European Journal of Orthodontics 26 (2004) 303–309
European Journal of Orthodontics vol. 26 no. 3

How predictable is orthognathic surgery?

Charlotte E. Eckhardt and Susan J. Cunningham
Orthodontic Department, Eastman Dental Institute, University College London, UK

SUMMARY There are a number of increasingly sophisticated techniques available for orthognathic treatment planning. All are based on the determination of the skeletal pattern and the position of the dentition. However, they all suffer from difficulties associated with predicting the soft tissue profile. The aim of this retrospective cephalometric investigation was, therefore, to compare the ability to predict accurately the outcome of orthognathic treatment using the 'hand planning' technique and the orthognathic planning and analysis (OPAL) computer program, with an emphasis on the soft tissue profile.

Seventy adult subjects were divided into two groups not specific for gender or age: the Class III patients had undergone bimaxillary surgery and the Class II patients sagittal split mandibular advancement. In each group, the pre-treatment and post-debond lateral cephalograms were utilized to calculate the actual orthodontic and surgical movements. These values were then used to produce a prediction using both the hand planning technique and the OPAL program. The resultant predictions were digitized using a customized computer program and compared with the actual outcome.

The results show that there was marked individual variation when planning by hand and using the OPAL program. In the mandibular surgery group, hand planning and OPAL were of similar accuracy and few points differed significantly between prediction and outcome. However, for the bimaxillary group, a number of points showed bias and the hand planning technique appeared to be more accurate than the OPAL program, particularly in the region of the lips. Although the usefulness of predictions is acknowledged, these results suggest that they should be used with a certain amount of caution.

Introduction

Orthodontic treatment aims to improve function by producing optimal occlusion and stability, and improving facial aesthetics. However, in some cases the malocclusion or facial disfigurement is so severe as to require more than orthodontics alone. These cases may be treated by a combination of orthodontics and orthognathic surgery.

Physical attractiveness is very important and the desire to improve facial appearance is a strong motivating factor in seeking treatment (Kiyak *et al.*, 1988). In view of this, the ability to predict the outcome of treatment is essential. The predictability of treatment depends on the relationship between the hard and soft tissues. However, it is generally accepted that soft tissue changes do not always mimic those of the underlying hard tissues due to a number of factors including the variation in the thickness of the soft tissues covering the face (Subtelny, 1959).

There are a number of techniques available for the planning of orthognathic treatment and these have become increasingly sophisticated over the years. These techniques have concentrated on determining the skeletal pattern and the position of the dentition, but a consistent problem has been the prediction of the soft tissue profile.

The aim of this study was to determine the accuracy of the 'hand planning' technique and the orthognathic planning and analysis (OPAL) computer program (Harradine and Birnie, 1997) for planning orthognathic treatment, with an emphasis on the soft tissue profile.

Materials and method

Materials

The investigation was retrospective, using lateral skull radiographs of subjects who had undergone a combination of orthodontics and orthognathic surgery. All were treated in the same hospital and as the magnification of all radiographs was identical, correction was not required. A sample size calculation was undertaken using Altman's nomogram (Altman, 1991) based on a power of 80 per cent and a significance level of 0.05. Group 1 comprised 30 patients with a Class II malocclusion who underwent a bilateral sagittal split advancement only. Group 2 comprised 40 patients with a Class III malocclusion who had undergone a bimaxillary osteotomy (with both antero-posterior and vertical maxillary movements). These two groups were selected as they are the two most common surgical procedures undertaken at the Eastman Dental Hospital. All subjects were adults and the groups were not gender specific.

The subjects were selected on the following criteria: they were all Caucasian; they had no congenital deformities, such as cleft lip and/or palate; they had undergone orthodontics and orthognathic surgery, but no genioplasty or rhinoplasty; lateral cephalometric radiographs were available at the following time points: T1 (pre-treatment) and T2 (debond).

