ORIGINAL ARTICLE

The predictability of maxillary repositioning in LeFort I orthognathic surgery

Ron Jacobson, DDS, MS,[a] and David M. Sarver, DMD, MS[b]
Chicago, Ill, and Birmingham, Ala

Our ability to predict the outcome of any orthognathic procedure relies on the surgeon's ability to accurately reproduce the desired skeletal movements and on our understanding of the soft tissue changes associated with those movements. The purpose of this retrospective study was to evaluate the surgical accuracy of maxillary repositioning by comparing the objectives obtained from cephalometric prediction tracings with the actual skeletal changes achieved during maxillary and maxillomandibular procedures. The sample consisted of 46 patients from the files of 1 orthodontist. Presurgical and immediately postsurgical cephalometric radiographs were digitized, and the original surgical prediction was reproduced with Dentofacial Planner (Dentofacial Software, Toronto, Ontario, Canada) software. Vertical and horizontal measurements to several skeletal landmarks were used to assess the differences between the predicted maxillary position and the actual maxillary postsurgical position. Statistical differences were found for some measurements, particularly those related to the vertical placement of the posterior maxilla. Three variables were evaluated for their effect on the surgical discrepancies. The 2 surgeons with the most patients were evaluated and found to have differences in the direction of discrepancies but not the amount (absolute values). Significant differences were also found when evaluating surgical complexity (single-jaw vs bimaxillary procedures) and the direction of movement (impaction vs advancement) in direction only. To assess the overall fit of individual predictions, we calculated an average discrepancy for each patient; 80% of the actual results fell within 2 mm of the prediction, and 43% fell within 1 mm of the prediction. (Am J Orthod Dentofacial Orthop 2002;122:142-54)

For over 30 years, orthognathic surgery has proved to be a significant tool in correcting severe dentofacial skeletal discrepancies. The combined efforts of orthodontists and oral surgeons have dramatically improved the quality of life for patients with functional and esthetic disharmonies.

Now, more than ever, the esthetic outcome of the surgical procedure is a major factor in determining its success, because patient expectations are very high. Since their introduction over 10 years ago, software programs used to alter patient images to predict the outcome of surgical procedures have been refined and now have greatly improved graphics and impressive interfaces. Printouts of surgical predictions have reached photographic quality. However, there has been an ongoing controversy, considering the litigious nature of our society, about whether these predictions should be shown to patients. Some fear that predictions might imply a guaranteed outcome.[1]

[a]Assistant Clinical Professor, Northwestern University.
[b]Adjunct Professor, University of North Carolina.
Reprint requests to: Dr Ron Jacobson, 4200 W Peterson Ave, No 116, Chicago, IL 60646; e-mail, drron@jacobsonortho.com.
Submitted, May 2001; revised and accepted, March 2002.

0889-5406/2002/$35.00 + 0 **8/1/125576**
doi:10.1067/mod.2002.125576

A major benefit of surgical prediction is its use as a communication tool among all involved persons. The amount of esthetic change desired by a patient can dramatically alter the treatment plan. It can determine the need for bimaxillary versus single-jaw procedures and whether adjunctive procedures such as genioplasty, rhinoplasty, or liposuction are necessary. The esthetic goal also determines the direction of presurgical orthodontic preparation.[2]

As orthodontists, we look at both the occlusion and the esthetics when evaluating the success of a surgery. However, the best occlusion will not satisfy a patient who is unhappy with the esthetic outcome. Every patient has a unique perspective of beauty, and some patients have strong feelings about whether a given result is acceptable. For example, the patient in Figure 1 presented with mandibular retrognathia. The orthodontist and the surgeon agreed that a mandibular advancement would be appropriate treatment. The orthodontist asked the patient to posture the mandible forward until the anterior teeth were in contact, a simple form of prediction, and the patient was pleased with the photograph (Fig 1, *B*). Without consulting the patient or the orthodontist, the surgeon, knowing that the patient's chief complaint was a weak chin, decided to advance the chin an additional 15 mm. The patient was not happy with the result. We cannot arbitrarily assume that patients and surgeons have the same

Fig 1. A, Presurgical cephalometric radiograph; **B**, prediction (posturing); **C**, actual result.

ideal esthetic goal as we do. After all, "beauty is in the eye of the beholder."

The importance of surgical prediction has been the subject of many orthodontic articles, presentations, and textbook chapters.[1-22] The accuracy of soft tissue prediction has been and continues to be an area of strong interest in both the orthodontic and surgical literature. Hing,[3] Jacobson,[4] Konstiantos et al,[5] Sinclair et al,[6] Aharon et al,[7] and Upton et al[22] have studied the accuracy of the soft tissue profiles generated by the 3 major software programs: Dentofacial Planner (Dentofacial Software, Toronto, Ontario, Canada) and Quick Ceph (Quick Ceph Systems, San Diego, Calif). These studies concentrated on the accuracy of the mathematics used to calculate the soft tissue changes by eliminating surgical accuracy and relapse from the variables.[3-7,22] The actual skeletal changes achieved during the surgery were measured and then reproduced in the prediction. The predicted profile was then compared with the surgical result, and the differences at key landmarks were measured. Although this technique provides an excellent test of the mathematics, it assumes that the achieved skeletal change matches the original goal. This method of evaluation tests software accuracy, but it does not answer a very important question: Did the treatment team actually achieve the treatment goals determined by computer prediction?

Three factors influence the patients' perceptions of the esthetic outcome of their procedures: their involvement in the treatment planning and their understanding of what is planned, the accuracy of the computer prediction software, and the surgical accuracy of the treatment team.[23,24] This study was designed to eliminate the first 2 factors as variables and to measure the surgical accuracy.

Recent studies have shown that despite variations, a single-jaw mandibular advancement is a predictable procedure. Specifically, there was no statistical difference in the position of the skeletal hard tissue between the ideal prediction and the actual result.[8,9] Gerbo et al[8] analyzed the accuracy of a prediction algorithm from Quick Ceph by comparing the predicted and actual results of 35 patients treated with mandibular advancement and setback surgery. Unlike the previous studies, the predictions were generated by moving the mandible parallel to the occlusal plane so that the mandibular incisors were positioned in an acceptable overjet and overbite relationship with the maxillary incisors. The Student *t* test showed that 10 of the 16 hard and soft tissue measurements did not differ significantly. Because the noted differences were less than 1.8 mm or 3.1°, the authors concluded that the magnitude of the differences were within clinically acceptable limits.[8]

Powers[9] conducted a similar study evaluating hard and soft tissue predictability in 39 cases of mandibular advancement surgery by a single surgeon. She also used ideal incisal relationships as the guide for prediction and found no significant difference in the results of the skeletal mandibular landmarks in a vertical or horizontal direction. However, statistically significant differences between actual and predicted mandibular plane angle existed that could be partially attributed to difficulty in the location and transfer of gonion on the cephalometric radiographs. No significant differences were found in the prediction of the soft tissue landmarks, with the exception of the lower lip and the vertical position of the chin. In judging clinical acceptability, Powers[9] considered errors greater than 3 mm significant and found that the upper lip and labiomental fold horizontally were very predictable with only a few patients (3% and 8%, respectively) exhibiting clinically significant errors. On the other hand, the lower lip exhibited clinically significant errors both vertically and horizontally (38% and 18%, respectively). Therefore, this study supports previous findings that, in general, predictions are fairly accurate in simulating both the hard tissue changes and the resultant soft tissue profile, except for the lower lip.

