Measuring Cosmetic Facial Plastic Surgery Outcomes

A Pilot Study

Ramsey Alsarraf, MD, MPH; Wayne F. Larrabee, Jr, MD, MPH; Shelia Anderson, RN; Craig S. Murakami, MD; Calvin M. Johnson, Jr, MD

Objective: To test 4 previously published outcomes instruments (the Facelift Outcomes Evaluation, the Rhinoplasty Outcomes Evaluation, the Blepharoplasty Outcomes Evaluation, and the Skin Rejuvenation Outcomes Evaluation) in terms of their reliability and validity in assessing patient-related outcomes of surgical intervention.

Design: A prospective pilot study of 78 patients in 3 similar private cosmetic surgery centers undergoing a total of 100 face-lift, rhinoplasty, blepharoplasty, and skin rejuvenation procedures. Patients were evaluated at 2 preoperative and 1 postoperative time points and the instruments were analyzed with regard to their test-retest reliability, internal consistency, and responsiveness to change.

Results: All 4 outcomes instruments had excellent reliability, consistency, and validity scores. Test-retest reliability was 0.74 to 0.83 (Pearson correlation coefficients), internal consistency scores were .83 to .88 (Cronbach α), and responsiveness to change was statistically significant for each instrument tested (P<.001). In addition, patients experienced significant quality of life improvement, with overall satisfaction increasing on average from 37% to more than 84% after these procedures.

Conclusions: These 4 instruments are reliable and valid and can be used to accurately assess patient-related satisfaction in studies of face-lift, rhinoplasty, blepharoplasty, and skin resurfacing outcomes. These brief questionnaires provide the cosmetic surgeon with quantitative tools to evaluate otherwise subjective and purely qualitative outcomes and are recommended for use in future prospective studies.

Arch Facial Plast Surg. 2001;3:198-201

Quantitative assessment of quality of life outcomes has become standard in most fields of otolaryngology–head and neck surgery, yet to date there has been a paucity of outcomes research activity in the realm of facial plastic and reconstructive surgery.\textsuperscript{1,2} This lack of studies is despite the fact that most outcomes in this field, particularly cosmetic facial plastic surgery, are entirely subjective, patient-related assessments of satisfaction. In recent years, assessment of cost outcomes in facial plastic surgery and general plastic surgery has found a presence in the literature, but measurements of effectiveness have not.\textsuperscript{3-5} There is thus no quantitative manner by which to assess the outcome of cosmetic facial plastic surgery procedures, and the physician and patient alike are left with the qualitative measurement of patient satisfaction as the only real means of analyzing the results of these procedures. To evaluate different surgical techniques, differentiate between various approaches or interventions, or compare the results of different surgeons in an objective fashion, there must be a yardstick by which these outcomes are evaluated.

Application of outcomes research methods to facial plastic surgery thus requires the availability of standardized instruments, or questionnaires, that have been pilot tested to illustrate their reliability and validity in such applications.\textsuperscript{6} One of us (R.A.) previously developed 4 new outcomes instruments for use in the realm of cosmetic facial plastic surgery: the Facelift Outcomes Evaluation (FOE), the Rhinoplasty Outcomes Evaluation (ROE), the Blepharoplasty Outcomes Evaluation (BOE), and the Skin Rejuvenation Outcomes Evaluation (SROE).\textsuperscript{2} Before recommending their use in prospective facial plastic surgery studies, the reliability and validity of these instruments in quantitatively assessing quality of life outcomes for these 4 procedures...
PATIENTS AND METHODS

A total of 78 patients were prospectively enrolled at the 3 similar private facial plastic surgery centers included in this pilot study (the Hedgewood Surgical Center, New Orleans, La; the Larrabee Center for Facial Plastic Surgery, Seattle, Wash; and the Division of Facial Plastic Surgery at Virginia Mason Medical Center, Seattle). A total of 100 of the following 4 procedures were performed: 35 face-lifts, 26 rhinoplasties, 30 blepharoplasties, and 9 skin rejuvenation procedures (carbon dioxide laser resurfacing). For each procedure, an individual outcomes instrument had been previously developed to assess quality of life change in a qualitative manner. Each instrument is composed of 6 questions capturing 3 quality of life domains: physical, mental/emotional, and social. Inclusion of these 3 domains is the recommended methodology in the quality of life literature. Each question is scored on a scale from 0 to 4 and is converted to a total score of 0 to 100 by dividing by 24 and multiplying by 100. These brief instruments take less than 1 minute to complete and are well received by patients during the preoperative and postoperative evaluations. For example, the ROE asks:

How well do you like the appearance of your nose?
How well are you able to breathe through your nose?
How much do you feel your friends and loved one like your nose?
Do you think your current nasal appearance limits your social or professional activities?
How confident are you that your nasal appearance is the best that it can be?
Would you like to surgically alter the appearance or function of your nose?