Method

For each subject, the pre-treatment and debond lateral cephalograms were traced by hand and also digitized using the OPAL program. The hand tracings and OPAL printouts were then digitized using a customized digitizing program (GELA). This program measured a series of hard and soft tissue points to x and y reference lines as well as a series of angles (Figures 1–4). From these measurements, the actual dental and skeletal changes which occurred as a result of treatment were calculated.

Using the pre-surgical orthodontic changes and surgical movements, an orthognathic prediction plan was carried out for each subject using both the hand planning technique, based on the use of tracing paper overlays, and the OPAL program (Figure 5).

The hand and OPAL predictions were then digitized using the customized GELA program. Finally, the hand planned prediction was compared with the end of treatment hand tracing and the OPAL prediction with the end of treatment OPAL printout.

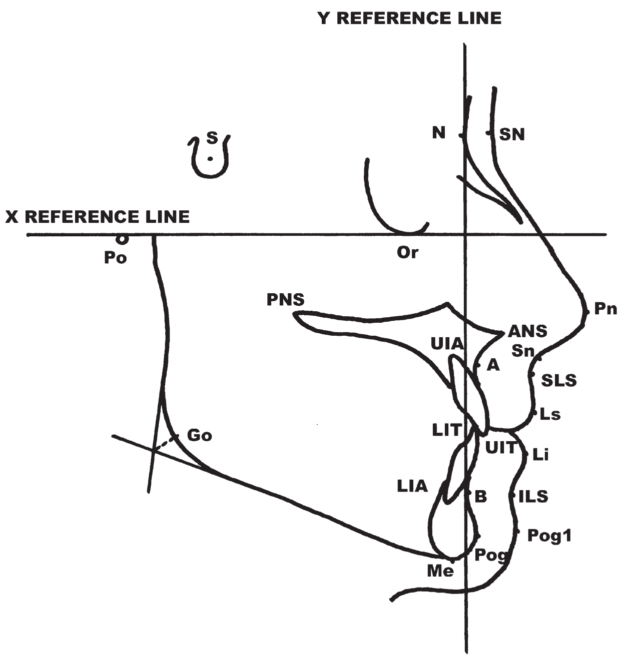

Figure 1 Cephalometric landmarks and reference lines. Soft and hard tissue points measured using the GELA computer program. S, sella; N, nasion; Po, porion; ANS, anterior nasal spine; PNS, posterior nasal spine; Or, orbitale; A, point A; UIT, upper incisor tip; UIA, upper incisor apex; LIA, lower incisor apex; LIT, lower incisor tip; B, point B; Pog, pogonion; Go, gonion; Me, menton; SN, soft tissue nasion; Pn, pronasale; Sn, subnasale; SLS, superior labial sulcus; Ls, labrale superius; Li, labrale inferius; ILS, inferior labial sulcus; Pog1, soft tissue pogonion.