Surgical procedures involving the maxilla, either alone or combined with mandibular osteotomies, can be very complex and more variable. The ideal placement of

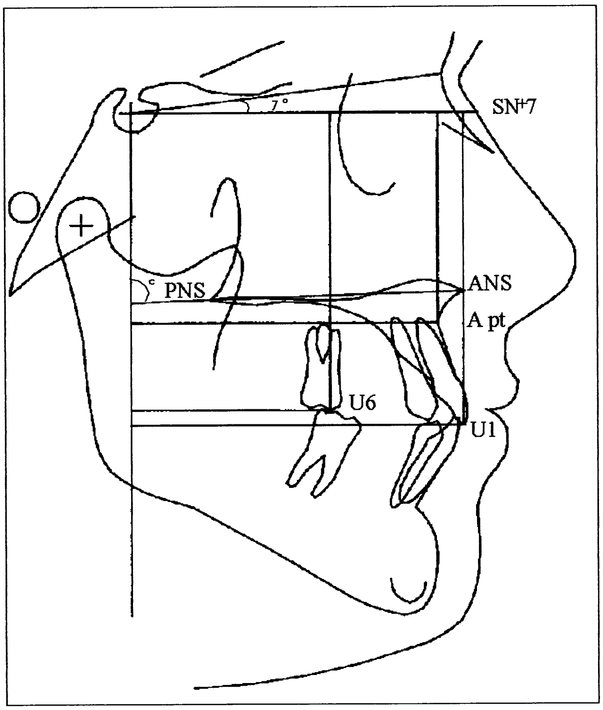

Fig 2. Custom analysis. Linear measurements (mm): A-point horizontal to $SN^{+}7$ perpendicular; A-point vertical to $SN^{+}7$; maxillary central incisor horizontal to $SN^{+}7$ perpendicular; maxillary central incisor vertical to $SN^{+}7$; maxillary first molar horizontal to $SN^{+}7$ perpendicular; maxillary first molar vertical to $SN^{+}7$. Angular measurement: palatal plane to $SN^{+}7$ perpendicular.

the maxilla can be a subject of debate and a matter of opinion. The vertical and horizontal placement of the maxilla will determine the amount of mandibular autorotation and the need for a mandibular osteotomy or adjunctive genioplasty. Ultimately, the placement of the maxilla during the surgical procedure has a significant effect on the esthetic result.

The predictability of maxillary surgery is influenced by the ability of the surgeon to accurately place the maxilla, the stability of the maxilla in its new location, and the variability of the soft tissue response. The purpose of this retrospective study was to evaluate the first of these factors by comparing the surgical objectives obtained from original presurgical cephalometric prediction tracings with the actual skeletal changes achieved in maxillary orthognathic surgery.

MATERIAL AND METHODS

Forty-six patients were selected from the files of the coauthor (D.M.S). The sample consisted of 31 females and 15 males with a mean age of 27.1 years (range, 13.4-56.0). Case selection was made on the basis of the following criteria: (1) patients received a LeFort I maxillary osteotomy alone or combined with a mandibular osteotomy or genioplasty; (2) preoperative and immediately postoperative lateral cephalometric radiographs were of good quality, and the selected landmarks were not obscured; (3) the original presurgical prescription was present (see the description below); (4) the radiographs showed no apparent misalignment of the subject's head in the cephalostat, with the same machine and technique used for each patient to minimize magnification error; and (5) cleft deformities were excluded because previous surgery and soft tissue anomalies might have influenced the surgical outcome.

The presurgical radiographs were taken at a mean time of 31 days (range, 1-74) before surgery. The postsurgical radiographs were taken at a mean time of 11 days (range, 2-37) after surgery.

The surgical treatment was divided among 8 surgeons. Twenty-one patients were treated by 1 surgeon, 14 by another, and the remaining 11 were treated by 6 surgeons who had 1 to 3 patients each.

For the purposes of the study, it was critical to obtain the original presurgical evaluation used by the orthodontist to determine the surgical objectives. These objectives were established after an extensive esthetic evaluation and an interactive consultation with the patient (to produce graphic images of functional and esthetic goals for patient approval) before the start of treatment and were confirmed before the surgical procedure.[2] The objectives used to evaluate the surgical outcome for the study were generated by the orthodontist and relayed to the surgeon as graphic and written prescriptions with specific instructions for maxillary movement. The following is an example of an actual surgical prescription used in the study: maxilla: impact 4.5 mm and advance 2 mm; mandible: autorotate; chin: advance 7 mm; rhinoplasty: none.

The surgical prescriptions were made on the basis of prediction tracings and model surgery performed by the orthodontist. A surgical splint, if used, was based on the same model surgery as the prediction and was also fabricated by the orthodontist. Ideally, the presurgical tracing plus the surgical prescription would equal the postsurgical tracing. All differences were recorded and labeled as surgical discrepancy.

The presurgical radiograph was traced onto acetate paper, and landmarks were identified. Outlines of the cranial base, the maxilla, and the mandible were drawn to aid in proper superimposition of the presurgical and postsurgical radiographs. Points corresponding to landmarks sella, nasion, porion, orbitale, basion, and pterygomaxillary fissure were marked on the postsurgical radio-

Table I. Summary of discrepancies (mm) between predicted and surgical results

Landmark	*Mean*	*SD*	*Range*	*Significance*	*Absolute mean*	*Absolute SD*	*>2 mm (%)*
A-pt horizontal	−0.5	1.6	−4.4/+2.7	NS	1.3	1.0	26
A-pt vertical	0.5	1.4	−2.4/+3.4	NS	1.1	.9	20
U1 horizontal	−0.4	1.9	−4.0/+4.7	NS	1.4	1.2	28
U1 vertical	0.3	1.6	−4.2/+3.2	NS	1.2	1.0	24
U6 horizontal	−0.4	1.9	−3.9/+4.8	NS	1.5	1.2	30
U6 vertical	1.0	1.3	−1.6/+3.8	*	1.3	0.9	22
SN+7 palatal plane‡	−0.7	2.8	−6.8/+5.5	*	2.2	1.8	39

*$P < .05$.
†Total sample: n = 46.
‡SN+7 palatal plane is measured in degrees.
pt, Point; NS, not significant; U, maxillary.

