Realistic Expectations: To Morph or Not to Morph?

Foot

Anurag Agarwal, M.D.
Edward Gracely, Ph.D.
William E. Silver, M.D.

*Naples, Fla.; Philadelphia, Pa.; and
Atlanta, Ga.*

Background: The purpose of this study was to determine whether computer-morphed images convey realistic expectations of the postoperative cosmetic result to patients seeking rhinoplasty.
Methods: This was a retrospective study in which preoperative, computer-morphed, and 1-year postoperative images of 25 consecutive rhinoplasty patients were analyzed by 12 anonymous, skilled observers and by the operating surgeon in a double-blind fashion. Each patient's photographs were accompanied by an identical four-question survey. The survey assessed whether the morphed image represented an obtainable goal, the degree of similarity between the morphed image and the actual postoperative result, and the quality of the surgical outcome as compared with the computer image.
Results: As expected based on variations in individual surgeons' aesthetic opinions, the cumulative data from the 12 blinded raters across all patients indicated that 82 percent of the raters would have performed the imaging in either the same way or, at most, a "slightly different" way from the operating surgeon. Approximately half thought that the computer-morphed image did realistically predict the postoperative result but that it could be a bit closer. Seventy-five percent felt that the actual surgical result was either the same or better than the predicted result.
Conclusions: Computer imaging for patients undergoing rhinoplasty does portray a realistic picture of the anticipated cosmetic result. However, because an average of 32 percent of respondents felt that the computer-morphed image did not predict the postoperative result, conservative imaging is encouraged to prevent false expectations. (*Plast. Reconstr. Surg.* 119: 1, 2007.)

The role of computer imaging in facial plastic surgery as a valuable communication tool has been well established. Patient satisfaction with cosmetic surgery after computer imaging has been documented to be higher than in those patients who did not receive imaging. This mode of communication often facilitates discussion about specific goals of the procedure and helps uncover any potential unrealistic expectations that the patient may have.[1]

However, on a daily basis, patients ask whether their morphed image conveys a realistic expec-

tation of the postsurgical outcome. Few studies in the literature exist that attempt to answer this question. In 1994, Bronz reported 100 consecutive cases of primary rhinoplasty with 1-year follow-up.[2] All of the patients who were operated on were shown their planned preoperative computer screen picture and asked to compare this to their postoperative image, on lateral view only. To objectively measure any differences in dorsal profile lines, both images were solarized and then overlapped. Despite differences in the two profile lines for 38 percent of the patients, each of the 100 patients considered the postoperative result to be identical to the planned image. One of the limitations of this study is that only the lateral view was shown, which offers a one-dimensional image of the nose. An open roof deformity, inverted-V deformity, crooked nose, or tip asymmetry may be missed without a frontal view. In addition, with any nonblinded survey of postoperative patients, there is an inherent bias possible based on the character of the

balt5/zpr-prs/zpr-prs/zpr00407/zpr0560-07z | xppws | S=1 | 1/24/07 | 15:24 | 4/Color Figure(s): F1-6 | Art: 180521 | Input-nlm

Plastic and Reconstructive Surgery • April 1, 2007

doctor–patient relationship and the patient's satisfaction or dissatisfaction with the surgical result.

Another study published in 1998 by Vuyk et al. evaluated 50 patients on whom a variety of procedures were performed.[3] It is not clear whether consecutive patients or selected patients were included in the study. All patients were surveyed with a mailed questionnaire accompanying their original computer-simulated image at a mean postoperative period of 17 months. The authors recognize that a patient's perception of the final surgical result, and the length of time elapsed since surgery, may influence their response on the questionnaire. The actual postoperative photographs of the patients were not sent to the patients but were compared with the computer-morphed images by the operating surgeon (average 8-month postoperative photographs). A lateral view was used throughout the evaluation. For rhinoplasty, the representative value of the computer predictions was found to be 83 percent.[3]

Sharp et al. analyzed responses to questionnaires administered to 56 consecutive rhinoplasty patients, 25 of whom had undergone preoperative computer imaging.[4] In their study, 64 percent of patients found the postoperative result to be equally or more aesthetic than predicted. Multiple photographic views were used when performing the imaging initially. However, the actual computer-morphed or postoperative images were not sent to the patients, who had to rely on memory to recall the preoperative computer predictions.