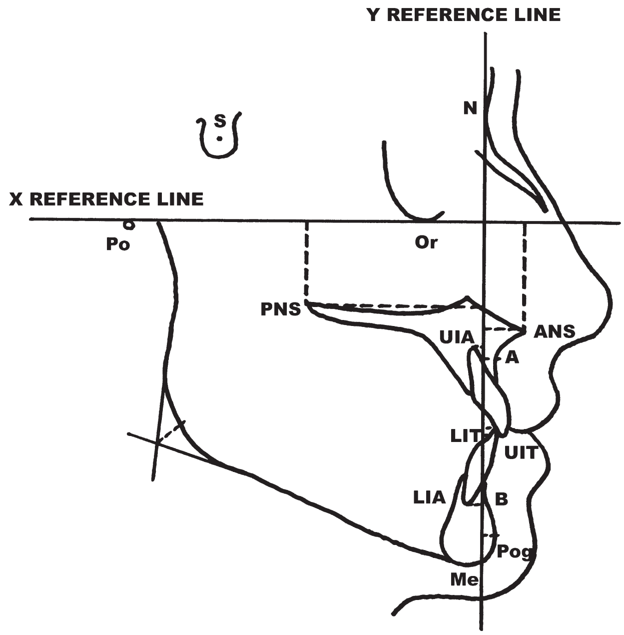

Figure 2 Hard tissue measurements. Hard tissue points measured using the GELA computer program. S, sella; N, nasion; UIA–Y, perpendicular distance from UIA to y reference line (YRL); UIT–Y, perpendicular distance from UIT to YRL; LIA–Y, perpendicular distance from LIA to YRL; LIT–Y, perpendicular distance from LIT to YRL; PNS–X, perpendicular distance from posterior nasal spine to x reference line (XRL); ANS–X, perpendicular distance from anterior nasal spine to XRL; PNS–Y, perpendicular distance from posterior nasal spine to YRL; ANS–Y, perpendicular distance from anterior nasal spine to YRL; A–Y, perpendicular distance from point A to YRL; B–Y, perpendicular distance from point B to YRL; Pog–Y, perpendicular distance from pogonion to YRL.

Error study

In order to establish the error of the method, 20 of the 70 cases were randomly selected: 10 from the mandibular advancement group and 10 from the bimaxillary group. In each of these cases, the pre-treatment cephalogram and the debond cephalogram were retraced and digitized using the OPAL program. The surgical movements were then recalculated for both groups.

In a second error study, the predictions were repeated using the original calculated treatment plan and compared with the original prediction.

Statistical method

The results were analysed using three statistical methods.

The paired t-test (Petrie and Watson, 1999) was used to determine whether the differences between each point for the actual outcome and the prediction were significant, having first established that the differences for each point were normally distributed. Due to the large number of comparisons being employed in this study, there was a possibility that some of the readings

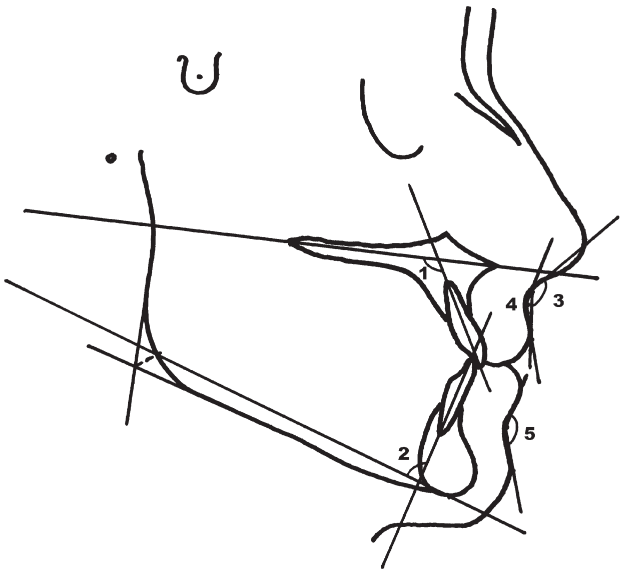

Figure 3 Cephalometric planes and angles. Angle 1, the angle subtended by the maxillary plane and the long axis of the most anterior maxillary incisor; angle 2, the angle subtended by the mandibular plane and the long axis of the most anterior incisor; angle 3 (nasolabial angle), the angle subtended by a line joining pronasale (Pn), subnasale (Sn) and labrale superius (Ls); angle 4 (upper lip contour angle), the angle subtended by a line joining subnasale (Sn), superior labial sulcus (SLS) and labrale superius (Ls); angle 5 (lower lip contour angle), the angle subtended by a line joining labrale inferius (Li), inferior labial sulcus (ILS) and soft tissue pogonion (Pog1).