Patients were asked to complete the appropriate outcomes instrument(s) (1) at the time of their initial consult, (2) on the day of surgery (0.5-9.5 weeks after the initial consult; mean, 3.5 weeks), and (3) approximately 5 months after surgery (11.5-32.0 weeks after surgery; mean, 19.3 weeks). The data from these 3 points were then analyzed to evaluate the reliability and validity of each outcomes instrument. This analysis included assessment of test-retest reliability via Pearson correlation coefficients, internal consistency via the Cronbach α, and validity by measuring responsiveness to change (preoperative vs postoperative scores, paired t test). Statistical analysis was completed using a computer software package (SPSS; SPSS Inc, Chicago, Ill), with significance levels provided in the tables.

Most patients in this study were women, with an average age of 48 years (Table 1). Most patients (78%) underwent a primary procedure, and almost half (46%) underwent multiple procedures at the time of their surgery. Most patients were otherwise healthy, with only 22% documenting any other significant medical history during their preoperative evaluation. Few patients were influenced by marketing (5%), and although the largest proportion were referred to their surgeon by friends (40%), similar percentages were referred by other physicians (33%) or self-referred (27%).

As for the evaluation of instrument reliability, Table 2 provides the test-retest reliability scores (Pearson correlation coefficients) of the 2 preoperative visits for the 4 instruments. The ROE demonstrated the best test-retest reliability (r = 0.83; P < .001), and all 4 instruments had reliability scores of 0.74 or greater. Average preoperative scores ranged from 32.3 to 43.7. Because these scores are measured on a scale from 0 to 100, these values represent a 32% to 44% preoperative patient satisfaction level.

In terms of consistency, each of the 4 instruments demonstrated excellent internal consistency scores (Cron-
retest reliability coefficients were between 0.74 and 0.83 to measure change after surgical interventions. Test- and internal consistency scores and the responsiveness meet these criteria, with excellent test-retest reliability in instruments. It seems that the FOE, ROE, BOE, and SROE surgery depends on the use of reliable and valid outcomes accurate assessment of outcomes in facial plastic surgery.

The FOE, ROE, BOE, and SROE all demonstrated the ability to measure this quality of life change after surgery, with statistically significant increases noted for all 4 procedures (P<.001 for all). Aging face procedures had a similar increase in patient satisfaction (47.8 points) compared with rhinoplasty alone (44.5 points). Blepharoplasty procedures were associated with the greatest increase in patient satisfaction (50.7 points) and rhinoplasty procedures with the least (44.5 points), but compared with the other procedures, these differences were not statistically significant. The highest average postoperative satisfaction score was 89.6 for the blepharoplasty procedures, and the lowest average postoperative satisfaction score was 79.2 for the skin rejuvenation (laser resurfacing) procedures. Face-lift and rhinoplasty procedures had excellent average postoperative satisfaction ratings as well, at 85.9 and 83.3, respectively.

### Table 3. Patient Satisfaction Validity Scores (Responsiveness to Change)*

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Mean Patient Satisfaction Score</th>
<th>Preoperative</th>
<th>Postoperative</th>
<th>Change</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROE</td>
<td>38.8</td>
<td>83.3</td>
<td>+44.5</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>FOE</td>
<td>40.5</td>
<td>85.9</td>
<td>+45.4</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>BOE</td>
<td>38.9</td>
<td>89.6</td>
<td>+50.7</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>SROE</td>
<td>32.0</td>
<td>79.2</td>
<td>+47.2</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Overall satisfaction</td>
<td>37.6</td>
<td>84.5</td>
<td>+46.9</td>
<td>. . .</td>
<td></td>
</tr>
</tbody>
</table>

*ROE indicates Rhinoplasty Outcomes Evaluation; FOE, Facelift Outcomes Evaluation; BOE, Blepharoplasty Outcomes Evaluation; and SROE, Skin Rejuvenation Outcomes Evaluation.