Despite the heterogeneity of available studies on the predictability of computer imaging, certain themes are consistently mentioned in these studies: (1) the patients should be informed that the computer images are simulations, and not a guarantee of the surgical result; (2) the operating surgeon should project morphed images within the confines of his or her surgical abilities; and (3) computer imaging facilitates the exchange of ideas.

The purpose of this study was to determine whether the computer-generated image in patients seeking rhinoplasty is an accurate predictor of the postoperative result. Although any assessment of this parameter is essentially subjective, we tried to eliminate any confounding factors. Because the previously outlined studies have either surveyed the operating surgeon or the patients themselves, independent double-blinded skilled raters were used. To highlight the potential bias that is inherent when asked to evaluate one's own surgical results, the first author (A.A.) administered an identical survey with the accompanying images to the operating surgeon (W.E.S.). In addition, a general survey assessing the role of computer imaging in a facial plastic surgeon's practice was sent to each blinded rater.

PATIENTS AND METHODS

Twenty-five consecutive rhinoplasty patients operated on by the senior surgeon (W.E.S.) for whom 1-year postoperative photographs were available were used for the study. All patients had their computer imaging performed by the operating surgeon using the Uniplast imaging software at the time of initial consultation. Proposed changes were demonstrated on both frontal and right lateral views in the Frankfort plane. The patients included those who presented for purely cosmetic changes and those who desired aesthetic improvements in conjunction with functional surgery. Their ages ranged from 17 to 60 years. Of the 25 patients, 23 were women and two were men. There were 19 Caucasians, four African Americans, and two South Asians.

Seventeen packets were then assembled, each of which included an anonymous cover letter with instructions on how to fill out the survey; a general survey designed to assess the role of computer imaging in the blinded rater's practice (Table 1); and the preoperative, computer-morphed, and 1-year postoperative images of each of the 25 consecutive patients (Fig. 1). Each patient's photo-

AQ: 1

T1

F1

Table 1. General Survey on the Cover Letter for Each of the 17 Packets

1. Do you currently use computer imaging software in your practice? (If No, then skip to the next page)
   Yes   No
2. If you do use computer imaging, do you do it yourself or is it done by someone else in the practice?
   Myself   Nurse   Office Administrator   Fellow   Other
3. If you do use computer imaging, do you give a printed copy of the image to your patients?
   Yes   No
4. If you do give the image to a patient, do you annotate it with a disclaimer (i.e., "This is intended as a communication tool only," etc.)
   Yes   No

Fig. 1. Sample of one of the 25 patients' preoperative (*left*), morphed (*center*), and 1-year postoperative (*right*) images (patient 5). In response to question 1, 50 percent of the raters felt that this morphed image was definitely surgically obtainable and 50 percent thought the image was probably obtainable. For question 2, 75 percent would have performed the imaging in a similar or identical fashion. Seventy-five percent of the raters felt that the computer imaging either realistically predicted the postoperative result or could be slightly closer for this patient. Thirty-three percent said that the actual result was better than predicted, and 25 percent said that the result was the same as imaged.

graphs were accompanied by an identical four-question survey (Table 2). This survey assessed whether the morphed image represented an obtainable goal, the degree to which the morphed image predicted the actual postoperative result, and the quality of the surgical outcome as compared with the computer image. There was no identifying patient or surgeon data within any packet.