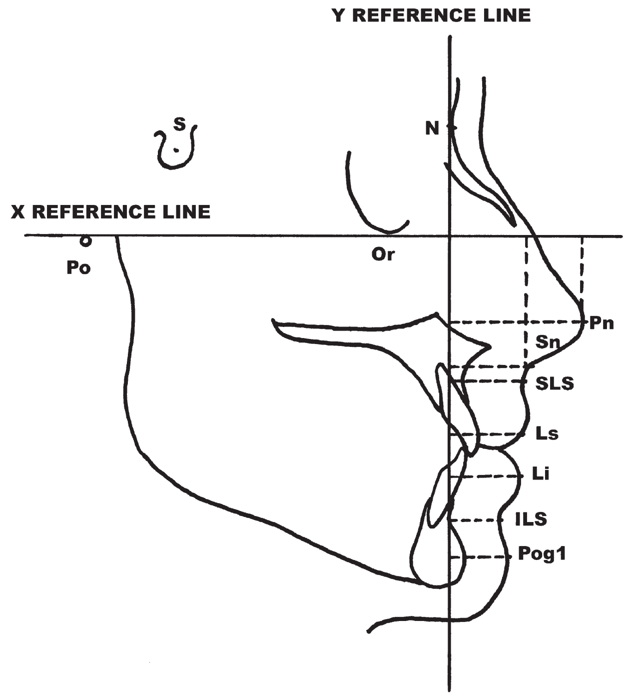

Figure 4 Soft tissue measurements. Soft tissue points measured using the GELA computer program. Pn–X, perpendicular distance from pronasale to *x* reference line (XRL); Pn–Y, perpendicular distance from pronasale to *y* reference line (YRL); Sn–X, perpendicular distance from subnasale to XRL; Sn–Y, perpendicular distance from subnasale to YRL; SLS–Y, perpendicular distance from superior labial sulcus to YRL; Ls–Y, perpendicular distance from labrale superius to YRL; Li–Y, perpendicular distance from labrale inferius to YRL; ILS–Y, perpendicular distance from inferior labial sulcus to YRL; Pog1–Y, perpendicular distance from soft tissue pogonion to YRL.

would be significant purely by chance. In view of this, the level of significance was increased from $P = 0.05$ to $P = 0.02$ to make the tests more stringent.

The Bland and Altman method (Altman, 1991) was used to compare the predictions with the actual outcomes and to establish whether there was any bias. In addition, limits of agreement were calculated. This test was also used in the error groups to establish whether the calculated plans (error study 1) and the predictions (error study 2) were repeatable.

An intraclass correlation coefficient (ICC) (Streiner and Norman, 1995) was also used to establish repeatability. An ICC is preferable to using a Pearson or Spearman correlation because it will produce a value of $r = 1.0$ only if all observations on each subject are identical and the intercept is at 0.

Results

The results of the error group studies showed that there were no significant differences between the first and second tests and the limits of agreement were within clinically acceptable boundaries.

The results of the main study are shown in Tables 1–3. Only the significant findings are presented.

Mandibular advancement group

In general, the results of the hand predictions for the mandibular advancement group were good, particularly for the linear values (Table 1). The paired *t*-test showed that none of the linear measurements were significantly different to the actual outcome. However, two of the angular measurements showed bias: the upper lip contour angle ($P = 0.002$) was consistently overestimated and the lower lip contour angle ($P = 0.001$) was consistently underestimated in the prediction. The ICC values (Table 2) were good for the linear measurements, with a minimum value of 0.93, although the angular measurements were less acceptable, with the nasolabial angle and the upper lip contour angle well below acceptable limits (upper lip contour angle $r = 0.19$ and nasolabial angle $r = 0.53$). This was reflected by the marked variation in the limits of agreement (Table 3).

For the OPAL program, there was systematic bias affecting two of the linear measurements (labrale superius and labrale inferius to the *y* reference line) with both values being overestimated in the prediction.