Table II. Effect of choice of surgeon on prediction accuracy

Landmark	*Mean*	*SD*	*Range*	*Significance*	*Absolute mean*	*Absolute SD*	*>2 mm (%)*
Surgeon 1	n = 21						
A-pt horizontal	0.1	1.4	−2.7/+2.7	NS	1.1	0.9	19
A-pt vertical	0.5	1.4	−2.4/+2.9	NS	1.2	0.9	21
U1 horizontal	0.2	1.9	−3.0/+4.7	NS	1.2	1.3	24
U1 vertical	0.2	1.6	−4.2/+2.5	NS	1.3	1.0	14
U6 horizontal	0.5	1.8	−2.8/+4.8	NS	1.3	1.3	24
U6 vertical	0.8	1.1	−1.2/+2.8	NS	1.1	0.8	19
SN+7-palatal plane†	−0.4	2.7	−6.5/+5.5	NS	2.0	1.8	38
Surgeon 2	n = 14						
A-pt horizontal	−1.4	1.2	−2.8/+1.0	*	1.5	1.0	43
A-pt vertical	0.7	1.5	−1.9/+3.4	NS	1.3	1.0	21
U1 horizontal	−1.5	1.9	−4.0/+1.4	*	2.1	1.1	43
U1 vertical	0.6	1.7	−2.6/+3.2	NS	1.4	1.1	36
U6 horizontal	−1.5	1.7	−3.9/+0.8	*	1.9	1.3	50
U6 vertical	1.2	1.5	−1.6/+3.8	*	1.6	1.0	29
SN+7-palatal plane	−0.4	2.3	−4.5/+2.2	NS	1.9	1.2	29

*$P < .05$.
†SN+7-palatal plane is measured in degrees.
pt, Point; NS, not significant; *U*, maxillary.

graph by superimposing the presurgical and postsurgical radiographs on the cranial base.[22] This was done to assure that these landmarks were coincident on the radiographs. The computer uses sella and nasion to superimpose the cephalometric tracings. In addition, maxillary regional superimposition of the postsurgical radiograph on the presurgical tracing was used as a guide to outline the maxilla on the postsurgical radiograph and to select the maxillary landmarks anterior nasal spine (ANS), A-point, maxillary incisor tip (U1), maxillary molar cusp tip (U6), and posterior nasal spine (PNS). These landmarks were used to assess changes in maxillary position, and, if they were not identified the same on both radiographs, the differences in landmark location would affect the measurement results.

The tracings were digitized with a Neumonics digitizer (Dentofacial Software) and Dentofacial Planner, a computerized diagnostic and treatment planning software system for IBM-compatible computers.

The original surgical prediction tracing was then recreated in Dentofacial Planner Plus so that a direct comparison could be made with the postsurgical tracing. This was accomplished by first loading the presurgical tracing onto the screen. The software has a surgical component that allows the user to move maxilla, mandible, and chin to simulate the desired surgery. The infor-

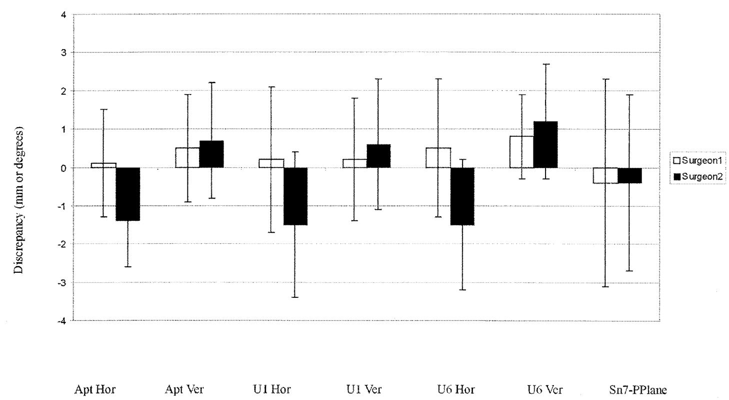

Fig 3. Effect of choice of surgeon on prediction accuracy.

mation from the surgical prescription was used as a guide to create the desired skeletal positions. The software then generated a predicted soft tissue profile on the basis of the movements of the skeletal structures. For this study, the soft tissue profile was not compared with the actual outcome because of facial swelling immediately after surgery. For similar reasons, the predicted mandibular position was not compared with the actual outcome because its position on the immediately postsurgical radiograph might not be accurate because of edema, muscular swelling, and, in some cases, the presence of a surgical splint.

A custom analysis was designed with the Tools (Dentofacial Software) supplemental software. The analysis measured the linear distances (in millimeters) from the maxillary landmarks (A-point, U1, and U6) to horizontal and vertical reference lines (Fig 2). The horizontal reference is a line 7° to sella-nasion (SN) and is referred to as SN^+7. The line was used to orient the tracings on the computer screen and the printouts. The vertical reference was established perpendicular to SN^+7, with its origin at sella (see previous articles). One angular measurement was included to determine discrepancies in the orientation of the palatal plane (PNS-ANS).

The surgical prediction and the actual immediately postsurgical tracing were superimposed on the computer on the basis of SN and oriented on the basis of SN^+7. The analysis was performed for the individual tracings, and differences between the tracings were presented and recorded as the raw data for the study. When evaluating the differences between the 2 tracings, it is helpful to remember that the differences were calculated by subtracting the predicted landmark location from the actual postsurgical landmark location. Therefore, for horizontal measurements, a negative value indicates that the actual result was posterior to predicted position; a positive value indicates that the actual result was anterior to the predicted position. For vertical measurements, a negative value indicates that the actual result was superior to the predicted position, whereas a positive value indicates that the actual result was inferior to the prediction.

An analysis of variance (ANOVA) was used to assess statistical significance. The independent variables of landmark and predicted/actual result were used in all analyses.

The primary interest of the study was how closely the predicted result matched the actual result at each landmark. Two different dependent variables were used to assess this. In the first set of analyses, the dependent variables were the distances from the landmarks to the reference lines (positive and negative), comparing the predicted results with the actual results. To determine the overall accuracy, we used a 2-way repeated-measures ANOVA, with landmark and predicted/actual results being the 2 repeated measures. To test the effect of a variable on the accuracy of the surgery, we divided the sample into a series of subgroups. The first series was 2 subgroups determined on the basis of the surgeon used. The second series involved 2 subgroups determined according to the level of surgical complexity (maxillary alone or bimaxillary). The third series had 2 subgroups based on the primary direction of maxillary movement (impaction vs advancement). The subgroups were compared in each series with 3-way repeated-measures ANOVAs with the same variables as above, but with the added independent variable specific to the series. In this analysis, we were primarily interested in interaction effects. When significance was obtained in an interaction effect, a Newman-Keuls post hoc analysis was used to establish which predicted values were different from the actual values.

