Accuracy of a LeFort I Maxillary Osteotomy

Steven Semaan[a]; Mithran S. Goonewardene[b]

Abstract: An optimal outcome of combined surgery and orthodontics involving the maxilla is dependent on many factors. Accurate placement of the maxilla by the surgical team is ultimately of paramount importance. The aim of this retrospective study was to evaluate the accuracy of LeFort I maxillary osteotomy with respect to the presurgical prediction. The sample comprised 42 patients (33 females, nine males) who had undergone LeFort I osteotomy procedure alone or in combination with a mandibular osteotomy with or without genioplasty. Tracings of presurgical and immediate postsurgical lateral cephalograms and surgical predictions were digitized and compared using Quick Ceph software analysis. Vertical and horizontal measurements to various skeletal landmarks were used to assess the discrepancy between the predicted maxillary position and the actual postsurgical result. Statistically significant differences were found between the predicted and actual postsurgical maxillary molar vertical position, and significant differences were also found for the palatal plane angular measurements. Two surgical teams were compared, and surgical team 1 had significantly less variation in the surgical outcomes than did surgical team 2. When single-jaw and bimaxillary surgery were compared, no significant differences were found. Similarly, there were no statistically significant differences found when assessing the primary direction of movement (impaction vs downgraft vs advancement). Overall, 66% of the results were within two mm of prediction and 26% of the results were within one mm of prediction. A LeFort I maxillary osteotomy can be an accurate procedure with a wide range of discrepancy. (*Angle Orthod* 2005;75:964–973.)

Key Words: Maxillary osteotomy; Surgical accuracy; LeFort I

INTRODUCTION

Orthognathic surgery is an effective method for correcting significant skeletal and dentofacial discrepancies. There are several articles in the orthodontic and surgical literature that discuss the accuracy of predicting surgical outcomes.[1–5] Most of these studies deal with assessing the algorithms used to predict the soft and hard tissue profiles produced by imaging programs. What is worth noting is that all these studies make an assumption that the surgeon always places the jaws in the planned position. Because things do not always go according to plan, the question arises of just how accurate is the repositioning of the jaws? That is, can the surgeon actually produce what the orthodontist has prescribed and to what degree of accuracy?

A recent study conducted at the University of Alabama by Jacobson and Sarver[6] addressed this question. It is one of only a few studies designed to focus exclusively on evaluating the accuracy of the surgical team. They found that in 43% of the subjects the overall results were within one mm of prediction and in 80% of the subjects the actual results were within two mm of the prediction. These results were based on evaluation of patients in one practice.

The opportunity also arose to compare the surgical outcomes of two surgical teams performing LeFort I repositioning. One team was involved exclusively in private practice, and the second team treated patients in a teaching hospital as part of a surgical training program.

[a] Orthodontic Resident, Dental School, Oral Health Centre of Western Australia, The University of Western Australia, Nedlands, Western Australia, Australia.

[b] Program Director, Orthodontics, Oral Health Centre of Western Australia, The University of Western Australia, Nedlands, Western Australia, Australia.

Corresponding author: Mithran S. Goonewardene BDSc, MMedSc, CertOrth, Program Director, Orthodontics, Oral Health Centre of Western Australia, The University of Western Australia, 17 Monash Avenue, Nedlands 6907, Western Australia (e-mail: mithran@ohcwa.uwa.edu.au).

Accepted: January 2005. Submitted: October 2004.

MATERIALS AND METHODS

The presurgical and postsurgical cephalometric radiographs and the hand-generated surgical prediction tracings of a total of 42 consecutively treated subjects were included in this retrospective study. The records of patients treated by team 1 were selected from the files of a private orthodontic practice (22 subjects; 16 females, aged 15 to 45 years and six males, aged 18 to 30 years). The records of patients treated by team 2 were selected from the files of the Orthodontic Department of the School of Dentistry, University of Western Australia (20 subjects; 17 females, aged 15 to 43 years and three males, aged 16 to 30 years). The LeFort I procedures took place between January 1996 and June 2004.

The inclusion criteria were that each subject had received a LeFort I maxillary osteotomy procedure with or without a genioplasty (13 subjects: eight private practice, five university) or a LeFort I osteotomy combined with a mandibular osteotomy (either a setback or advancement) with or without genioplasty (29 subjects in total: 15 private practice, 14 university) by either of the two surgeons and treated under the supervision of the same orthodontist. Presurgical radiographs were taken at a mean of 51 days before the date of surgery, and postsurgical radiographs were taken at a mean of 11 days after surgery.

The reasons for exclusion included:

- Incomplete records;
- Presence of any congenital craniofacial anomaly or syndrome;
- Presence of any significant skeletal asymmetries; or
- Radiographs taken at different radiology centers or not of an acceptable quality.

The orthognathic surgery was performed by two oral maxillofacial surgical teams. A total of 23 patients were treated by surgical team 1 and 18 by surgical team 2.

Surgical splints were used for each patient and were based on hand-generated surgical prediction tracings and subsequent model surgery performed by both the surgeon and orthodontic clinician.

The pre- and postsurgical radiographs and surgical predictions for each subject were traced onto acetate paper with 0.5 mm pencil. A total of 12 landmarks were then identified on the presurgical radiograph and transferred to the postsurgical radiographs tracing and surgical prediction tracing.

