74             | 84.96               | 9.95                                      | 0                  | 4.548941  | 0.310899 | 3.939580  | 5.158302  | 6.980593             | 8.589562              |
| Di Rosa et al, 2020                       | 100           | 52.02          | 1.03        | 100            | 65.37               | 0.59                                      | 0                  | 15.844936 | 0.804676 | 14.267771 | 17.422102 | 1.042048             | 7.812065              |
| Bacaglia et al, 2020                      | 21            | 20.25          | 7.37        | 21             | 80.75               | 6.24                                      | 0                  | 8.692865  | 0.995627 | 6.741437  | 10.644293 | 0.680670             | 7.394365              |
| Khan et al, 2019                          | 90            | 30.50          | 4.65        | 90             | 72.50               | 2.70                                      | 0                  | 10.999808 | 0.598443 | 9.826859  | 12.172757 | 1.884012             | 8.202073              |
| Amer et al, 2019                          | 36            | 20.94          | 8.67        | 36             | 79.56               | 10.65                                     | 0                  | 5.971793  | 0.549565 | 4.894645  | 7.048941  | 2.234040             | 8.282252              |
| Amer et al, 2018                          | 25            | 19.78          | 5.74        | 25             | 81.22               | 10.31                                     | 0                  | 7.247745  | 0.776405 | 5.725991  | 8.769499  | 1.119317             | 7.869870              |
| Budut et al, 2018a                        | 51            | 41.10          | 16.00       | 51             | 74.30               | 24.10                                     | 0                  | 1.610875  | 0.226602 | 1.166736  | 2.055015  | 13.140228            | 8.660459              |
| Budut et al, 2018b                        | 102           | 39.80          | 15.20       | 102            | 75.10               | 24.00                                     | 0                  | 1.750752  | 0.164240 | 1.428841  | 2.072663  | 25.013224            | 8.699084              |
| Riedel et al, 2021                        | 64            | 28.20          | 12.80       | 64             | 61.90               | 24.20                                     | 0                  | 1.730481  | 0.206339 | 1.326056  | 2.134906  | 15.847699            | 8.674322              |
| Haddadi et al, 2021                       | 60            | 51.27          | 10.54       | 60             | 79.60               | 9.67                                      | 0                  | 2.783128  | 0.255312 | 2.282716  | 3.283540  | 10.351116            | 8.638686              |

Fail-N-safe and meta regression. AIC=Alkake's information criterion; BIC=Bayesian information criterion; Omnibus: See text; Cond. No=Number of conditionality on the distribution curve; Inf=Only inferior limits; OLS=Ordinary least square; NA=Not available; SD=Standard deviation; SEg=Standard error of g

observation studies (of which 6 prospective and 6 retrospective). Owing to the inherent weakness of the study design, most of the included studies exhibited medium to low qualitative rigor, and one high-quality study. This hallmark shows the weakness in the clinical reports in rhinoplasty for ROE. The main reasons were the weak inclusion and exclusion criteria, methodological flaws with the statistical analysis, and poor data synthesis and evaluation in many papers. Moreover, a huge variability was recorded among studies, as denoted by an  $I^2$  statistics close to 100%. This variability likely reflects differences in surgical skills among centers.

The main limitations were due to the weak inclusion and exclusion criteria, methodological flaws with the statistical analysis, and poor data synthesis and evaluation in many papers about ROE. Actually, this study has limitations for which further research is in progress in our surgical units. A weakness in the inclusion and exclusion criteria is due to the relative paucity in homogeneously selected paper, due to a significant variability in different surgical techniques employed. Time range is restricted to a relatively short period because the reliability of ROE data was particularly concentrated in the past 3–4 years, thanks to a fundamental improvement in statistics and surgical standardisation of methodology. A huge variability in the different surgical methods and expertise and the different applications of ROE is a further source of bias and weakness of the current meta-analysis, for which rigorous statistical tools have been adopted. Finally, the different use of ROE scoring in performing statistics represents another source of variability in the retrieval of a systematic literature.

Finally, even patients' age is a possible confounder in the reliability of ROE.<sup>[44]</sup> Yet, Arima *et al.* reported that in 61 out of 112 patients responding to the ROE questionnaire, a mean difference of 50.5 ( $P < 0.0001$ ) from preoperative to postoperative, was observed,<sup>[44]</sup> a result perfectly comparable with our systematic survey (36.13,  $P < 0.001$ ) if considering that these authors reported that the mean difference is lower than 50 in patients aged  $<30$  years.<sup>[44]</sup> Interestingly, as a confirm of our meta-analysis, the authors assessed that ROE did not depend on the kind of surgical approach or on sex or postoperative follow-up processes.<sup>[45]</sup> Further studies in very recent publications confirmed this trend.<sup>[46,47]</sup>

## CONCLUSIONS

Despite some limitations, the systematic review and meta-analysis here presented confirm the major role of ROE in assessing rhinoplasty outcomes aside from the different rhinoplasty approaches used in clinics. A higher number of reports should increase the homogeneity of retrieved data, assessing the preliminary results here reported. Looking for a feasible, suitable, easy-to-read, and reliable tool to evaluate patient's satisfaction following rhinoplasty is of major interest for experts in the fields and surgeons, who are endowed with a tool able to provide insights about also the functional perception and outcome of surgical rhinoplasty. Further research is needed to improve ROE usefulness and increase knowledge in rhinoplasty.