The second set of analyses used the absolute value of the discrepancies between predicted and actual measurements as the dependent variable. The independent variables were landmark, surgeon, complexity, and direction. The statistic used for analysis was a 2-way repeated-

Table III. Effect of surgical complexity on prediction accuracy

Landmark	*Mean*	*SD*	*Range*	*Significance*	*Absolute mean*	*Absolute SD*	*>2 mm (%)*
Maxillary	n = 22						
A-pt horizontal	−0.4	1.5	−2.8/+2.5	NS	1.2	0.9	23
A-pt vertical	0.5	1.3	−1.9/+2.8	NS	1.0	0.9	18
U1 horizontal	−0.9	1.0	−4.0/+1.5	NS	1.4	1.2	23
U1 vertical	0.4	1.4	−2.6/+2.7	NS	1.1	0.9	27
U6 horizontal	−0.8	1.6	−3.9/+1.3	*	1.3	1.2	27
U6 vertical	0.5	1.4	−1.6/+3.8	NS	1.1	1.0	14
$SN^{+}7$-palatal plane	0.6	2.4	−3.9/+5.5	NS	1.8	1.5	32
Bimaxillary	n = 24						
A-pt horizontal	−0.6	1.7	−4.4/+2.7	NS	1.3	1.1	29
A-pt vertical	0.4	1.5	−2.4/+3.4	NS	1.2	0.9	21
U1 horizontal	0.1	2.0	−2.9/+4.7	NS	1.5	1.3	33
U1 vertical	0.1	1.7	−4.2/+3.2	NS	1.4	1.0	21
U6 horizontal	0.0	2.0	−2.8/+4.8	NS	1.6	1.2	33
U6 vertical	1.4	1.0	−0.5/+3.2	*	1.5	0.9	29
$SN^{+}7$-palatal plane	−1.9	2.7	−6.8/+2.9	*	2.6	2.0	46

pt, Point; *NS,* not significant; *U,* maxillary.

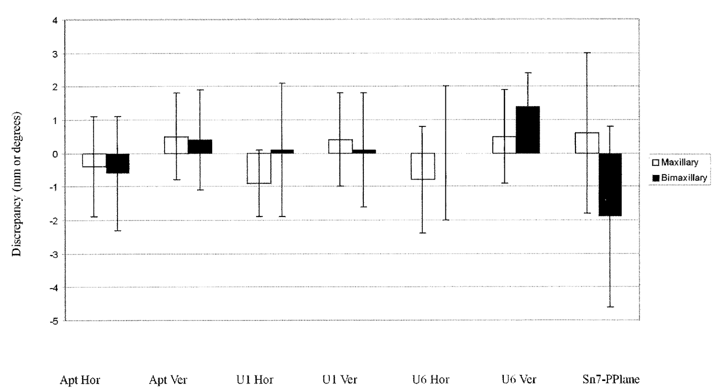

Fig 4. Effect of surgical complexity on prediction accuracy.

measures ANOVA, with landmark as the repeated variable.

In any study based on cephalometrics, the error in the method must be assessed and considered when interpreting the results. To assess the error in this study, we randomly selected 5 subjects and repeated the complete process of presurgical tracing, postsurgical tracing, superimposition, and digitizing 2 to 3 months after the initial evaluation. The tracings were superimposed by the computer and compared with the custom analysis created for the study. The mean differences for the linear landmark measurements were within 0.65 mm. The mean difference for the angular measurement was 0.96.

RESULTS

Surgical prediction and actual outcome measurements

A 2-way repeated-measures ANOVA was performed to look for differences between the prediction and the actual outcome. We examined the effect of 2 independent variables (predicted/actual and landmark) on the distance from the reference lines to the landmarks. Both independent variables were examined in a repeated-measures manner, because all measurements came from the same subject. This analysis measured the significance of 3 effects: predicted versus actual, landmark differences, and interaction between the 2

Table IV. Effect of direction of maxillary movement on prediction accuracy

Landmark	*Mean*	*SD*	*Range*	*Significance*	*Absolute mean*	*Absolute SD*	*>2 mm (%)*
Impact	n = 23						
A-pt horizontal	−0.5	1.6	−4.4/+2.7	NS	1.2	1.1	17
A-pt vertical	0.2	1.6	−2.4/+2.9	NS	1.3	0.9	26
U1 horizontal	0.1	2.0	−3.9/+4.7	NS	1.4	1.4	26
U1 vertical	−0.1	1.8	−4.2/+2.7	NS	1.5	1.0	35
U6 horizontal	0.3	1.9	−3.6/+4.8	NS	1.4	1.4	26
U6 vertical	1.3	1.3	−1.2/+3.8	*	1.5	1.1	30
SN^+7-palatal plane	−2.2	2.5	−6.8/+2.7	*	2.6	2.0	48
Advance	n = 20						
A-pt horizontal	−0.3	1.6	−2.8/+2.5	NS	1.3	1.0	35
A-pt vertical	0.8	1.1	−0.7/+3.4	*	1.0	0.9	15
U1 horizontal	−0.7	1.8	−4.0/+3.7	*	1.5	1.1	30
U1 vertical	0.8	1.1	−0.8/+3.2	NS	1.1	0.9	15
U6 horizontal	−0.9	1.6	−3.9/+1.3	*	1.5	1.0	35
U6 vertical	0.6	1.2	−1.6/+2.8	NS	1.1	0.8	15
SN^+7-palatal plane	1.0	2.3	−3.9/+5.5	*	2.0	1.5	35

*$P > .05$.
pt, Point; *NS*, not significant; *U*, maxillary.

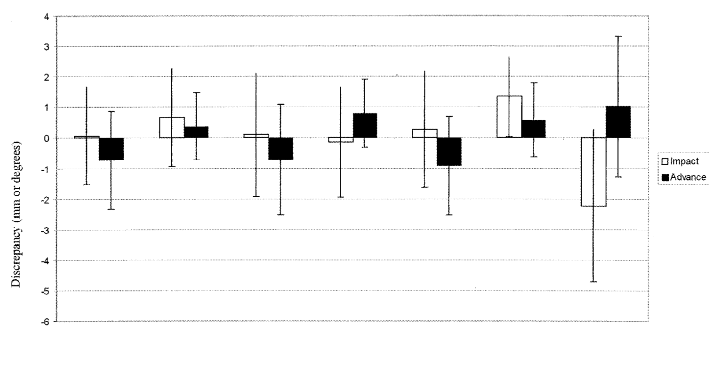

Fig 5. Effect of direction of maxillary movement on prediction accuracy.

independent variables. The mean discrepancies, standard deviations, and ranges for landmark location between the surgical prediction and the actual outcome for the total sample are summarized in Table I. No significant overall differences were found between predicted and actual locations of the landmarks. As expected, the landmark effect was significant, because different landmarks naturally occur at different distances from the reference lines. This is irrelevant to this study. The third effect, the interaction effect, shows whether the delta (discrepancy) between predicted and actual location of each landmark is significantly different from 0 or from other landmarks. The analysis showed that this interaction effect was significant.