Seven landmarks (Nasion, Sella, Porion, Basion, Pterygoid, Orbitale, and PM) were transferred by cranial base best-fit superimposition. Five other landmarks (ANS, PNS, A-point, upper incisal edge, and distal upper molar cusp tip) were transferred by maxillary regional superimposition. This was done to improve overall method accuracy and also to ensure that these landmarks were coincident. A similar method has been used in previous reports.[6]

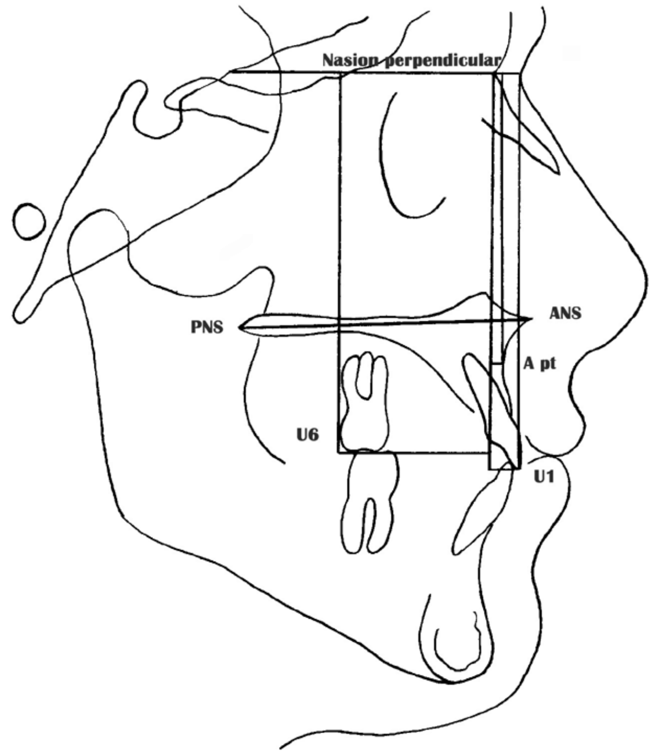

FIGURE 1. Custom Analysis.

The three tracings for each subject were then scanned to create standard jpeg files. These scanned tracings were imported into Quick Ceph software (San Diego, Calif) on an Apple Macintosh personal computer and analyzed using a customized cephalometric analysis. The analysis created an xy coordinate system that enabled measurement of linear distances from horizontal and vertical reference lines.

The horizontal reference was called Nasion Perpendicular. It was parallel to the Frankfort-Horizontal line and passed through Nasion. The vertical reference line was drawn at right angles to the Nasion Perpendicular starting at Nasion and extended inferiorly (Figure 1). The images were enlarged during digitization to aid the landmark identification. All linear and angular measurements were made from five maxillary landmarks with respect to the vertical and horizontal reference lines and calculated by the software to the nearest 0.1 mm and 0.1°, respectively. Adjustments were made to account for the varying magnifications inherent in the lateral cephalogram of each subject.

To assess surgical accuracy, differences between actual and predicted landmarks were calculated by subtracting the predicted landmark location from the actual postsurgical landmark location for the five maxillary landmarks. For vertical measurements, a nega-

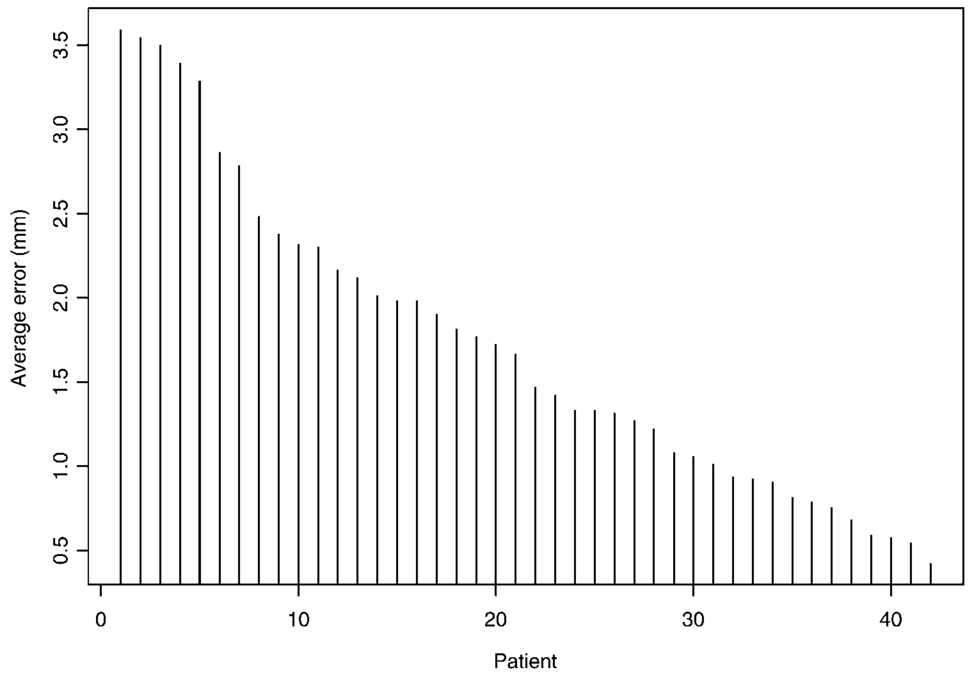

FIGURE 2. Total sample: average discrepancy (mm) for each patient.

tive value indicated that the actual resulting landmark position was superior to the predicted position whereas a positive value indicated the actual result was inferior to the predicted position. For horizontal measurements, a negative value indicated that the actual result was posterior to the predicted position and a negative value indicated the actual result was anterior to the predicted position.

Analysis of variance tests were then used to assess the statistical significance of the overall difference between the discrepancies between predicted and actual surgical results. The effects that a particular variable had on the accuracy of surgery were examined by dividing the sample into three subgroups:

- Effect of choice of surgical team;
- Effect of choice of surgical complexity; and
- Effect of primary direction of maxillary movement (impaction, advancement, or downgraft).

The overall discrepancy between predicted and actual results for each landmark was calculated. *P* values were obtained using paired *t*-tests to determine whether the sample mean was significantly different from zero. To determine the accuracy of the scientific method applied, cephalograms of five randomly selected subjects were retraced and redigitized on two separate occasions at a minimal interval of 10 days. The mean difference for the linear measurements was 0.6 mm (SD $\pm$ 0.3) and for the angular measurements was 1.5 degrees (SD $\pm$ 1.6).