## Financial support and sponsorship

Nil.

## Conflicts of interest

There are no conflicts of interest.

## REFERENCES

1. Marinozzi S, Sanese G, Messineo D, Raposio E, Codolini L, Carbonaro R, *et al.* The art of rhinoplasty: Researching technical and cultural foundations of western world rhinosurgery, from the middle ages to the renaissance. *Aesthetic Plast Surg* 2021;45:2886-95.
2. Harb A, Brewster CT. The nonsurgical rhinoplasty: A retrospective review of 5000 treatments. *Plast Reconstr Surg* 2020;145:661-7.
3. Bertossi D, Giampaoli G, Verner I, Pirayesh A, Nocini R, Nocini P. Complications and management after a nonsurgical rhinoplasty: A literature review. *Dermatol Ther* 2019;32:e12978.
4. Johnson ON 3<sup>rd</sup>, Kontis TC. Nonsurgical rhinoplasty. *Facial Plast Surg* 2016;32:500-6.
5. Segreto F, Marangi GF, Cerbone V, Alessandri-Bonetti M, Caldaria E, Persichetti P. Nonsurgical rhinoplasty: A graft-based technique. *Plast Reconstr Surg Glob Open* 2019;7:e2241.
6. Bertossi D, Lanaro L, Dorelan S, Johanssen K, Nocini P. Nonsurgical rhinoplasty: Nasal grid analysis and nasal injecting protocol. *Plast Reconstr Surg* 2019;143:428-39.
7. Liang X, Wang K, Malay S, Chung KC, Ma J. A systematic review and meta-analysis of comparison between autologous costal cartilage and alloplastic materials in rhinoplasty. *J Plast Reconstr Aesthet Surg* 2018;71:1164-73.
8. Kondamuri NS, Miller AL, Rathi VK, Miller L, Bergmark RW, Patel TS, *et al.* Trends in ambulatory surgery center utilization for otolaryngologic procedures among medicare beneficiaries, 2010-2017. *Otolaryngol Head Neck Surg* 2020;162:873-80.
9. Izu SC, Kosugi EM, Brandão KV, Lopes AS, Garcia LB, Suguri VM, *et al.* Normal values for the Rhinoplasty Outcome Evaluation (ROE) questionnaire. *Braz J Otorhinolaryngol* 2012;78:76-9.
10. Bilgin E, Say MA, Baklaçi D. Assessment of patient satisfaction with primary septorhinoplasty using the rhinoplasty outcome evaluation questionnaire. *Cureus* 2020;12:e11777.
11. Metin M, Avcu M. The effect on patient satisfaction of the postoperative nasal topographic, demographic, and functional results of open and closed septorhinoplasty techniques. *J Craniofac Surg* 2021;32:868-73.
12. Izu SC, Kosugi EM, Lopes AS, Brandão KV, Sousa LB, Suguri VM, *et al.* Validation of the Rhinoplasty Outcomes Evaluation (ROE) questionnaire adapted to Brazilian Portuguese. *Qual Life Res* 2014;23:953-8.
13. Khan N, Rashid M, Khan I, Ur Rehman Sarwar S, Ur Rashid H, Khurshid M, *et al.* Satisfaction in patients after rhinoplasty using the rhinoplasty outcome evaluation questionnaire. *Cureus* 2019;11:e5283.
14. Gökçe Küçük S, Arıkan OK. Evaluation of the effects of open and closed rhinoplasty on the psychosocial stress level and quality of life of rhinoplasty patients. *J Plast Reconstr Aesthet Surg* 2019;72:1347-54.
15. Bulut OC, Wallner F, Oladokun D, Kayser C, Plath M, Schulz E, *et al.* Long-term quality of life changes after primary septorhinoplasty. *Qual Life Res* 2018;27:987-91.
16. Floyd EM, Ho S, Patel P, Rosenfeld RM, Gordin E. Systematic review and meta-analysis of studies evaluating functional rhinoplasty outcomes with the NOSE score. *Otolaryngol Head Neck Surg* 2017;156:809-15.
17. Öztürk G. New approaches for the let-down technique. *Aesthetic Plast Surg* 2020;44:1725-36.
18. Robotti E. Invited discussion on: New approaches for the let-down technique. *Aesthetic Plast Surg* 2020;44:1737-41.
19. Robotti E. Invited discussion on: Combination of the pushdown and letdown techniques: Mix-down approaches. *Aesthetic Plast Surg* 2021;45:1150-3.
20. Varoquier M, Rumeau C, Vuissoz PA, Perez M, Hossu G, Jankowski R. Do the upper lateral nasal cartilages exist? The concept of septolateral cartilages. *Eur Ann Otorhinolaryngol Head Neck Dis* 2021;138:77-81.
21. Tawfik GM, Dila KA, Mohamed MY, Tam D, Kien ND, Ahmed AM,