Next, a Newman-Keuls post hoc analysis was performed to determine which landmarks were significantly different from 0. For most measurements, the mean differences were very close to 0. The placement of the maxillary molar in a vertical dimension tended to be inferior to the prediction and led to the strongest statistical difference between predicted and actual positions. The angular measurement SN^+7-palatal plane tended toward a negative direction, indicating that the slope of the palatal plane was more upward (or less downward) than predicted.

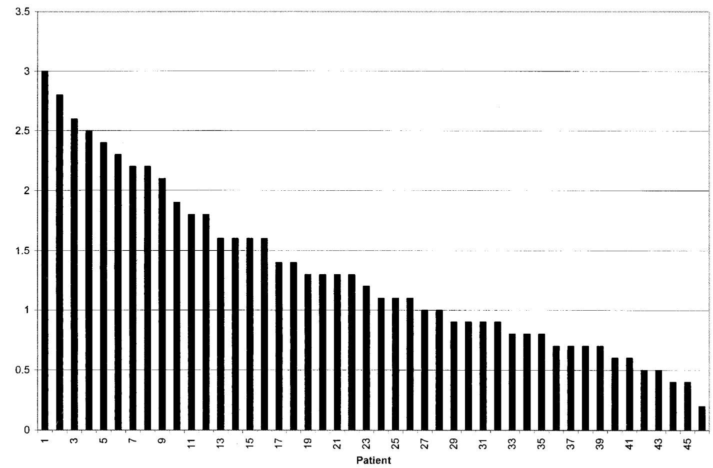

Fig 6. Total sample: average discrepancy (mm) for each patient.

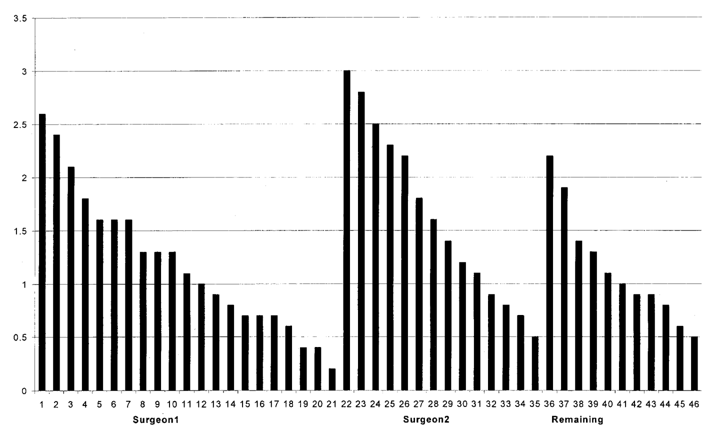

Fig 7. Comparison of average discrepancy (mm) by surgeon.

We used the preceding analysis to determine the significance of the directionality of differences between predicted and actual results. We also wanted to determine if any landmarks were more susceptible to discrepancy, regardless of the direction of the discrepancy. To illustrate the amount of the discrepancy, we removed the directional component by looking at absolute values of the data (Table I). Using these figures, we performed an ANOVA, which showed that only the angle $SN^{+}7$-palatal plane was significantly different from the other measurements ($P = .000014$).

To evaluate clinical relevance, the percentage of patients with a discrepancy of 2 mm or more between predicted and actual landmark locations is also shown in Table I. For the angular measurement, $Sn^{+}7$ perpendicular to the palatal plane, 2° was used as the clinical level of significance. For the total sample, a 2-mm or greater discrepancy was noted for 20% to 30% of the patients, and a 2^{o} or more discrepancy was noted for 39% of the patients.

Surgical accuracy between surgeons

To evaluate the effect of choice of surgeon on prediction accuracy, we used a subset of the data. Surgeon 1 treated 21 patients, and surgeon 2 treated 14 patients. The remaining 6 surgeons treated between 1 and 3 patients each, and their subgroups were too small to evaluate meaningfully. The first 2 surgeons' measurement results are presented in Table II. The directional data showed significance between surgeons 1 and 2, and an interaction effect between choice of surgeon and landmark discrepancy.

Surgeon 1 (21 patients) had no statistically significant discrepancies between predicted and actual locations, and was more consistent in the placement of the maxilla according to the prediction. The percentage of surgeon 1's

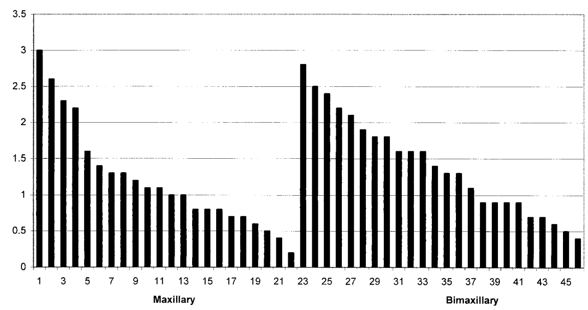

Fig 8. Comparison of average discrepancy (mm) by surgical complexity.

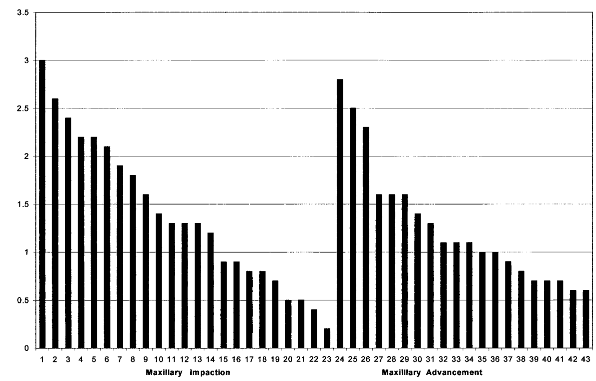

Fig 9. Comparison of average discrepancy by direction of maxillary movement.

patients with discrepancies larger than 2 mm was less than the total sample and less than surgeon 2's patients (range, 19%-24%).

Surgeon 2 had significant differences in predicted versus actual locations for 4 of the landmarks because of a tendency to place the maxilla posterior to its predicted position. The differences in landmark placement between the surgeons are illustrated in Figure 3. As shown in Table II, surgeon 2 had higher percentages of patients, with clinically significant discrepancies (21%-50%); unfortunately, the samples were not of equal size, and the surgical procedures were not exactly the same, so a direct clinical comparison was problematic.