RESULTS

Overall average discrepancy between predicted and actual results

The overall average discrepancy between predicted and actual results was measured. The absolute values of the linear measurements were used to avoid misinterpreting the results because discrepancies in opposite directions would cancel each other out, thus giving the impression that the results were more accurate than they actually were. The values were organized from least accurate to most accurate. Approximately two-thirds (66%) of the subjects had their maxilla placed within two mm of prediction and about one-quarter (26%) of the maxillae were placed within one mm of the predicted outcome (Figure 2). The mean discrepancies, standard deviations, ranges, and *P* values for landmark locations between the surgical prediction and actual outcome for the total sample are summarized in Table 1.

There was moderate evidence that the mean difference in actual measurements for two of the landmarks was nonzero. These were the maxillary molar (U6) vertical measurements (two-sided *P* value, .036; 95% confidence interval, −1.05 to −0.04) and palatal plane measurements (two-sided *P* value, .027; 95% confidence interval, −1.91 to −0.12).

Effect of choice of surgical team

To determine the effect of the choice of surgical team on surgical accuracy, the sample was divided

TABLE 1. Summary of Discrepancies Between Predicted and Surgical Results[a]

Landmark	Mean	SD	Range	Significance (*P* value)	Absolute Mean	Absolute SD	>2 units (%)
A-pt horizontal	−0.27	2.61	−6.09 to 6.55	NS (.505)	2.06	1.60	42.9
A-pt vertical	0.30	2.01	−3.82 to 4.74	NS (.333)	1.61	1.22	26.2
U1 horizontal	−0.07	2.58	−6.82 to 5.45	NS (.870)	1.91	1.71	33.3
U1 vertical	0.37	1.90	−3.82 to 4.46	NS (.218)	1.50	1.20	33.3
U6 horizontal	0.03	2.46	−6.64 to 5.73	NS (.946)	1.82	1.63	33.3
U6 vertical	−0.54	1.62	−3.18 to 4.09	S (.036)	1.42	0.94	23.8
Palatal plane	−1.01	2.87	−6.55 to 7.36	S (.027)	2.43	1.80	47.6

[a] Please note Palatal plane is measured in degrees. Other landmarks are measured in millimetres. >2 units means the absolute value of the discrepancy is greater than 2 units. NS indicates not significant; S, significant at the 5% level. Total sample: n = 42.

TABLE 2. Effect of Choice of Surgical Team

Landmark	Mean	SD	Range	Significance (*P* value)[a]	Absolute Mean	Absolute SD	>2 units (%)
Surgical team 1	n = 23						
A-pt horizontal	−0.72	2.44	−6.09 to 3.09	NS (.173)	1.96	1.57	43.5
A-pt vertical	0.14	2.04	−3.82 to 4.74	NS (.749)	1.61	1.21	21.7
U1 horizontal	−0.18	1.75	−3.90 to 2.73	NS (.623)	1.34	1.11	21.7
U1 vertical	0.14	1.82	−3.82 to 4.46	NS (.714)	1.36	1.19	21.7
U6 horizontal	−0.10	1.71	−4.18 to 3.18	NS (.777)	1.31	1.07	21.7
U6 vertical	−0.47	1.76	−3.18 to 4.09	NS (.215)	1.47	1.04	21.7
Palatal plane	−1.15	2.43	−4.73 to 4.00	S (.034)	2.21	1.48	52.2
Surgical team 2	n = 18						
A-pt horizontal	0.33	2.84	−4.36 to 6.55	NS (.630)	2.25	1.68	44.4
A-pt vertical	0.40	2.01	−2.55 to 4.09	NS (.412)	1.55	1.29	27.8
U1 horizontal	0.10	3.46	−6.82 to 5.45	NS (.903)	2.72	2.04	50.0
U1 vertical	0.56	2.02	−2.82 to 4.18	NS (.256)	1.65	1.24	44.4
U6 horizontal	0.22	3.27	−6.64 to 5.73	NS (.782)	2.54	1.98	50.0
U6 vertical	−0.77	1.40	−2.73 to 2.00	S (.033)	1.33	0.84	27.8
Palatal plane	−0.94	3.47	−6.55 to 7.36	NS (.266)	2.80	2.16	44.4

[a] NS indicates not significant; S, significant at the 5% level.

into two subgroups: surgical team 1 (23 patients, exclusive private practices) and surgical team 2 (18 patients, teaching hospitals). The measurement results are displayed in Table 2.

For surgical team 1, the only landmark that showed a significant discrepancy between predicted and actual outcome was the palatal plane measurement (two-sided *P* value, .034; 95% confidence interval, −2.20 to −0.10). This meant that surgical team 1 tended to either over-rotate the maxilla if a clockwise rotation was predicted or under-rotate the maxilla if anticlockwise rotation was predicted. For surgical team 2, the only landmark that showed a significant discrepancy between predicted and actual outcome was the maxillary molar vertical measurement (U6 vertical) (two-sided *P* value, .033; 95% confidence interval, −1.46 to −0.07). This meant that surgical team 2 tended to place the maxillary molar more superiorly than predicted.

TABLE 3. Comparison Between Teams

Landmark	*t*-test *P* value	Wilcoxon Test *P* value	Absolute Values Wilcoxon Test *P* value[a]
A-pt horizontal	NS .213	.344	NS .528
A-pt vertical	NS .684	.813	NS .803
U1 horizontal	NS .754[b]	.773	S .010
U1 vertical	NS .490	.431	NS .494
U6 horizontal	NS .710[b]	.599	S .494
U6 vertical	NS .562	.753	NS .693
Palatal plane	NS .820	.885	S .655

[a] NS indicates not significant; S, significant at the 5% level.
[b] Welch modified *t*-test used.