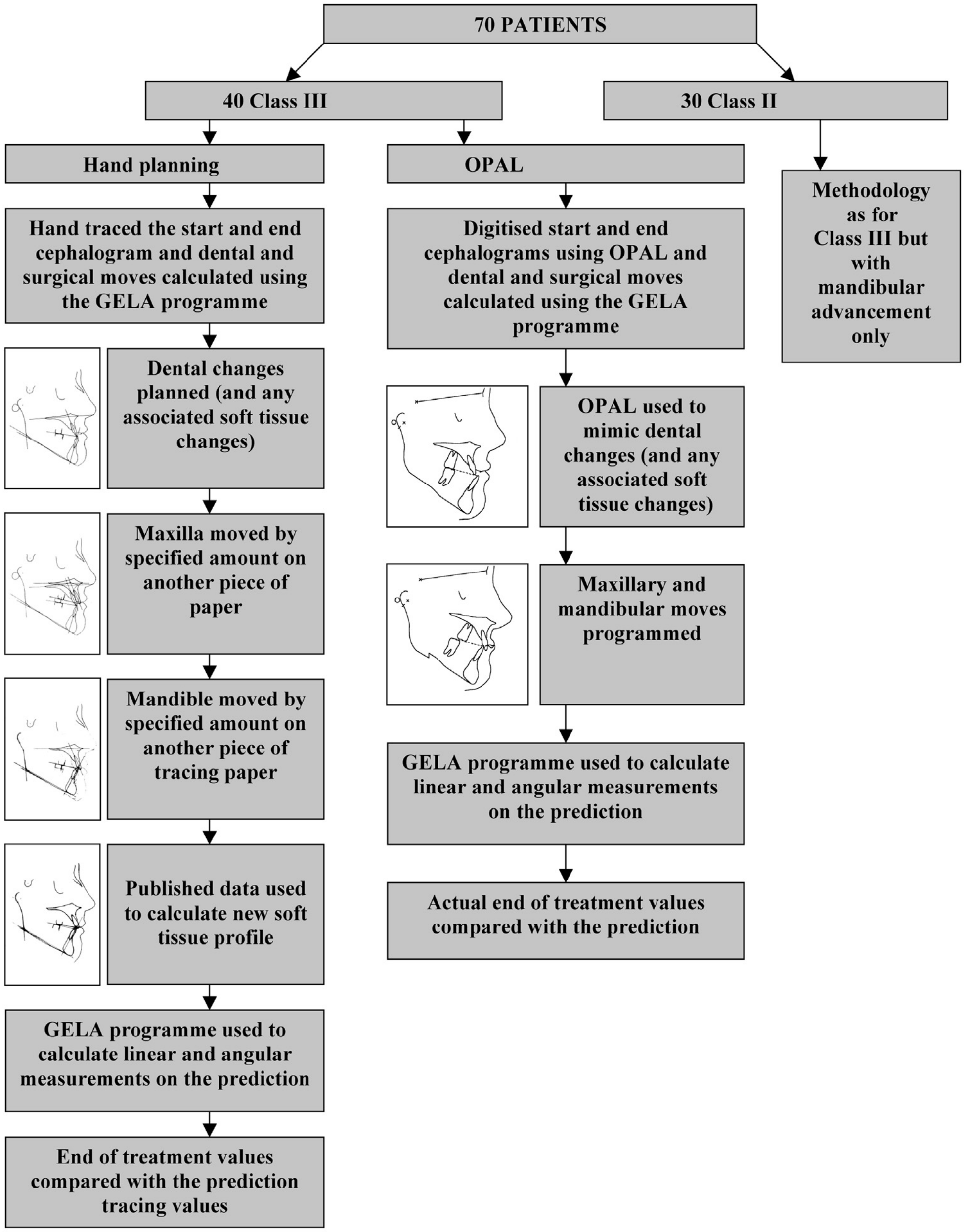

Figure 5 A flow diagram illustrating the methodology.

In addition, one of the angular measurements (lower lip contour angle) showed bias and was consistently underestimated. Although the ICCs were good for the linear measurements, again, the angular measurements were less acceptable, with all three angles having a coefficient below the acceptable value of 0.70. This was also reflected in the marked individual variation shown by the limits of agreement.

Bimaxillary group

In the bimaxillary group, the hand predictions showed bias for four of the linear and one of the angular measurements (Table 1). Pronasale to *x* axis, labrale superius to *y* axis and labrale inferius to *y* axis were all underestimated and inferior labial sulcus to *y* axis and lower lip contour angle were both overestimated in the prediction tracing. It is noteworthy that there were more significant differences for the bimaxillary group than for the single jaw group.