Evaluation of absolute values (amount of discrepancy regardless of direction) showed that surgeon 2 had consistently higher means, but the differences between the surgeons were not significant (repeated-measures ANOVA with the absolute differences between actual and predicted locations ($P = .130$), and the surgeons did not differ from each other by landmark (interaction effect $P = .464$).

Complexity of surgical procedure

To evaluate the effect of surgical complexity as a variable in the reliability of the prediction, we divided the total sample into 2 complexity subgroups. The first group consisted of 22 LeFort I procedures with or without a genioplasty involving some degree of mandibular rotation but no concurrent mandibular osteotomy. The second group consisted of 24 LeFort I procedures combined with bilateral sagittal split osteotomy mandibular advancement or setback and occasional adjunctive genioplasty. The results of the landmark measurements are shown in Table III and Figure 4. A 3-way ANOVA showed no overall differences in landmark discrepancies because of complexity. However, the deltas (discrepancies between predicted and actual placement) were significant for certain surgeries at certain landmarks. Specifically, the bimaxillary procedures tended to have negative discrepancies, indicating that the slope of the palatal plane was more upward (or less downward) than predicted. The bimaxillary procedures also resulted in more positive maxillary first molar vertical change than predicted. This might have resulted from less posterior impaction than predicted. The maxillary procedures were more negative on maxillary first molar horizontal than predicted.

There was no significant effect of complexity on the absolute amount of discrepancy ($P = .16$), nor was there an interaction effect between complexity and landmark ($P = .81$) as measured by a repeated-measures 2-way ANOVA.

Effect of primary direction of maxillary movement on reliability

To evaluate this variable, we established 2 subgroups. Those involving primarily maxillary impaction were placed in group 1 (n = 23), and those involving primarily advancement were placed in group 2 (n = 20). Some patients received a combination of movements. For example, the maxilla might have been moved upward 4 mm and forward 2 mm. This patient would have been placed in the impaction group because that was the greater movement. Three patients had primarily downward movement and were excluded from the analysis. The comparison of the subgroups is summarized in Table IV and Figure 5.

An ANOVA showed a difference between the subgroups in the amount of surgical change: greater overall changes were predicted and realized for the advancement surgeries. More to the interest of this study was an interaction effect between the direction of movement and the landmark measurements. For the advancement patients, there was a significant tendency for the surgeon to fall short of the desired advancement. In 25% of those subjects, the maxilla was underadvanced by 2 mm or more. In addition, there was a significant tendency for the

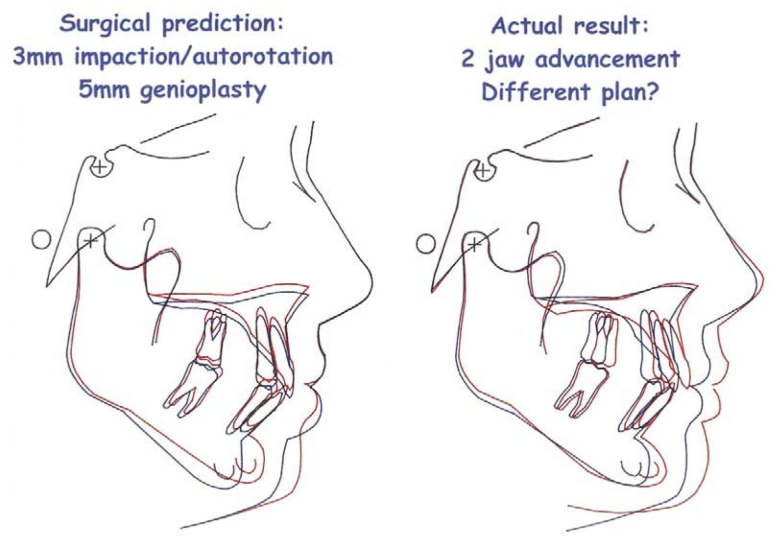

Fig 10. Example of 5-mm horizontal discrepancy (average discrepancy, 2.5 mm). Mandible is not in occlusion in postsurgical radiograph.

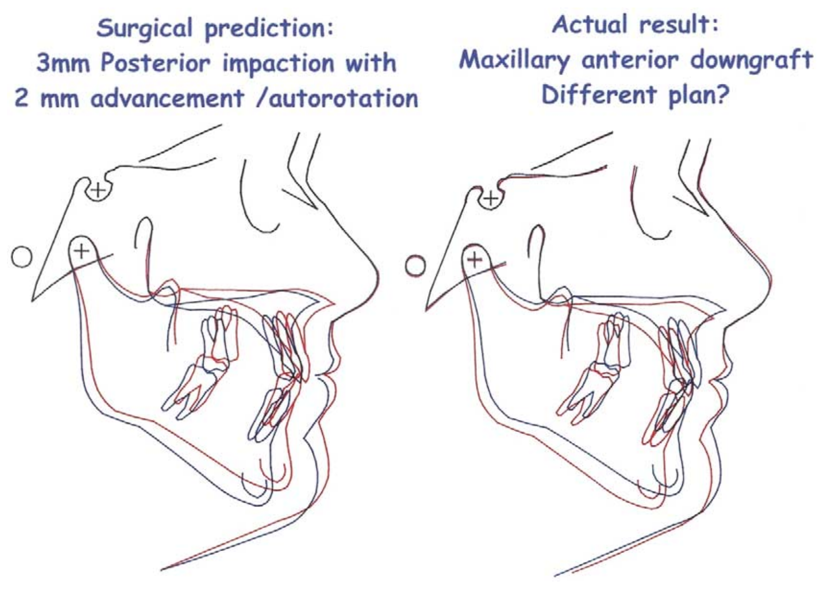

Fig 11. Vertical and horizontal discrepancy (average discrepancy, 3 mm).

maxilla to have downward movement that was not anticipated by the prediction. For the impaction subgroup, the maxillary first molar vertical measurement and the $SN^{+}7$-palatal plane were significant. Clinically, in 30% of the cases, the surgeon underimpacted the maxilla by 2 mm or more.

Direction of maxillary movement had no significant effect on absolute amounts of discrepancy as measured by a 2-way repeated measures ANOVA ($P = .33$), nor was there an interaction effect between landmark and direction ($P = .31$).