Sixteen identical packets were then sent to a liaison at the American Academy of Facial Plastic and Reconstructive Surgery to perform the study in a double-blind fashion. This liaison then distributed the packets to 16 American Academy of

balt5/zpr-prs/zpr-prs/zpr00407/zpr0560-07z | xppws | S=1 | 1/24/07 | 15:24 | 4/Color Figure(s): F1-6 | Art: 180521 | Input-nlm

Plastic and Reconstructive Surgery • April 1, 2007

Table 2. Survey That Accompanied Each of the 25 Patients' Photographs

Computer Imaging Survey (circle one)

1. In your hands, does the morphed image represent a surgically obtainable result?
   Probably not   Not sure   Probably   Definitely
2. If you were performing the morphing (patient requests aside), how similar would your computer-generated image be to the one displayed here?
   Completely different   Slightly different   Similar   The same
3. Did the computer-morphed image realistically predict the postoperative result?
   No   Yes, but could be a bit closer   Yes, fully
4. In your opinion, how does the actual postoperative result compare to the morphed image?
   Worse   Same   Better

Facial Plastic and Reconstructive Surgery fellowship directors, 12 of whom returned completed packets to the liaison. These 12 packets were then mailed back to the corresponding author (A.A.), without any identifying data as to the identity of the blinded raters. The seventeenth packet was given to the operating surgeon (W.E.S.).

Across all blinded raters and patients, the percentage of answers to each of the four questions was determined. If all 12 blinded raters answered a question for each of the 25 patients, there would be 300 responses to that question in the denominator (12 × 25 = 300). As an example, for question 1, there were a total of 139 "probably" responses among 300 total responses to this question (46 percent). In addition, specific patients who were determined by the blinded raters to have the most discrepancies between their computer-morphed image and actual result were highlighted. This was done to assess whether certain patient characteristics may predict a poor correlation between the morphed image and the postoperative outcome, suggesting a limitation to computer imaging for patients with these features.

Finally, a statistical analysis was performed to assess the test-retest reliability of the survey. Intraclass correlation coefficients were calculated for the 12 blinded raters as a set and for individual raters. Pearson and Spearman correlations between individual raters were then calculated to confirm the findings from the single rater intraclass correlation coefficients.
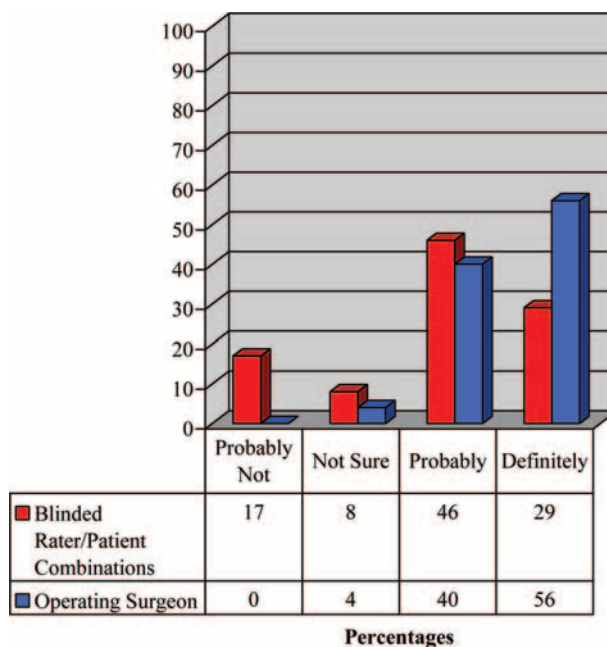
RESULTS