In the OPAL group, there was also bias for several linear measurements and two of the angular measurements. Pronasale to *x* axis, subnasale to *y* axis, superior labial sulcus to *y* axis and upper lip contour were all underestimated and inferior labial sulcus to *y* axis and the lower lip contour angle were both overestimated in

Table 1 Paired *t*-test to establish whether there was systematic bias between planned and actual outcomes.

		Measurement	Mean difference	*P*-value
Mandibular advancement group (n = 30)				
	Hand planning	SnSLSLs (upper lip contour angle)	–0.81	0.002
		LiILSPog1 (lower lip contour angle)	9.1	0.001
	OPAL	Ls–Y (labrale superius to *y* axis)	–1.8	0.001
		Li–Y (labrale inferius to *y* axis)	–1.8	0.002
		LiILSPog1 (lower lip contour angle)	23.4	0.001
Bimaxillary group (n = 40)				
	Hand planning	Pn–X (pronasale to *x* axis)	0.5	0.010
		Ls–Y (labrale superius to *y* axis)	1.0	<0.001
		Li–Y (labrale inferius to *y* axis)	0.9	0.015
		ILS–Y (inferior labial sulcus to *y* axis)	–0.6	0.019
		LiILSPog1 (lower lip contour angle)	–10.5	<0.001
	OPAL	Pn–X (pronasale to *x* axis)	1.1	0.002
		Sn–Y (subnasale to *y* axis)	1.3	0.004
		SLS–Y (superior labial sulcus to *y* axis)	1.6	0.007
		ILS–Y (inferior labial sulcus to *y* axis)	–2.1	0.007
		SnSLSLs (upper lip contour angle)	7.7	0.007
		LiILSPog1 (lower lip contour angle)	–11.8	<0.001

OPAL, orthognathic planning and analysis computer program.

Table 2 The intraclass correlation coefficient ranges for all the linear and angular measurements.

		Mandibular advancement	Bimaxillary surgery
Hand planning	Linear (mm)	0.93 to 0.99	0.91 to 0.99
	Angular (degrees)	0.19 to 0.77	0.54 to 0.91
OPAL	Linear (mm)	0.91 to 9.7	0.89 to 0.96
	Angular (degrees)	0.41 to 0.70	–0.18 to 0.58

OPAL, orthognathic planning and analysis computer program.
Minimum acceptable value normally 0.70.

the computer prediction. Figure 6 illustrates the Bland and Altman plot for the lower lip contour angle. The majority of points lie above the zero difference line and this shows the consistent overestimation in the predictions.

The limits of agreement appeared slightly better for hand planning than OPAL (Table 3). In addition, the ICCs tended to be better for the hand planning method. However, the ICCs for the angular measurements were, in the most part, unacceptably low, with the exception of the nasolabial angle in hand planning ($r = 0.91$; Table 2).

Discussion

A number of methods are now available for the prediction of outcomes of orthognathic treatment. However, few studies have compared these methods to establish the accuracy of the predictions with respect to the soft tissue profile (Eales *et al.*, 1994).

Statistical analysis of the mandibular surgery group showed that, in general, few points were significantly different to the actual outcomes. However, for the

Table 3 Limits of agreement for significant findings from the Bland and Altman method for several angular and linear measurements.

		Hand planning	OPAL
Mandibular advancement group (n = 30)			
	Linear (mm)	–	Ls–Y (labrale superius to *y* axis) –6.8 to 3.2 Li–Y (labrale inferius to *y* axis) –7.7 to 3.9
	Angular (degrees)	SnSLSLs (upper lip contour angle) –33.7 to 17.5 LiILSPog1 (lower lip contour angle) –15.3 to 33.5	LiILSPog1 (lower lip contour angle) –9.0 to 55.9
Bimaxillary group (n = 40)			
	Linear (mm)	Pn–X (pronasale to *x* axis) –1.9 to 2.9 Ls–Y (labrale superius to *y* axis) –2.0 to 4.0 Li–Y (labrale inferius to *y* axis) –3.5 to 5.3 ILS–Y (inferior labial sulcus to *y* axis) –4.1 to 2.7	Pn–X (pronasale to *x* axis) –3.1 to 5.3 Sn–Y (subnasale to *y* axis) –3.9 to 6.5 SLS–Y (superior labial sulcus to *y* axis) –5.4 to 8.6 ILS–Y (inferior labial sulcus to *y* axis) –11.1 to 6.9
	Angular (degrees)	LiILSPog1 (lower lip contour angle) –30.7 to 9.7	SnSLSLs (upper lip contour angle) –25.7 to 41.1 LiILSPog1 (lower lip contour angle) –33.3 to 9.9