Evaluation of individual cases

Evaluating the means and standard deviations of the measurements can yield information about tendencies in direction and amount of discrepancy for a specific landmark, but in evaluating the accuracy of a prediction for a patient, it is helpful to look at the maxilla as a whole. An overall discrepancy in millimeters for each patient was calculated as the total of the absolute values of the linear measurements. The absolute values were used so that discrepancies in different directions would

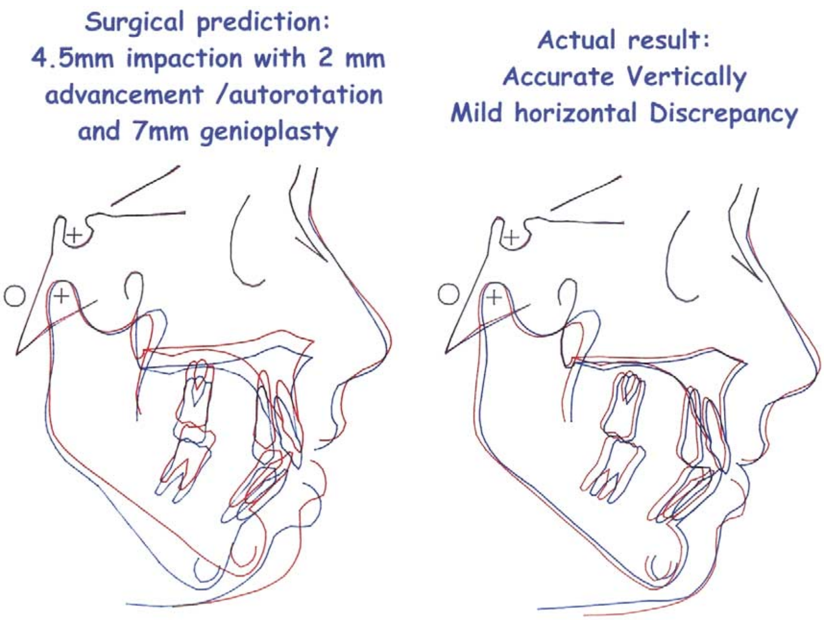

Fig 12. Mild average discrepancy (1.5 mm).

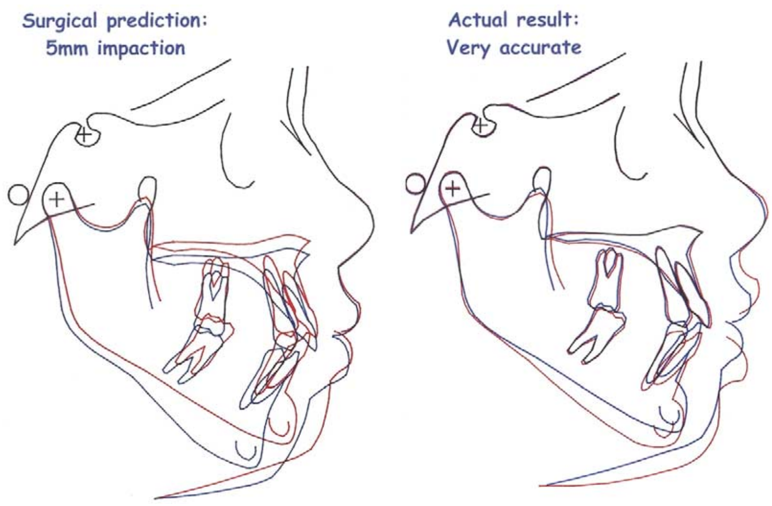

Fig 13. Minimal average discrepancy (0.6 mm).

not cancel each other. The average discrepancy in millimeters for the patient is the overall discrepancy divided by 6 (number of millimeter measurements). By assigning a value to the overall fit of the actual result to the prediction, it was possible to organize the sample from least accurate to most accurate (Fig 6). The average discrepancy in millimeters does not take into account the angular data, so similar calculations combining the measurements in millimeters and angular data were also performed. The results were very close (with a slight difference in the sorting of the patients), so the 2 types of measurements were not combined.

As the graph illustrates, many predictions had very little discrepancy from the actual result. In fact, 43% of the predictions had less than 1 mm average discrepancy, and 80% had less than 2 mm average discrepancy.

Organizing the data by average discrepancy also allowed for a visual comparison between the series of subgroups that were compared in the study. Figure 7 shows the average discrepancy for the predictions in the

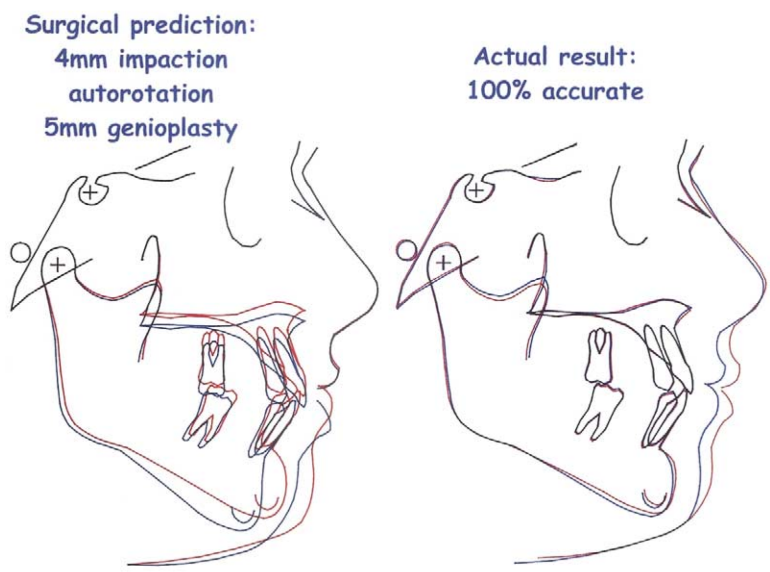

Fig 14. Least average discrepancy (0.2 mm).

subgroups determined by the surgeon involved in the treatment. This confirms that there was little difference between the surgeons. Figure 8 shows the average discrepancy in the predictions for the subgroups determined by surgical complexity. There appears to be less average discrepancy with single-jaw procedures than with double-jaw procedures, but the difference is not significant. Figure 9 shows the average discrepancy for the predictions in the subgroups determined by the direction of maxillary movement. Again, the graphs are similar.

DISCUSSION

The results of this study suggest that LeFort I maxillary orthognathic surgery can be very accurate. However, the results showed a fairly wide range. Most patients (80%) had good approximation between the prediction and the actual result, and several were extremely good. What happened to the patients for whom predictions did not match surgical outcome? In some, for valid reasons, the surgeon could not, or chose not to, follow the final presurgical plan.

Figure 10 shows the surgical plan for a patient and the actual result superimposed on the prediction. This patient demonstrates 1 of the highest average millimeter discrepancies. In this plan, a 3-mm maxillary impaction was prescribed with subsequent mandibular autorotation and a 5-mm advancement genioplasty. The surgeon impacted the maxilla 3 mm but also advanced it 5 mm. Then, because the mandibular autorotation was not sufficient, a mandibular advancement was performed with a concurrent advancement genioplasty. This does not necessarily indicate that the result of the surgical procedure was poor, but that it was different from the original plan. In fact, the surgeon might have viewed the patient's esthetic needs differently. This demonstrates the need for surgical prediction and consultation between the involved persons but also illustrates a shortcoming of the current method of investigation. To label the measured discrepancies between prediction and actual result a "surgical error," we would need to establish with confidence that the surgeon attempted to follow the agreed-on prediction.