Comparison between teams

The data in Table 3 shows that the variance of the differences in upper incisor horizontal (U1H) measurements and upper molar horizontal (U6H) measure-

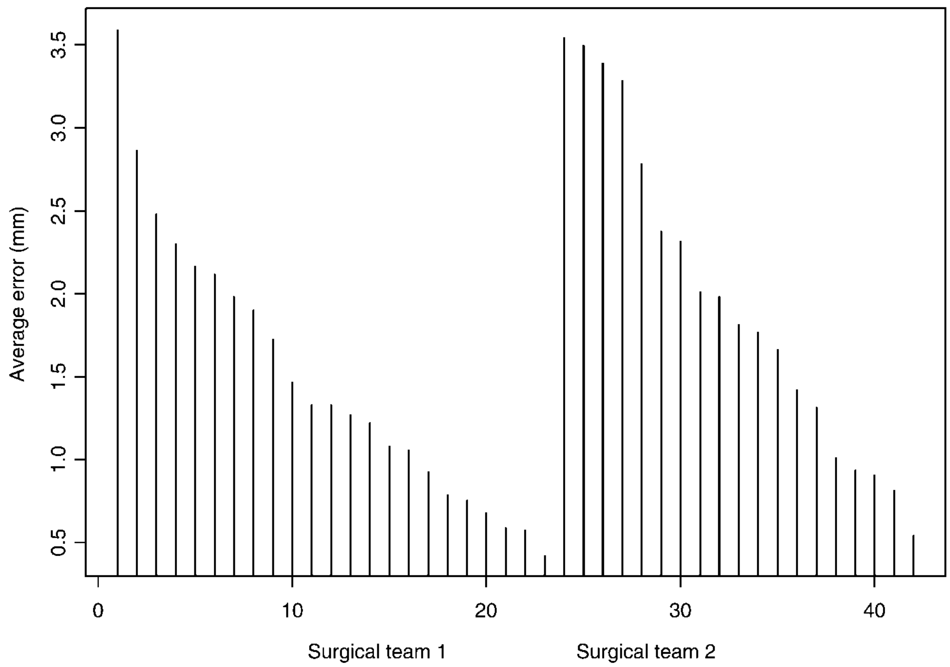

FIGURE 3. Comparison of average discrepancy (mm) by surgical team.

ments are significantly different for the two teams (two-sided *P* values are .003 and .005, respectively, from an *F*-test for equality of variance). Overall, *P* values for surgical team 1 showed significantly less variation in discrepancy than surgical team 2.

Organizing the data by average discrepancy, the difference between the two teams can be observed (Figure 3). Interestingly, there was no significant difference detected by *t*-tests in the mean differences between the two surgical teams. Wilcoxon tests were used for equality of medians because absolute discrepancies were not normally distributed. We found significant differences between the teams for U1H and U6H measurements. The differences in individual landmark placement between the surgical teams are shown in Figure 4. As shown in Table 2, surgical team 2 had a higher percentage of patients with clinically significant discrepancies (27.8–50%). Unfortunately, because the surgical procedures were not exactly the same, direct clinical comparisons are problematic.[6]

For surgery performed in a private practice (surgical team 1), approximately three-quarters (74%) of the subjects had their maxilla placed within two mm of prediction and about one-third (30%) of the maxillae were placed within one mm of the predicted outcome. For surgery performed in the teaching hospital (surgical team 2), approximately half (50%) of the subjects had their maxilla placed within two mm of prediction and about one-fifth (22%) of the maxillae were placed within one mm of the predicted outcome (Figure 3).

Effect of choice of surgical complexity

To analyze the effect that surgical complexity may have on the prediction, the sample was divided into two subgroups: maxillary surgery (ie, those who underwent LeFort I procedures with or without genioplasty) and bimaxillary surgery (ie, those who underwent LeFort I procedures combined with mandibular osteotomy with or without genioplasty).

The results of the landmark measurements are shown in Table 4. The differences in landmark placement between the maxillary and bimaxillary groups are shown in Figure 5. There was insufficient evidence that the mean differences were nonzero for any of the landmarks for maxillary surgery. However, there was strong evidence that the mean difference for bimaxillary surgery was nonzero for two landmarks: U6 vertical measurements (two-sided *P* value, .005; 95% confidence interval, -1.29 to -0.25) and palatal plane measurements (two-sided *P* value, .007; 95% confidence interval, -2.59 to -0.46). This meant that 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 negative vertical change, indicating that the actual result tended to be superior to the predicted position.

When comparing maxillary and bimaxillary surgery, there were no significant differences detected by *t*-tests (or Wilcoxon rank sum tests) for any landmarks