Of the 12 raters, 75 percent use computer imaging within their practice. Of those surgeons who use computer imaging, 78 percent perform the morphing themselves, whereas 22 percent have the imaging performed by another person within the practice (i.e., nurse, fellow, other). Sixty-seven percent of those who use imaging software give printed copies of the image to their patients, and all of them are annotated with a disclaimer (e.g., "This is intended as a communication tool only").
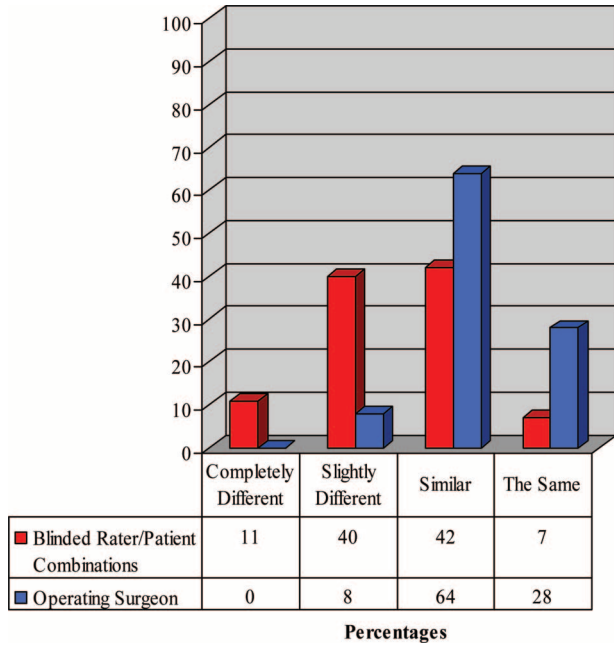
The cumulative total responses across all 12 blinded raters and 25 patients to each of the four questions are shown as percentages in Figures 2 through 5 as speckled black columns. The data obtained from the operating surgeons' surveys is depicted as speckled white for comparison. The percentage of patient/rater responses to question 1 indicates that the morphed image was probably (46 percent) or definitely (29 percent) a surgically obtainable result. In question 2, responses suggest that they would have performed the morphing in a slightly different fashion for 40 percent of the patients but in a similar way for 42 percent of the patients. Figure 4 reveals that as a group, the 12 raters thought that 50 percent of the computer images did realistically predict the postoperative

	Probably Not	Not Sure	Probably	Definitely
Blinded Rater/Patient Combinations	17	8	46	29
Operating Surgeon	0	4	40	56

Percentages
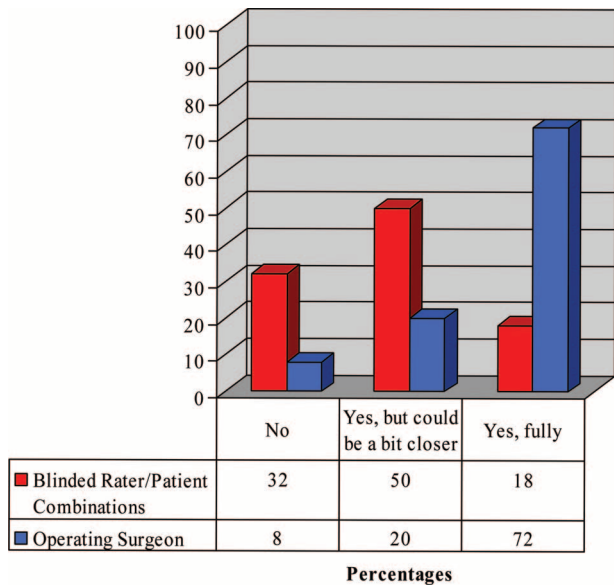
Fig. 2. Chart of survey results to the question, "In your hands, does the morphed image represent a surgically obtainable result?"

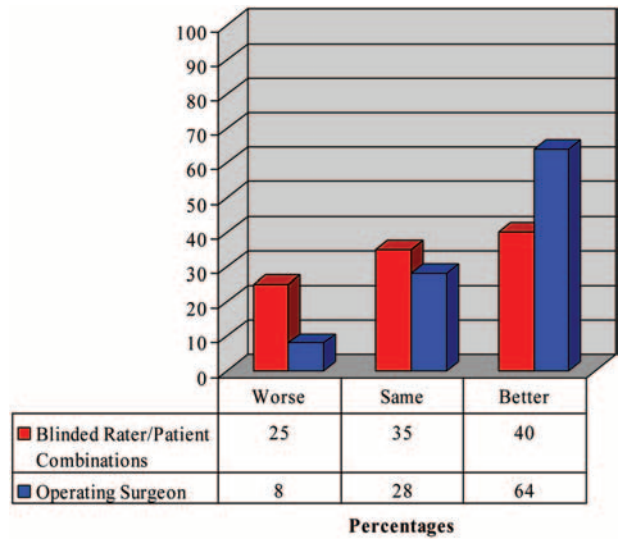|	Completely Different	Slightly Different	Similar	The Same
■ Blinded Rater/Patient Combinations	11	40	42	7
■ Operating Surgeon	0	8	64	28

Percentages

Fig. 3. Chart of survey results to the question, "If you were performing the morphing (patient requests aside), how similar would your computer-generated image be to the one displayed here?"