Another example of a patient with a high average discrepancy is shown in Figure 11. In this patient, the surgical prescription called for a 3-mm posterior impaction and a 2-mm advancement to accommodate the mandibular autorotation. Instead, the surgeon downgrafted the anterior portion of the maxilla and left the posterior maxilla alone. This meant that no mandibular autorotation occurred, so there was no need for the maxillary advancement. Was this surgical error or disagreement with the plan? The result was the highest average millimeter discrepancy (3 mm).

Fortunately, most results were considerably closer to the prediction than the last 2 examples. Figure 12 illustrates a typical case that falls in the middle of the sample when sorted by the average discrepancy. The surgical prescription called for a 4.5-mm maxillary impaction with a 2-mm advancement to allow for mandibular autorotation and a 7-mm advancement genioplasty. Figures 13 and 14 show examples of extremely accurate results.

This study suggests that, to evaluate the error in the surgical technique, the method should include the surgeon's own predictions. This would eliminate the variable of communication and agreement between the clinicians.

Why would the clinicians not agree? Because this study is retrospective, that is a difficult question to answer. Most patients in the sample originally sought orthodontic treatment and were referred to the surgeons. However, some were referred for orthodontic treatment by the surgeons involved. Did that make a difference in the perception of who was responsible for deciding the final esthetic goals? A prospective study involving consecutive cases and confirmed agreement would be ideal. The next step in evaluating the existing sample would be to look at the posttreatment radiograph at least 6 months after the surgery to evaluate the effect of the discrepancies in maxillary position on the soft tissue outline, ie, to determine if the surgical goals were met.

CONCLUSIONS

1. LeFort I maxillary and maxillomandibular surgery can be extremely accurate.
2. In 80% of the sample, the results of the surgery were, on average, within 2 mm of the prediction.
3. There were statistical differences between 2 surgeons in the direction of specific landmark discrepancies but not in the amount.
4. There were no statistical differences between the surgical groups when divided by surgical complexity (single-jaw vs bimaxillary procedures) with the exception of the direction of the palatal plane discrepancy.
5. There were statistical differences between the surgical groups when divided by direction of movement (impaction vs advancement) in the direction of landmark discrepancies. For advancements, the surgeons tended to place the maxilla posterior and inferior to its predicted position. For impactions, the surgeons tended to underimpact the posterior maxilla. However, the amounts were not significant when absolute values were considered.
6. The study supports the use of surgical prediction tracing and demonstrates the need for communication between the involved clinicians to ensure that they agree on a plan.

REFERENCES

1. Pospisil OA. Reliability and feasibility of prediction tracing in orthognathic surgery. J Craniomaxillofac Surg 1987;154:79-83.
2. Sarver DM. Esthetic orthodontics and orthognathic surgery. St Louis: Mosby-Year Book; 1998.
3. Hing NR. The accuracy of computer generated prediction tracings. Int J Oral Maxillofac Surg 1989;18:148-51.
4. Jacobson RS. The reliability of soft tissue profile prediction in orthognathic surgery [thesis]. Chicago: Northwestern University; 1989.
5. Konstiantos KA, O'Reilly MT, Close J. The validity of the prediction of soft tissue profile changes after LeFort I osteotomy using the dentofacial planner (computer software). Am J Orthod Dentofacial Orthop 1994;105:241-9.
6. Sinclair PM, Kipelainen P, Phillips C, White RP, Rogers L, Sarver DM. The accuracy of video imaging in orthognathic surgery. Am J Orthod Dentofacial Orthop 1995;107:177-85.
7. Aharon PA, Eisig S, Cisneros GJ. Surgical prediction reliability: a comparison of two computer software systems. Int J Adult Orthod Orthognath Surg 1997;12:65-78.
8. Gerbo LR, Poulton DR, Covell DA. A comparison of a computer-based orthognathic surgery prediction system to postsurgical results. Int J Adult Orthod Orthognath Surg 1997;12:55-62.
9. Powers B. The accuracy of computer generated profile predictions associated with mandibular advancement surgery [thesis]. Chicago: Northwestern University; 1998.
10. Turvey T, Hall DJ, Fish LC, Epker BN. Surgical-orthodontic treatment planning for simultaneous mobilization of the maxilla and mandible in the correction of dentofacial deformities. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1982;54:491-8.
11. Harradine NWT, Birnie DJ. Computerized prediction of the results of orthognathic surgery. J Maxillofac Surg 1985;13:245-9.
12. Friede H, Kahnberg K, Ragnar A, Ridell A. Accuracy of cephalometric prediction in orthognathic surgery. J Oral Maxillofac Surg 1987;45:754-60.
13. Bryan DC, Hunt NP. Surgical accuracy in orthognathic surgery. Br J Oral Maxillofac Surg 1993;31:343-50.
14. Bryan DC. An investigation into the accuracy and validity of three points used in the assessment of autorotation in orthognathic surgery. Br J Oral Maxillofac Surg 1994;32:363-72.
15. Johnson DG. Intraoperative measurement of maxillary repositioning: an ancillary technique. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1985;60:266-8.
16. Van Sickels JE, Larsen AJ, Triplett R. Predictability of maxillary surgery: a comparison of internal and external reference marks. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1986;61:542-5.
17. Stanchina R, Ellis E III, Gallo WJ, Fonseca RJ. A comparison of two measures for repositioning the maxilla during orthognathic surgery. Int J Adult Orthod Orthognath Surg 1988;3:149-54.
18. Wylie GA, Epker BN, Mossop JS. A technique to improve the accuracy of total maxillary surgery. Int J Adult Orthod Orthognath Surg 1988;3:143-7.
19. Polido WD, Ellis E III, Sinn DP. An assessment of the predictability of maxillary surgery. J Oral Maxillofac Surg 1990;48:697-701.
20. Ellis E III, Tharanon W, Gambrell K. Accuracy of face-bow transfer: effect on surgical prediction and postsurgical result. J Oral Maxillofac Surg 1992;50:562-7.
21. Wiesenfeld D. Improving the accuracy of orthognathic surgery. Ann R Aust Coll Dent Surg 1994;12:225-6.
22. Upton PM, Sadowsky PL, Sarver DM, Heaven TJ. Evaluation of video imaging prediction in combined maxillary and mandibular orthognathic surgery. Am J Orthod Dentofacial Orthop 1997;112:656-65.
23. Kiyak HA, Bell R. Psychosocial considerations in surgery and orthodontics. In: Profitt WR, White RP, editors. Surgical-orthodontic treatment. St Louis: Mosby; 1991.
24. Kiyak HA, Vitaliano PP, Crinean J. Patients' expectations as predictors of orthognathic surgery outcomes. Health Psychol 1988;7:251-68.