Journal of Orthodontics, Vol. 30, 2003, 149–154

SCIENTIFIC SECTION

The validity of computerized orthognathic predictions

R. R. J. Cousley, E. Grant and J. D. Kindelan
York District Hospital, York, UK

Abstract

Objective: utilizing OPAL™ cephalometric prediction software.

Design: A retrospective investigation involving the random selection of Class II orthognathic patients from surgical records.

Subjects: These 25 cases had undergone treatment aimed at producing Class I incisors. This involved fixed orthodontic appliances and a mandibular advancement osteotomy with rigid internal fixation.

Methods: Lateral cephalographs from three key stages were digitized and processed using the OPAL software. Pre-treatment predictions were generated and compared with the actual clinical changes.

Results: Prediction of some of the principal OPAL variables (SNA, ANB, LAFH%, OJ, OB) was reasonably accurate in terms of mean values. However, there were large individual variations for most measurements, and prediction of Wits, MxP/MnP, LAFH, and LPFH was prone to systematic error. In particular, there was a tendency towards over-prediction of the surgically-induced backward mandibular rotation.

Conclusion: In lieu of further validation caution should be exercised with the interpretation of individual OPAL predictions, especially vertical skeletal changes, and an explanation given to patients that orthognathic predictions are based on generalizations.

Index words:
Class II, computerized prediction, OPAL, orthognathic surgery

Received 29 August 2002; accepted 26 September 2002

Introduction

The prediction of treatment outcome is an important part of orthognathic planning and the process of informed consent. The orthodontic and surgical changes must be described accurately prior to treatment in order to assess the treatment's feasibility and optimize case management. Nowadays a variety of computerized analyses of lateral cephalographs are used to predict treatment change in the antero-posterior and vertical facial planes, e.g. Dentofacial Planner™,[1–4] OPAL™, Quick Ceph™,[1,5–7] and TIOPS™.[8–10] Since 1982, many UK teams have used a prediction package called COGSOFT™ (Consultant Orthodontists Group Software).[11] Recently, this has been updated and released in a Microsoft Windows™ format as OPAL (Orthognathic Prediction Analysis). This software enables simulation of the effects of incisor decompensation and surgical jaw movements, and illustrates these changes in terms of quantitative values and a jaw/profile silhouette (based on established hard-soft tissue ratios). However, whilst the OPAL software is widely used there is only limited literature available on its validity[12] or on the computerized prediction of skeletal changes in general.[3,5,7,8–10] Instead, the relevant orthognathic literature has focused on the analysis of soft tissue/profile prediction errors (from the perspective of visual treatment objectives and patient counselling) and surgical accuracy.[13] The few hard tissue prediction studies have had a number of limitations, including the use of heterogeneous samples (in terms of malocclusions types or surgical procedures) and small case numbers.[3,5,7,8–10] In addition, there has not been a thorough evaluation of the pre-treatment prediction stage. Consequently, this study aimed to assess the accuracy of the OPAL orthognathic prediction process in the calculation of the total treatment hard tissue and dental changes.

Address for correspondence: Richard Cousley, Orthodontic Department, Peterborough District Hospital, Thorpe Road, Peterborough, Cambridgeshire, PE3 6DA, UK. Email: Richard.Cousley@pbh-tr.nhs.uk

Material and methods

Data collection

All adult Class II surgical cases were identified retrospectively from operating theatre records at York District Hospital. Each patient was assigned a number and a random sample of 25 cases was then selected from this cohort according to the following criteria:

1. Surgery was performed between 1990 and 1999 to allow for an appropriate follow-up time.
2. Clinical management, including treatment planning, had been carried out by a combined team of two orthodontists and two surgeons, where the agreed objective was a Class I incisor relationship. No deliberate surgical over-correction was performed.
3. Treatment must have involved orthodontic fixed appliances combined with a bilateral sagittal split mandibular advancement osteotomy and internal fixation. Cases involving bimaxillary osteotomies or any additional surgical procedures were excluded.
4. Sufficient clinical and radiographic records had to be available and additional selection criteria were that the cephalographs had been taken at a standardized magnification, with the head in its natural head posture, the mandible in the retruded position, and the labial soft tissues at rest. The pre-treatment, pre-surgery, and 1-year post-surgery lateral cephalographs were utilized in this study.