- et al.* A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health* 2019;47:46.
22. Sales JO, Gubisch W, Duarte RR, Moreno AS, Oliveira FM, Coura LM. Caudal extension graft of the lower lateral cartilage: Technique and aesthetic and functional results. *Facial Plast Surg* 2021;37:666-72.
  23. Rabaioli L, Oppermann PO, Pilati NP, Klein CF, Bernardi BL, Migliavacca R, *et al.* Evaluation of postoperative satisfaction with rhinoseptoplasty in patients with symptoms of body dysmorphic disorder. *Braz J Otorhinolaryngol* 2020;S1808-8694(20)30143-9.
  24. Jahandideh H, Maleki Delarestaghi M, Jan D, Sanaei A. Assessing the clinical value of performing CT scan before rhinoplasty surgery. *Int J Otolaryngol* 2020;2020:5929754.
  25. Di Rosa L, Cerulli G, De Pasquale A. Psychological analysis of non-surgical rhinoplasty. *Aesthetic Plast Surg* 2020;44:131-8.
  26. Bracaglia R, Servillo M, Fortunato R, Pino V, Albanese R, Gentileschi S. Rainbow graft for tip reconstruction in revision rhinoplasty. *J Plast Reconstr Aesthet Surg* 2020;73:556-62.
  27. Amer MA, Tomoum MO, Shehata EM, Mandour MF. Effectiveness of chondrofascial 'cigar' graft in contouring the nasal dorsum. *Int J Oral Maxillofac Surg* 2019;48:1552-7.
  28. Amer MA, Tomoum MO. Long-term follow-up of repair of external nasal deformities in rhinoscleroma patients. *J Craniomaxillofac Surg* 2018;46:2138-43.
  29. Bulut OC, Wallner F, Oladokun D, Plinkert PK, Baumann I, Hohenberger R. Patients screening positive for body dysmorphic disorder show no significant health-related quality of life gain after functional septorhinoplasty at a tertiary referral center. *Facial Plast Surg* 2018;34:318-24.
  30. Riedel F, Wähmann M, Veit JA, Bulut OC. Quality of life outcome in revision rhinoplasty in regards to number of revision surgeries and cartilage donor site. *Auris Nasus Larynx* 2022;49:286-90.
  31. Haddady Abianeh S, Moradi Bajestani S, Rahmati J, Shahrbaf MA, Fatehi Meybodi A. Evaluation of aesthetic and functional outcomes after open rhinoplasty: A Quasi-experimental Study by the Aid of ROE and RHINO Questionnaires. *Aesthetic Plast Surg* 2021;45:663-9.
  32. Sazgar AA, Hajjalipour S, Razfar A, Rahavi-Ezabadi S. The effectiveness of skeletal reconstruction in severe post-rhinoplasty nasal deformity. *Aesthet Surg J* 2021;41:905-18.
  33. Sözen T, Dizdar D, Göksel A. Awareness of facial asymmetry and its impact on postoperative satisfaction of rhinoplasty patient. *Aesthetic Plast Surg* 2021;45:214-20.
  34. Ke X, Yang YC, Shen Y, Liu J, Huang JJ, Hong SL. Application of the subjective and objective evaluation in functional rhinoplasty. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 2020;55:223-9.
  35. Fragkos KC, Tsagris M, Frangos CC. Publication bias in meta-analysis: Confidence intervals for Rosenthal's fail-safe number. *Int Sch Res Notices* 2014;2014:825383.
  36. Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979;86:638.
  37. Rosenberg MS. The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 2005;59:464-8.
  38. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
  39. Taylor SM, Rigby MH. The Taylor saddle effacement: A new technique for correction of saddle nose deformity. *J Otolaryngol Head Neck Surg* 2008;37:105-11.
  40. Tastan E, Demirci M, Aydin E, Aydogan F, Kazikdas KC, Kurkcuoglu M, *et al.* A novel method for internal nasal valve reconstruction: H-graft technique. *Laryngoscope* 2011;121:480-6.
  41. Simsek T, Erdoğan MM, Özçetinkaya Erdoğan S, Kazaz H, Tezcan E, Seyhan S. Assessment of functional and aesthetic outcomes in septorhinoplasty. *Eur Arch Otorhinolaryngol* 2021;278:1089-97.
  42. 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;273:2569-73.
  43. Cingi C, Songu M, Bal C. Outcomes research in rhinoplasty: Body image and quality of life. *Am J Rhinol Allergy* 2011;25:263-7.
  44. Arima LM, Velasco LC, Tiago RS. Influence of age on rhinoplasty outcomes evaluation: A preliminary study. *Aesthetic Plast Surg* 2012;36:248-53.
  45. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
  46. Plath M, Sand M, Cavaliere C, Plinkert PK, Baumann I, Zaoui K. How to predict the outcome of septorhinoplasty? A normative study of ROE and FROI-17 scores. *Acta Otorhinolaryngol Ital* 2021;41:327-35.
  47. Bulut OC, Wallner F, Hohenberger R, Plinkert PK, Baumann I. Quality of life after primary septorhinoplasty in deviated- and non-deviated nose measured with ROE, FROI-17 and SF-36. *Rhinology* 2017;55:75-80.