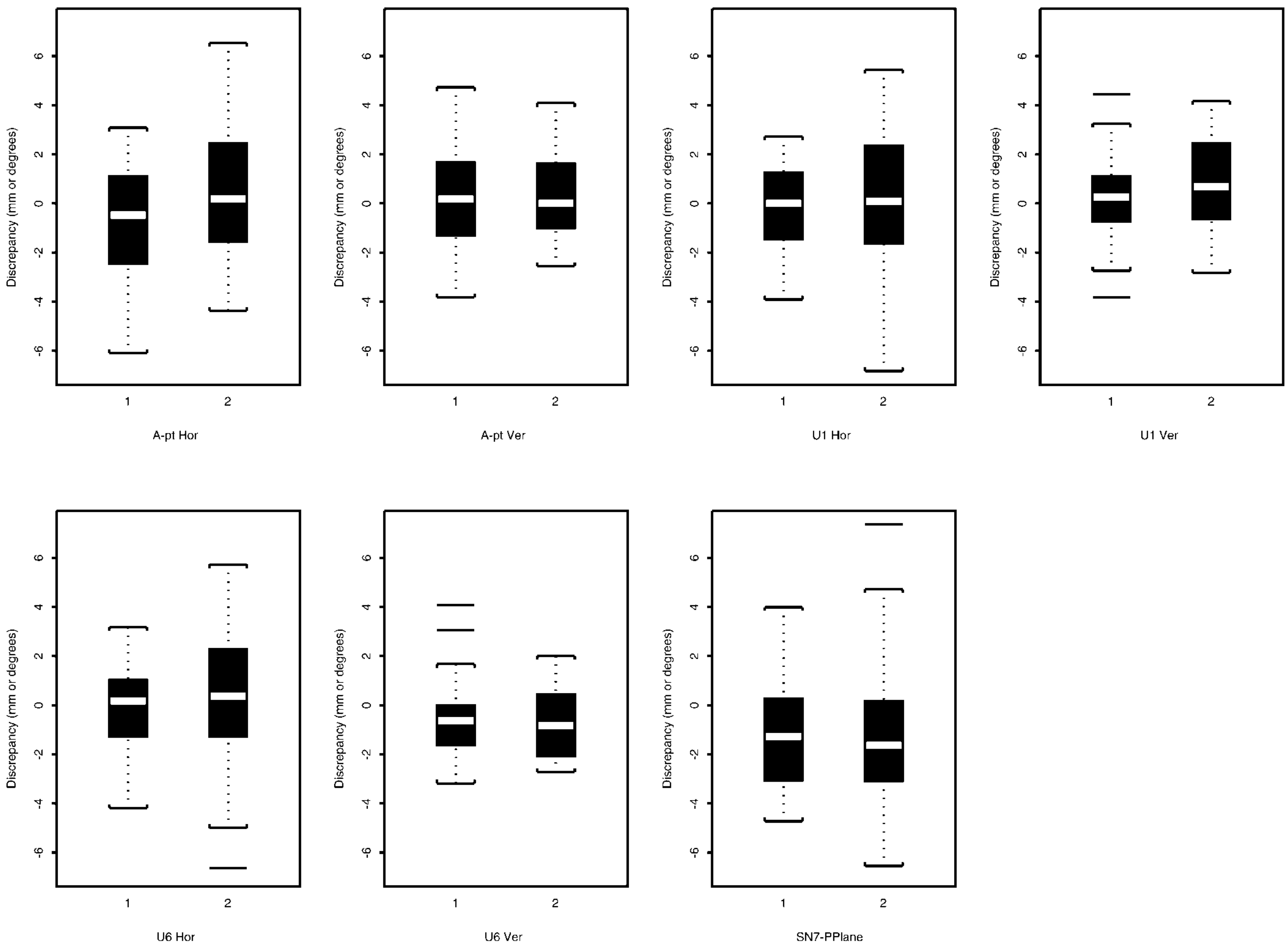

FIGURE 4. Effect of choice of surgical team.

TABLE 4. Effect of Choice of Surgical Complexity

Landmark	Mean	SD	Range	Significance (P value)[a]	Absolute Mean	Absolute SD	>2 units (%)
Maxillary	n = 13						
A-pt horizontal	−1.36	2.79	−6.09 to 4	NS (.103)	2.47	1.78	38.5
A-pt vertical	0.20	2.51	−3.82 to 4.74	NS (.776)	2.00	1.42	30.8
U1 horizontal	0.62	2.19	−3.90 to 5.45	NS (.324)	1.58	1.58	23.1
U1 vertical	0.33	2.21	−3.82 to 4.46	NS (.605)	1.63	1.46	38.5
U6 horizontal	0.42	2.09	−4.18 to 4.36	NS (.482)	1.50	1.45	23.1
U6 vertical	−0.04	2.07	−2.73 to 4.09	NS (.950)	1.72	1.04	23.1
Palatal plane	0.13	2.80	−4.55 to 4.73	NS (.874)	2.32	1.42	46.2
Bimaxillary	n = 29						
A-pt horizontal	0.22	2.41	−4.36 to 6.55	NS (.628)	1.87	1.50	44.8
A-pt vertical	0.35	1.79	−2.55 to 4.09	NS (.302)	1.43	1.10	24.1
U1 horizontal	−0.37	2.71	−6.82 to 5.45	NS (.463)	2.05	1.77	37.9
U1 vertical	0.38	1.78	−2.82 to 4.18	NS (.255)	1.44	1.08	31.0
U6 horizontal	−0.15	2.62	−6.64 to 5.73	NS (.759)	1.96	1.71	37.9
U6 vertical	−0.77	1.36	−3.18 to 2	S (.005)	1.29	0.88	24.1
Palatal plane	−1.53	2.80	−6.55 to 7.36	S (.007)	2.48	1.97	48.3

[a] NS indicates not significant; S, significant at the 5% level.