OPAL, orthognathic planning and analysis computer program.

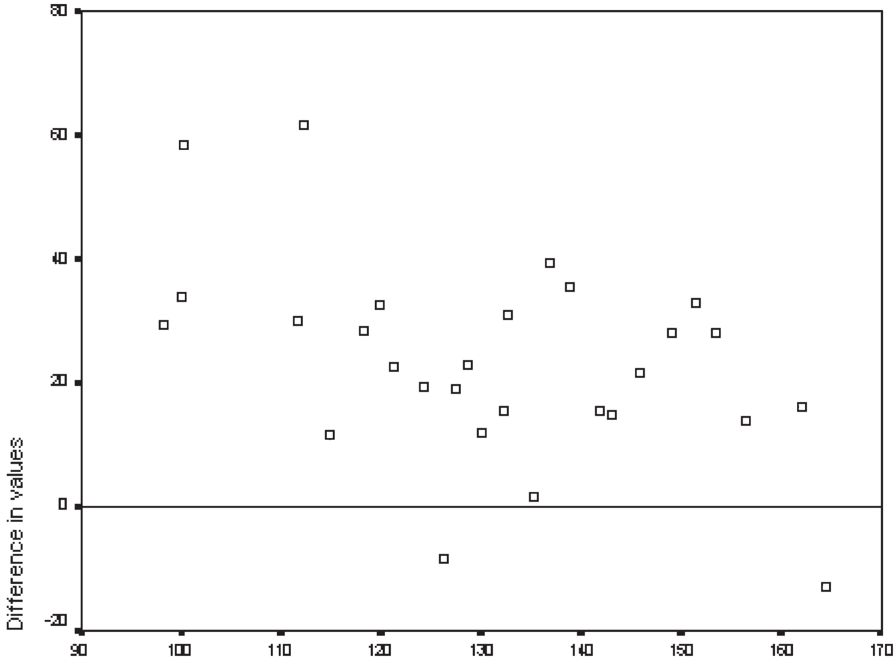

Figure 6 Bland and Altman plot for the lower lip contour angle in the bimaxillary OPAL group. X axis = average value, i.e. (actual value + predicted value)/2. Y axis = difference between the two values, i.e. predicted value – actual value.

bimaxillary group, a number of points showed variation for both the hand planning and OPAL techniques. This finding is not altogether unexpected as the actual surgical treatment plans and the predictions are more complicated. In addition, the mandibular movements are, to some extent, reliant on the maxillary movements and, therefore, the potential to compound the error is greater.

The linear measurements were generally well predicted, particularly for the mandibular group, although less so for the bimaxillary group, but the angular measurements were much more problematic. In the mandibular group, two of the three angular measurements (lower lip contour angle in both groups) were underestimated in the prediction and one was overestimated (upper lip contour angle in the hand planning group). In the bimaxillary group, the reverse applied, with the lower lip contour angle being overestimated in both groups and the upper lip contour angle being underestimated by OPAL. The problems associated with the prediction of angular measurements is probably due to the angles being based on three points. Therefore, the variation in these points is compounded (i.e. the points themselves may not be significantly different but when added together produce a very different angle). For example, the lower lip contour angle showed limits of agreement between –33.3 and 9.9 degrees for the bimaxillary OPAL prediction group. This reflects the vague instructions provided in the data used for the planning of this region. For example, the published data used in the hand planning method give instructions to 'increase' or 'decrease and evert' with regard to the nose and lip, respectively; instructions which are open to interpretation and, therefore, error. Other authors have attempted to quantify the soft tissue changes following bony moves, i.e. the upper lip will shorten approximately 10–40 per cent of the anterior impaction of the maxilla, or the nasal tip will elevate approximately 20 per cent of the anterior impaction of the maxilla (Wolford and Fields, 1999). However, these are still vague instructions and are open to interpretation by the clinician.