	No	Yes, but could be a bit closer	Yes, fully
■ Blinded Rater/Patient Combinations	32	50	18
■ Operating Surgeon	8	20	72

Percentages

Fig. 4. Chart of survey results to the question, "Did the computer-morphed image realistically predict the postoperative result?"

result, but could be a bit closer. In 18 percent, the morphed images were fully accurate, but in 32 percent, they were not predictive. Based on Figure 5, it is clear that the blinded raters felt that the actual surgical result was the same as (35 percent) or better than (40 percent) the computer image.

	Worse	Same	Better
■ Blinded Rater/Patient Combinations	25	35	40
■ Operating Surgeon	8	28	64

Percentages

Fig. 5. Chart of survey results to the question, "In your opinion, how does the actual postoperative result compare with the morphed image?"

However, the postoperative result was worse than the predicted image 25 percent of the time.

When comparing individual patients based on the tabulated survey results, four patients stood out (patients 8, 9, 12, and 15). The 12 blinded raters felt that in their hands, the morphed images for these patients would probably not be surgically obtainable, that they would have performed the imaging in a completely different manner, that the morphed images did not realistically predict the postoperative result, and that the actual surgical outcome was worse than the projected result (Fig. 6).

The intraclass correlation coefficient for each of the four questions is shown in Table 3. Both the average measures for the group of 12 blinded raters and the correlation between individual raters are shown, with their respective 95 percent confidence intervals. When evaluating intraclass correlation coefficients, 0 indicates a complete lack of agreement between raters, whereas 1.0 would indicate identical responses. Using the intraclass correlation coefficients for the 12 raters reveals that as a set, their average response to an image is quite reliable. Pearson and Spearman's rho correlations were also calculated, which confirmed the low correlation noted by the low intraclass correlation coefficient for individual raters.

DISCUSSION

Computer imaging has become a routine part of our facial plastic surgery practice. It is used to

F6

T3

Fig. 6. Preoperative (*left*), morphed (*center*), and 1-year postoperative (*right*) images of patient 9. The blinded raters felt that significant discrepancies between the morphed image and actual surgical result were present.

simulate results for patients undergoing rhinoplasty, endoscopic forehead lifting, cervicofacial rhytidectomy, chin augmentation, lip augmentation, and upper and lower eyelid blepharoplasty. Perhaps the greatest benefit is in reconciling any subtle differences between the surgeon's aesthetic opinion and the patient's desired result. This improvement in preoperative communication benefits both parties. One of the advantages of having the surgeon perform the imaging is that the surgeon has the best concept of what he or she can achieve surgically.[1] We also feel that that it is a critical tool for preoperative rhinoplasty analysis, assisting the surgeon in planning surgical maneuvers. We routinely give a printout of the preoperative frontal and lateral photographs to the patient, along with the corresponding morphed views. A disclaimer statement is included on the photographs. The patients then have a chance to review the proposed changes after the initial consultation in a more relaxed setting, and with their friends and family.

balt5/zpr-prs/zpr-prs/zpr00407/zpr0560-07z | xppws | S=1 | 1/24/07 | 15:24 | 4/Color Figure(s): F1-6 | Art: 180521 | Input-nlm

Volume 119, Number 4 • Assessment of Morphed Images

Table 3. Intraclass Correlation Coefficient*

Question	Measures	ICC†	95% Confidence Interval Lower Bound	Upper Bound
1	Individual	0.456‡	0.313	0.635
	Average	0.910	0.845	0.954
2	Individual	0.257‡	0.144	0.433
	Average	0.792	0.645	0.894
3	Individual	0.336‡	0.210	0.518
	Average	0.859	0.761	0.928
4	Individual	0.160	0.076	0.309
	Average	0.696	0.494	0.843

ICC, intraclass correlation coefficient.
*Two-way random effects model where both people effects and measures effects are random.
†Type A intraclass correlation coefficients using an absolute agreement definition.
‡The estimator is the same, whether the interaction effect is present or not.