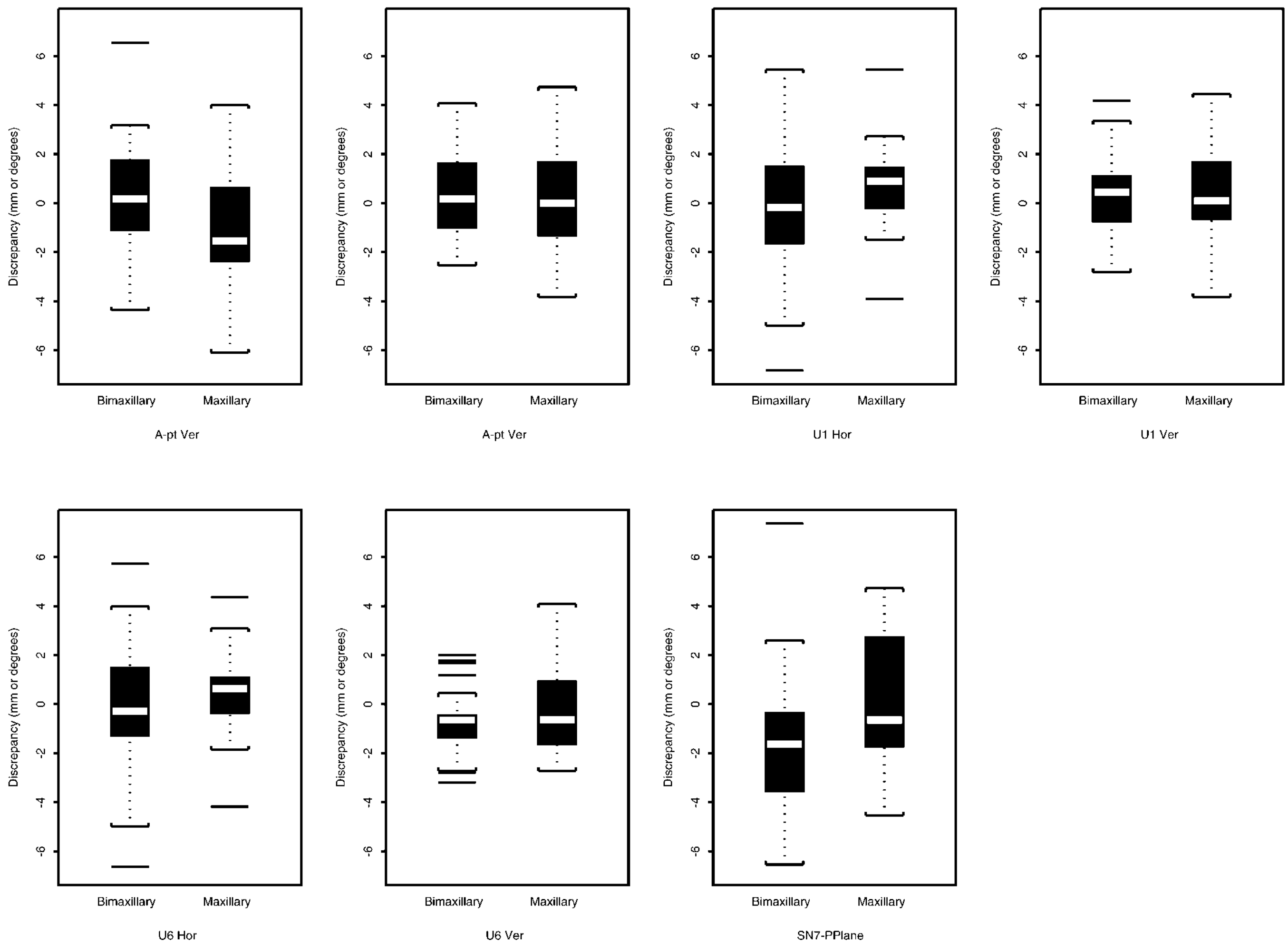

FIGURE 5. Effect of choice of surgical complexity.

TABLE 5. Comparison Between Maxillary and Bimaxillary Surgery

Landmark	*t*-test *P* value	Wilcoxon Test *P* value	Absolute Values Wilcoxon Test *P* value
A-pt horizontal	.068	.089	.295
A-pt vertical	.831	.870	.270
U1 horizontal	.250	.226	.270
U1 vertical	.928	.957	.913
U6 horizontal	.493	.414	.308
U6 vertical	.178	.454	.178
Palatal plane	.085	.146	.775

(Table 5). Despite the lack of a statistically significant difference in predicted landmark locations between the two groups, around 50% of the entire maxillary group was within two mm of prediction whereas around 70% of the bimaxillary group was within two mm of prediction. There were similar percentages of cases within one mm of prediction for both maxillary and bimaxillary cases (31% and 27%, respectively). The comparison of average discrepancy by surgical complexity is shown in Figure 6.

Effect of direction of primary maxillary movement

To evaluate the effect of the direction of maxillary movement, we divided the sample into three subgroups: downgraft ($n = 9$), impaction ($n = 15$), and advancement ($n = 16$) (Table 6). For patients who underwent combined advancement and impaction, for example five-mm advancement and three-mm impaction, the subject was placed into the advancement subgroup because that was the greater movement. Subjects were excluded if the amount of impaction and advancement was equal. If the maxilla was downgrafted, it was placed into the downgraft group. For surgical downgraft, there was only moderate evidence that the mean difference was nonzero for one measurement: U6 vertical measurements (two-sided *P* value, .032; 95% confidence interval, −1.74 to −0.10). This meant that the surgical teams on an average tended not to downgraft the maxilla as much as predicted.

For surgical impaction, there was insufficient evidence that the mean differences were nonzero for any of the landmarks. In the case of surgical advancement,