Each lateral cephalograph was digitized directly using a backlight-illuminated platform and a digitizing pad. Twenty-three radiographic landmarks were identified and digitized, along with the continuous (stream mode) tracing of five soft and hard tissue outlines. All of these tracings were undertaken by one operator (RC) and the radiographs for each individual were digitized serially in order to minimize random error variance, as recommended by Houston.[14] The data were processed using the OPAL software. Significantly, the standard OPAL analysis utilizes values based on the Eastman Analysis,[15] which is widely used by UK orthodontists. Twelve specific linear and angular dental and skeletal measurements were identified as being most relevant for this study: SNA, SNB, ANB, MxP/MnP, LAFH%, LAFH, LPFH, OJ, OB, U1/MxP, L1/MnP, and Wits.[16] These were calculated and rounded to the nearest 0.1 mm/degree.

Method error study

Twenty-five cephalographs were randomly selected and retraced at least 1 month after the original recordings. As recommended by Houston[14] and Battagel,[17] the method error was assessed for error variance using Dahlberg's formula: mean square error (SE^2) = $\Sigma d^2/2n$ (where d is the difference between repeated measurements and n is the number of radiographs recorded). A repeated measures t-test was also performed to assess systematic error.

Cephalometric analysis

Each individual's cephalometric tracings were superimposed to evaluate the actual changes that occurred during the treatment period. The OPAL software was used to generate a pre-treatment prediction of the orthodontic and surgical treatment, where the incisor inclinations were normalized (according to Eastman values) and the mandible advanced to produce a Class I incisor relationship. Predicted changes were calculated as the differences between the prediction analysis and baseline measurements. This prediction was then compared to the actual post-treatment measurements, and the differences analysed in terms of data distribution and by a paired t-test.

Results

Study sample details

The mean pre-treatment cephalometric values of this sample (Table 1) confirm that it was representative of patients with mandibular retrognathia undergoing orthodontic treatment combined with mandibular advancement surgery. The sample was homogenous in terms of the initial Class II skeletal discrepancy despite the inclusion of both division 1 and 2 cases. The influence of these initial incisor inclinations was also negated by reliance on relative changes as opposed to absolute values. At the end of treatment all of the mean angular values approximated standard Class I values (Table 1). Surgery increased the mean values for LAFH, LAFH%, and MxP/MnP by 2.9 mm, 1.1% and 4.1 degrees respectively, but the LPFH mean reduced by 3.8 mm.

Method error study

Students' t-tests of the repeated measures showed a systematic error in only one variable: L1/MnP (t = 3.48,

Table 1 Cephalometric data for the study sample before and after combined treatment.

Variable	Pre–treatment				1 Year post–treatment				Mean post–op change*	Mean total change
	Mean	–95% CI	+95% CI	SD	Mean	–95% CI	+95% CI	SD		
SNA	80.7	79.3	82.1	3.4	80.4	78.8	82.0	3.9	–0.3	–0.3
SNB	73.9	72.4	75.4	3.6	76.5	74.8	78.2	4.2	2.7	2.6
ANB	6.8	6.2	7.5	1.6	3.9	3.1	4.7	1.9	–2.9	–3.0
MxP/MnP	23.8	21.0	26.7	6.9	27.9	24.8	31.0	7.5	4.1	4.1
LAFH%	53.4	52.3	54.5	2.7	55.2	54.1	56.3	2.6	1.1	1.8
LAFH	58.1	56.1	60.2	4.9	63.8	61.4	66.1	5.6	2.9	5.7
LPFH	42.4	40.0	44.8	5.8	40.5	37.7	43.3	6.8	–3.8	–1.8
Wits	5.9	4.6	7.1	3.0	2.1	1.3	3.0	2.1	–4.7	–3.7
OJ	10.2	8.9	11.4	3.0	3.1	2.7	3.4	0.8	–6.0	–7.1
OB	4.8	3.7	6.0	2.8	2.2	1.7	2.8	1.3	–1.5	–2.6
U1/MxP	111.1	107.1	115.2	9.8	109.0	105.6	112.4	8.3	–3.0	–2.1
L1/MnP	92.9	89.4	96.4	8.4	91.4	88.4	94.4	7.2	–3.2	–1.4