In the case of the OPAL program, poor prediction of the lip region was due to the volume of the soft tissues remaining constant, thus the profile line becomes stretched and distorted. This finding was also noted by Eales *et al.* (1994) in their study of the COG Soft 3.4 program, which was the predecessor to the OPAL program. Although the program has now been updated, it would appear that prediction in the region of the lips is still problematic. In contrast, hand tracing allowed more 'artistic licence'. However, this in itself can be problematic.

It must also be borne in mind that just because a difference is not statistically significant does not mean it would not have clinical relevance. Some of the points which showed no significant difference with the *t*-test still had limits of agreement which may be considered clinically significant. As medicine becomes more litigious it is increasingly important to be able to account for treatment outcomes, especially in elective procedures.

Patients are more demanding and, as such, clinicians must be more accountable, both prior to treatment and in defence of the treatment carried out. Due to the complex nature of combined orthodontic–orthognathic treatment, the prediction of outcomes is very difficult. This is, in part, due to individual differences. For example, soft tissues vary in their tonicity, which in turn alters the way they adapt to changes in the underlying hard tissues. To improve the ability to predict outcomes, larger studies must be carried out incorporating all possible surgical moves. Only by doing this will it be possible to calculate the ratios of soft tissue:hard tissue movements more accurately.

The findings reinforce the importance of using clinical judgement and experience in planning orthognathic cases. It is also important to warn patients that predictions are purely a guide to the outcome and not a guarantee.

Conclusions

1. There was a great deal of individual variation when planning by hand and using the OPAL method.
2. On average, in the mandibular surgery groups, few points/angles differed significantly between prediction and outcome. However, in the bimaxillary surgery groups, a number of points/angles showed bias.
3. Linear measurements were predicted better than angular measurements.
4. For the mandibular surgery group there was little difference between hand and OPAL planning. In contrast, for the bimaxillary group, the accuracy attained when planning by hand appeared to be better than with OPAL.
5. The main problems with the OPAL computer program were in the region of the lips.

Address for correspondence

C. E. Eckhardt
Orthodontic Department
Eastman Dental Institute
256 Gray's Inn Road
London WC1X 8LD
UK

Acknowledgement

The authors would like to thank Mr S. L. Newell for writing the customized computer program (GELA).

References

Altman D G 1991 Practical statistics for medical research. Chapman & Hall/CRC, Boca Raton, Florida

Eales E A, Jones M L, Newton C, Sugar A W 1994 The accuracy of computerised prediction of the soft tissue profile: a study of 25 patients treated by means of the Le Fort I osteotomy. International Journal of Adult Orthodontics and Orthognathic Surgery 9: 141–152

Harradine N W T, Birnie D J (written by Chauvet G) 1997 Orthognathic planning and analysis (OPAL) computer program. Bristol University, UK

Kiyak H A, Vitaliano P P, Crinean J 1988 Patients' expectations as predictors of orthognathic surgery outcomes. Health Psychology 7: 251–268

Petrie A, Watson P 1999 Statistics for veterinary and animal science. Blackwell Science, Oxford

Streiner D L, Norman G R 1995 Health measurement scales. A practical guide to their development and use, 2nd edn. Oxford University Press, Oxford

Subtelny J D 1959 A longitudinal study of soft-tissue facial structures and their profile characteristics defined in relation to underlying skeletal structures. American Journal of Orthodontics 45: 481–507

Wolford L M, Fields R T 1999 Surgical planning. In: Booth P W, Schendel S A, Hausamen J-E (eds) Maxillofacial surgery. Churchill Livingstone, Edinburgh, pp. 1205–1257