TABLE 6. Effect of Primary Direction of Maxillary Movement

Landmark	Mean	SD	Range	Significance (P value)[a]	Absolute Mean	Absolute SD	>2 units (%)
Downgraft	n = 9						
A-pt horizontal	−0.88	2.72	−6.09 to 2.82	NS (.359)	2.11	1.80	33.3
A-pt vertical	−0.42	2.51	−3.82 to 4.09	NS (.631)	1.99	1.42	33.3
U1 horizontal	−0.27	2.05	−4.82 to 2.27	NS (.703)	1.35	1.50	22.2
U1 vertical	−0.12	2.30	−3.82 to 4.18	NS (.877)	1.56	1.60	33.2
U6 horizontal	−0.19	1.78	−4 to 2.45	NS (.757)	1.20	1.25	22.2
U6 vertical	−0.92	1.07	−2.73 to 0.64	S (.032)	1.14	0.79	11.1
Palatal plane	−1.50	3.47	−6.55 to 4.73	NS (.231)	3.16	1.84	66.7
Impact	n = 15						
A-pt horizontal	0.93	2.69	−5.39 to 6.55	NS (.202)	1.98	2.00	33.3
A-pt vertical	0.43	1.71	−2.45 to 3.35	NS (.347)	1.47	0.89	26.7
U1 horizontal	1.16	2.65	−3.90 to 5.45	NS (.112)	2.16	1.85	46.7
U1 vertical	0.27	1.59	−2.82 to 2.45	NS (.521)	1.29	0.90	33.3
U6 horizontal	1.17	2.55	−4.18 to 5.73	NS (.096)	2.13	1.76	46.7
U6 vertical	−0.18	1.87	−3.18 to 3.07	NS (.709)	1.57	0.95	33.3
Palatal plane	−0.08	3.07	−4.73 to 7.36	NS (.920)	2.24	2.01	46.7
Advance	n = 16						
A-pt horizontal	−0.86	2.34	−4.36 to 3.09	NS (.164)	2.12	1.21	56.2
A-pt vertical	0.66	2.11	−2.55 to 4.74	NS (.230)	1.66	1.40	25.0
U1 horizontal	−0.77	2.51	−6.82 to 2.73	NS (.235)	1.89	1.77	25.0
U1 vertical	0.77	2.06	−2.73 to 4.46	NS (.158)	1.76	1.25	37.5
U6 horizontal	−0.64	2.47	−6.64 to 3.18	NS (.320)	1.81	1.75	25.0
U6 vertical	−0.61	1.75	−2.79 to 4.09	NS (.186)	1.47	1.07	25.0
Palatal plane	−1.55	2.44	−6.09 to 3.18	S (.023)	2.31	1.68	43.8

[a] NS indicates not significant; S, significant at the 5% level.

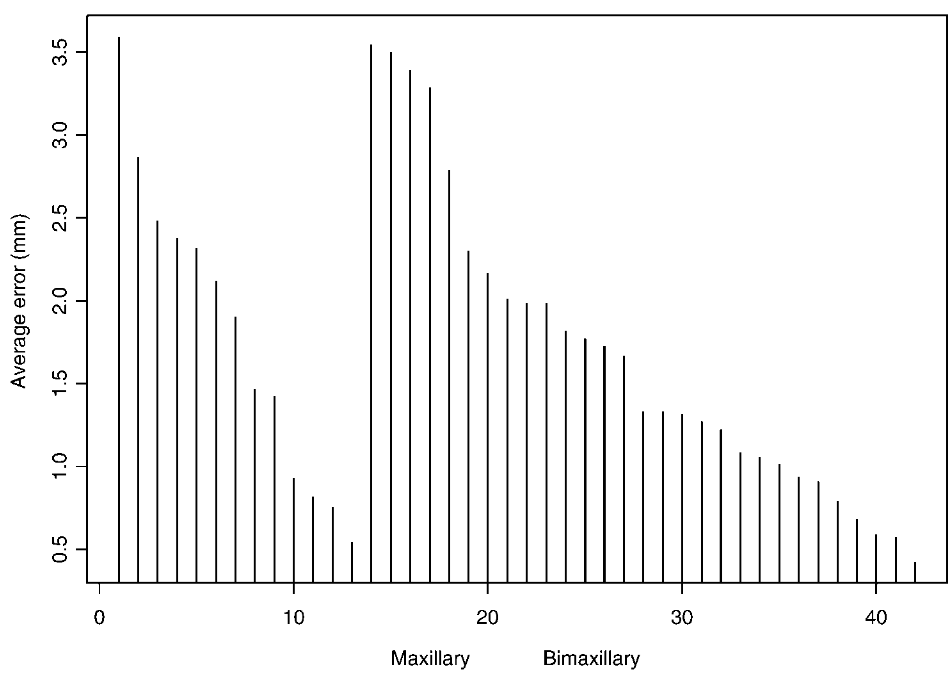

FIGURE 6. Comparison of average discrepancy (mm) by surgical complexity.

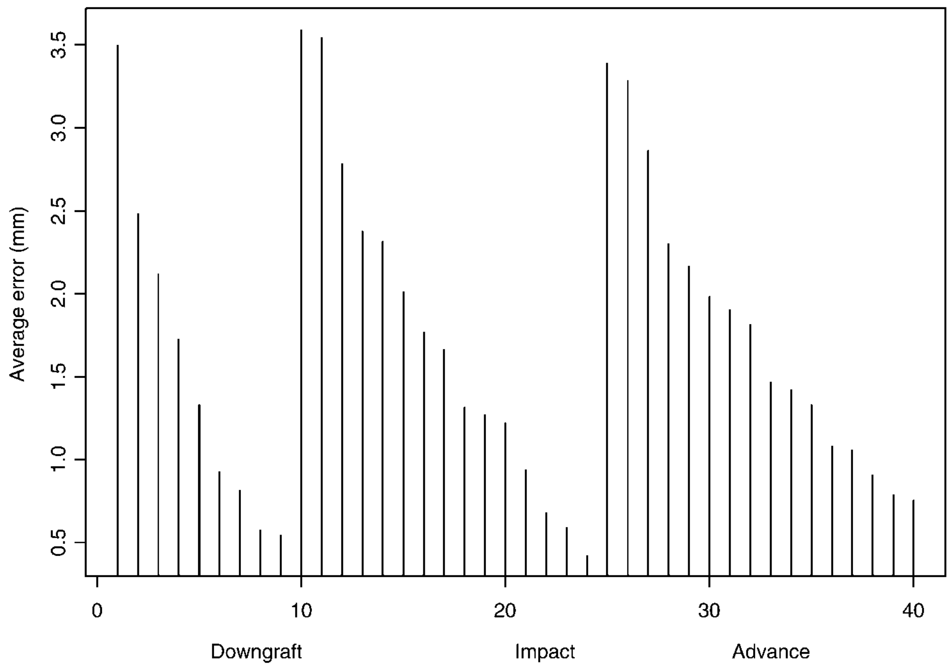

FIGURE 7. Comparison of average discrepancy (mm) by direction of maxillary movement.