However, on several occasions throughout the clinic day, patients have asked the following critical question: "How close can you get to this morphed image?" Although a few studies in the literature have attempted to answer this, it is difficult to compare them because of different methodologies. We have tried to eliminate the biases inherent in evaluating one's own surgical results and those associated with patients' responses to questionnaires, because they may be influenced by their own perception of their result. In fact, Figure 4 demonstrates that the operating surgeon felt that the computer-morphed image fully predicted the postoperative result in 72 percent of the patients, compared with only 18 percent when the percentages of cumulative total responses across all patients and the 12 American Academy of Facial Plastic and Reconstructive Surgery fellowship directors were calculated. In addition, the operating surgeon felt that in 64 percent of the patients, the actual result was better than the planned result, compared with 40 percent when judged by the blinded raters. Several factors may influence a surgeon's perception of his or her results, including his or her individual aesthetic opinion and knowledge of the patient's satisfaction or dissatisfaction with the result.

The choice of American Academy of Facial Plastic and Reconstructive Surgery fellowship directors as blinded raters was made because they are experienced facial plastic surgeons who could provide an independent, critical assessment of the photographs. Seventy-five percent of the 12 raters who responded use computer imaging software on a routine basis, with the majority performing the morphing themselves. Of those who perform imaging, only two-thirds provide a hard copy to their patients. A disclaimer statement is always included. In several articles addressing the legal issues of computer imaging, the risk of liability has been shown to be quite low, provided that certain recommendations are followed: computer imaging should be supplemented with standard informed consent forms; computer images must be retained as part of the medical record and not destroyed; conservative morphing, perhaps even to demonstrate less favorable outcomes, is worthwhile; and it seems prudent to include disclaimer statements that indicate that the computer images do not guarantee the actual surgical result.[5,6]

In designing the survey to accompany the 25 patient photographs, questions were geared to elicit whether the operating surgeon conveyed realistic projections to his or her patients and to highlight the variations in aesthetic opinions between the facial plastic surgeons. The percentage of patient/rater responses for question 1 as depicted in Figure 2 reveals that most of the blinded raters (75 percent) agreed with the operating surgeon and thought that the morphed image was probably or definitely a surgically obtainable result. This indicates that unattainable surgical outcomes were not projected to the patients.

As expected based on individual surgeons' preferences, responses to question 2 indicate that without knowledge of the patient's requests, most (82 percent) of the blinded raters would have performed the morphing in a slightly different or similar manner. Very few would have imaged the patient completely different or exactly the same as the operating surgeon. Therefore, the operating surgeon's aesthetic goals are generally congruent with those of the 12 American Academy of Facial Plastic and Reconstructive Surgery fellowship directors.

Question 3 is perhaps the most important of the survey. Approximately half of the cumulative total responses across all patients and raters indicate that the morphed image did realistically predict the postoperative result but that it could have been a bit closer, whereas only 18 percent felt that it was a completely accurate predictor. Of note, 32 percent thought that the morphed image did not realistically predict the surgical outcome. Comparison between these data and the literature is not possible. Prior studies have assessed whether the postoperative result is more or less aesthetic than the predicted image and equated this to predictive value. For example, Vuyk et al. found that 83 percent of computer predictions were repre-

sentative, because these surgical results were judged as identical to or even more aesthetic than the morphed images.[3] However, in our study, question 3 assessed the predictability of computer imaging independent of whether the blinded raters thought the image was better or worse than the outcome. Instead, it is Figure 5 that demonstrates that the actual postoperative result was felt to be the same (35 percent) or better (40 percent) than the morphed image. This latter number is comparable to prior studies in which the surgical result was felt to be more aesthetic than the predicted image (37 percent in Vuyk et al.[3] and 44 percent in Sharp et al.[4]).