Accurate assessment of outcomes in facial plastic surgery depends on the use of reliable and valid outcomes instruments. It seems that the FOE, ROE, BOE, and SROE meet these criteria, with excellent test-retest reliability and internal consistency scores and the responsiveness to measure change after surgical interventions. Test-retest reliability coefficients were between 0.74 and 0.83 (P=0.03), and Cronbach α scores were between .83 and .88, with scores of .70 or greater generally recommended for adequate instrument reliability and internal consistency in quality of life evaluation. In addition, at 5-month follow-up, patients seemed to experience significant quality of life improvement after each of these 4 procedures, with an average increase in patient satisfaction from 37.6 to 84.5. The FOE, ROE, BOE, and SROE scores all responded dramatically to the significant quality of life change after each surgical procedure, a standard measure of instrument validity (P<.001 for all). Other fields of otolaryngology—head and neck surgery have for years used quantitative tools for the measurement and evaluation of otherwise subjective quality of life outcomes. In the field of facial plastic and reconstructive surgery, particularly cosmetic facial plastic surgery, most outcomes are qualitative assessments of either patient or physician satisfaction. For this reason, it seems that application of outcomes research methods to the field of facial plastic surgery would enhance the study of these results and provide a more rigorous means by which surgical outcomes could be objectively evaluated. Tools such as the FOE, ROE, BOE, and SROE are essential components of this potential application. The use of reliable and valid outcomes instruments might allow the cosmetic facial plastic surgeon to quantitatively assess those results that hitherto have been subject to the oftentimes vague conjecture of “patient satisfaction.”

A recent editorial in the plastic surgery literature calls for an increased effort to move beyond these vague terms and begin to characterize plastic surgery outcomes in a more objective and quantitative manner. In this editorial, Luce stresses the importance of using tested and validated instruments such as these to achieve this end. The difficulty in attaining this goal, in Luce’s mind, is the quantification of the various domains that have an impact on quality of life. Our hope, with this study, is to provide the starting point for this quantification. It is clear that these outcomes measures rely on the subjective evaluation of the patient and that this evaluation might conflict with the surgeon’s own assessment of the success of a given surgical procedure. This fact, however, does not limit the usefulness of such patient-related measures but implicates their intrinsic worth. As discussed elsewhere, the overriding purpose of outcomes research is to quantify and measure these otherwise subjective and qualitative aspects of patient care. For cosmetic facial plastic surgery, this analysis must begin with the use of reliable and valid instruments that measure not the surgeon’s evaluation of success or failure but the patient’s own estimation of this inherent value. In fact, the ever-present reliance on terms such as “patient satisfaction” in the current plastic surgery literature as a measure of a successful surgical outcome emphasizes the manner in which our specialty, to a large extent, already relies on such subjectively based assessments to determine a given procedure’s success.

Limitations of this study include small numbers in the skin resurfacing population (n=9); however, the SROE still demonstrated good reliability, excellent internal consistency, and a high level of responsiveness to change. The other 3 instruments, with higher numbers of patients studied, tended to demonstrate higher degrees of reliability and validity. Five-month follow-up might not be long enough to capture the ultimate quality of life changes associated with procedures such as rhinoplasty or laser skin resurfacing; however, this period seems to demonstrate significant improvements that might be long lasting. The purpose of this study, however, was not to
assess the success or failure of these specific procedures or individual surgeons or distinct techniques but, rather, to assess the reliability and validity of these 4 new outcomes instruments. For this reason, a more limited follow-up period such as 5 months should represent enough clinical change to allow the testing of such instruments in terms of the basic outcomes research tenets of test-retest reliability, internal consistency, and responsiveness to change. Because we were not comparing the outcomes of different surgical techniques in this pilot study, the use of 3 distinct surgeons should not bias these results in any way. In addition, there were no significant differences between the 3 surgeons included in this pilot study with regard to the measures of instrument reliability and validity evaluated in this analysis.

Although this study provides initial validation data with regard to improvements in patient satisfaction and quality of life after face-lift, rhinoplasty, blepharoplasty, and skin resurfacing procedures, the purpose of this study was, again, not to evaluate these procedures themselves but to demonstrate the ability of these instruments to accurately characterize these changes. The vast improvement in patient-related quality of life that was quantitatively documented, however, seems to parallel the qualitative satisfaction that most cosmetic facial plastic surgery patients enjoy after their individual procedures. With the use of reliable and valid outcomes instruments such as the FOE, ROE, BOE, and SROE, the subjective assessment of patient satisfaction can now find a more objective method of analysis for the cosmetic facial plastic surgery community as a whole.

CONCLUSIONS

The FOE, ROE, BOE, and SROE are reliable and valid outcomes instruments that accurately characterize patient-related quality of life satisfaction. These 4 questionnaires are useful for the quantitative measurement and assessment of cosmetic facial plastic surgery outcomes and provide the surgeon with a more objective means of comparing various treatments, surgical techniques, or approaches. Use of these instruments is recommended for such quantitative evaluation in future studies.

Accepted for publication March 28, 2001.

Corresponding author: Ramsey Alsarraf, MD, MPH, The Newbury Center, Cosmetic Facial Plastic Surgery, 69 Newbury St, Boston, MA 02116 (e-mail: ralsarraf@earthlink.net).

REFERENCES