TABLE 7. Comparison Between Surgical Downgraft, Impact, and Advance

Landmark	*P* value[a]
A-pt horizontal	.154
A-pt vertical	.619
U1 horizontal	.078
U1 vertical	.722
U6 horizontal	.092
U6 vertical	.712
Palatal plane	.495

[a] *P* values are from analysis of variance tests.

there was moderate evidence that the mean difference in actual palatal plane measurements and predicted palatal plane measurements was nonzero (two-sided *P* value, .023; 95% confidence interval, −2.85 to −0.25). This meant that, in cases where advancement was the greater movement, the surgical teams on an average tended to either over-rotate the maxilla if a clockwise rotation was predicted or under-rotate it if anticlockwise rotation was predicted.

Comparison of surgical downgraft, impaction, and advancement

There were no significant differences detected by analysis of variance tests between the three types of surgery (Figure 7; Table 7).

DISCUSSION

The focus of this study was on the accuracy of LeFort I maxillary osteotomy procedures. The results suggest that it is a very accurate procedure but has a wide range of variability. Approximately two-thirds of patients had their maxilla placed within two mm of prediction, and about a quarter were within one mm of prediction.

In this sample of patients, the environment in which the surgery was performed significantly influenced the accuracy of maxillary repositioning. Providing surgical treatment in a teaching hospital had an influence that cannot be underestimated. In this study, surgical treatment in a teaching hospital produced greater variation between the prediction and the outcome when compared with surgery delivered in the private environment (Figure 3). Although the accuracy of surgery in the private environment was similar to those quoted in a recent study by Jacobson and Sarver,[6] the results from the teaching hospital were not as impressive as those of Jacobson and Sarver, who reported 80% within two mm and 43% within one mm of prediction.

Why were some procedures not as accurate as others? One possible reason is that the surgeon could not or chose not to follow the surgical plan. This is one of the limitations of this study. That is, we are assuming that the surgeon has agreed with the surgical plan and carried it out. In this study, there are other factors that we have had to assume had no significant bearing on the discrepancy between the surgical outcomes and the surgical prediction. These include:

- Quality of all the relevant presurgical records (such as photos used in surgical prediction, presurgical study models, and bite registrations);

- Accuracy of the surgical prediction method(s);
- Accuracy of the model surgery;
- Method of surgical splint fabrication and subsequent fit before surgery;
- Experience and skill of the surgeon/surgical team;
- Type of surgical technique(s) used by the surgical team;
- Measurement methods and reference points used during surgery reposition the maxilla (for eg, use of a glabella screw).

It is worthwhile to note that any one of these factors may well have played a significant role in the accuracy of surgically repositioning the maxilla.

In light of such shortcomings, our results thus only provide an overall impression of how accurate the surgical outcome was with respect to the surgical prediction. Furthermore, Jacobson and Sarver[6] have commented this does not necessarily suggest that the surgical procedure was poor but only that it was different from the surgical plan.

Extensions of this study might consist of redoing the study without so many assumptions. For example, to exclude any cases, which we could not establish with confidence that the surgeon did follow the surgical plan, or excluding any cases that did not use the exact same surgical technique. We might also examine whether there is a possible correlation between the years of experience a surgeon had and the degree of discrepancy between surgical prediction and outcome.

CONCLUSIONS

- LeFort I maxillary osteotomy is a very accurate procedure but has a wide range of variability.
- Statistically significant differences were found between the predicted and actual postsurgical maxillary molar vertical position, and significant differences were also found for the palatal plane angular measurements.
- Surgical team 1 from the private environment had significantly less variation in discrepancy than surgical team 2 from the teaching hospital environment.
- There were no significant differences found when comparing single-jaw and bimaxillary surgery.
- There were no statistically significant differences found when comparing the primary direction of movement (impaction vs downgraft vs advancement).
- Overall, 66% of the results were within two mm of prediction and 26% of the results were within one mm of prediction.
- This study supports the use of a surgical prediction tracing and the need for surgeons to consult with the orthodontist regarding the surgical plan and agree to follow it.
- A number of factors influence the overall accuracy of surgical repositioning of the maxilla. Therefore, our results only provide us with an overall impression of how accurate the surgery result was with respect to the surgical prediction in our given sample. A large surgical discrepancy does not necessarily mean that the surgical procedure was poor but only that it was significantly different from the surgical plan.

ACKNOWLEDGMENTS

This study was funded by the Australian Society of Orthodontists Foundation for Orthodontic Research and the University of Western Australia Orthodontic Research Fund. The authors would like to acknowledge the assistance of Dr Alexandra Bremner, PhD, Biostatistical Consulting Service at the University of Western Australia School of Population Health, in managing the statistical analysis.

REFERENCES

1. 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.
2. Bryan DC, Hunt NP. Surgical accuracy in orthognathic surgery. *Br J Oral Maxillofac Surg.* 1993;1:343–349.
3. Cangialosi TJ, Chung JM, Elliot DF, Meistrell ME. Reliability of computer-generated prediction tracing. *Angle Orthod.* 1995;65:277–284.
4. 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;112:656–665.
5. Zhang X, Wang X, Zhang Z. [Accuracy of computerized aid diagnosis, surgical simulation and facial appearance prediction orthognathic surgery]. *Zhonghua Kou Qiang Za Zhi.* 1998;33:6–9.
6. Jacobson R, Sarver DM. The predictability of maxillary repositioning in Le Forte I orthognathic surgery. *Am J Orthod Dentofacial Orthop.* 2002;122:142–154.