Further analyses of the data were then performed, this time categorizing responses across the surveys by patient. The raters found that the morphed images for patients 8, 9, 12, and 15 were probably not surgically obtainable, and that they would have performed the imaging completely differently for these patients. The imaging did not realistically predict the postoperative result in these patients, and their actual surgical result was generally worse than the morphed image. After reviewing the photographs of these four patients, one feature stood out: three of the four patients were African American. The fourth was Caucasian. Although the numbers are small, it seems that there is significant discrepancy between what is considered aesthetically ideal in this patient population. In patients 8, 9, and 15 (African American women), the nostril base width was morphed and reduced to the intercanthal width, which may have been considered by some raters as a less ethnic, more Caucasian feature. Of the remaining 21 patients in whom significant concordance existed between the raters and the operating surgeon, only one was African American.

The test-retest reliability was then calculated for each rater across the entire set of patients by question. Correlations between individual raters were low to moderate (intraclass correlation coefficient, 0.16 to 0.456), indicating considerable variability in responding. However, the average response of the 12 raters is likely to be reliable and reproducible. Thus, the average of the 12 provides a good estimate of the perceived quality of the image compared with expectation, with average measures of intraclass correlation coefficient for the set of raters ranging from 0.696 to 0.910.

Despite our attempts to eliminate sources of bias by using independent, skilled observers in a double-blind fashion, there are some limitations to this type of study. The 25 consecutive patients included in the study were those who returned for at least 1-year follow-up photographs. In the senior surgeon's practice, postoperative photographs are taken at the 6-month, 1-year, and annual follow-up visits thereafter. However, approximately 25 percent of postoperative rhinoplasty patients fail to return for follow-up after their 3- or 6-month visit. Therefore, a truly consecutive sample could not be obtained. Furthermore, because frontal and lateral views were included, there is more potential for discrepancy between the morphed image and the postoperative result, compared with prior studies in which a lateral view only was used. However, we feel that both views are important when counseling patients. Furthermore, the use of American Academy of Facial Plastic and Reconstructive Surgery fellowship directors as blinded raters is not necessarily representative of the patient's perspective. We sought the most critical observers to conservatively underestimate the predictive value of computer imaging. As noted by Sharp et al., failure to surgically attain the "ideal" nose generated by the computer image does not necessarily result in poor overall patient satisfaction.[4]

CONCLUSIONS

Computer imaging for patients undergoing rhinoplasty does portray a realistic picture of the anticipated cosmetic result. Sixty-eight percent of patients' morphed images were either fully accurate or, more likely, good predictors of the surgical outcome, which could be a bit closer. Three of every four postoperative results were judged to be the same or even better than the computer prediction by 12 independent, critical observers. However, 25 percent of the time, the actual result was rated worse than the morphed image. Also, because the cumulative totals across all blinded raters indicate that 32 percent of the computer-morphed images did not predict the postoperative results, patients must be made aware of the limitations of computer imaging. In our study, when analyzing the responses for individual patients, the most heterogeneity and the least accurate predictive value of computer imaging was noted to be in the African American patients.

Anurag Agarwal, M.D.
Aesthetic Surgery Center
11181 Health Park Boulevard, Suite 1115
Naples, Fla. 34110
anurag57@hotmail.edu

DISCLOSURE

None of the authors has any financial interest in any of the products, devices, or drugs mentioned in this article.

REFERENCES

1. Hopping, S. *Image thyself. Facial Plast. Surg.* 7: 45, 1990.
2. Bronz, G. Predictability of the computer imaging system in primary rhinoplasty. *Aesthetic Plast. Surg.* 18: 175, 1994.
3. Vuyk, H., Stroomer, J., and Vinayak, B. The role of computer imaging in facial plastic surgery consultation: A clinical study. *Clin. Otolaryngol.* 23: 235, 1998.
4. Sharp, H., Tingay, R., Coman, S., et al. Computer imaging and patient satisfaction in rhinoplasty surgery. *J. Laryngol. Otol.* 116: 1009, 2002.
5. Chavez, A., Dagum, P., Koch, R., et al. Legal issues of computer imaging in plastic surgery: A primer. *Plast. Reconstr. Surg.* 100: 1601, 1997.
6. Mavroforou, A., Giannoukas, A., and Michalodimitrakis, E. Medical litigation in cosmetic plastic surgery. *Med. Law* 23: 479, 2004.