*Where the post-operative change compares the pre-operative and long-term values.

$P < 0.01$, df = 24). Dahlberg's values demonstrated that random error for angular measurements ranged from 0.27 to 1.21 degrees for SNB and L1/MnP respectively. Similarly, linear measurement random errors ranged from 0.18 mm for the OJ to 0.65 mm for LPFH. Since some measurements, especially L1/MnP, U1/MxP and MxP/MnP, were susceptible to error caution has been exercised in extrapolating from the data where the differences were less than 0.5 unit. This arbitrary level was suggested following a meta-analysis of cephalometric errors.[18]

Comparison of the predicted and actual cephalometric changes

Five out of the 12 OPAL measurements (SNA, ANB, LAFH%, OJ and OB) displayed small mean differences (less than 0.5 unit) between the predicted and actual values (Table 2). The SNA, LAFH%, OJ, and OB results also showed reasonable consistency with narrow data distributions. Conversely, the largest mean discrepancy was observed for the Wits variable (5.9 mm) and three other variables had clinically relevant differences: MxP/MnP, LAFH and LPFH. Paired *t*-tests also revealed statistically significant differences ($P < 0.05$) for the Wits, LAFH and LAFH% values, and LPFH had marginal significance ($P = 0.06$). However, interpretation of these apparent prediction errors is complicated by the occurrence of residual vertical facial growth in many cases during the treatment period. In reality, the LAFH increased and LPFH reduced between the pre- and post-treatment stages (Table 1), but OPAL underpredicted these changes. This was associated with a clinically significant systematic error in MxP/MnP. This value was over-predicted by a mean of 1.9 degrees, although this result was not statistically significant. MxP/MnP and four other variables (LPFH, Wits, L1/MnP and U1/Man) also had high standard deviations and wide confidence intervals.

Discussion

There was little mean difference between the actual and predicted values for five of the 12 OPAL cephalometric measurements (SNA, ANB, LAFH%, OJ and OB) and the differences were not statistically significant for nine values (Table 2). SNA, LAFH%, OJ and OB also exhibited small standard deviations, although the difference in LAFH% was statistically significant ($P = 0.03$). These findings are comparable to those of Eales *et al.*,[12] where OPAL's precursor, COGSOFT (3.4)^{TM}, was utilized in the prediction of soft tissue changes. Similarly, a study of 16 mandibular advancement cases concluded that Quick Ceph prediction tracings were generally similar to the post-operative appearance and that computerized prediction of skeletal change is more accurate than the manual tracing technique.[6] Comparable findings were also observed in a recent study of Quick Ceph Image Pro^{TM} software in 28 heterogeneous cases.[7] These authors found statistically insignificant differences in 10 out of 14 measurements and concluded that the mean differences in the other four variables (ANB, FMA, U1/Max, Wits) were not clinically significant. However, many orthodontists may arguably view

Table 2 Relative differences between the pre-treatment OPAL prediction and the cephalometric changes associated with the treatment period

Variable	Mean	–95% CI	+95% CI	Min	Max	SD	*t* value	*P* value
SNA	0.35	–0.18	0.88	–1.4	3.9	1.29	1.36	0.19
SNB	0.61	–0.30	1.53	–4.5	4.4	2.22	1.38	0.18
ANB	–0.24	–1.20	0.71	–4	5.2	2.32	–0.53	0.60
MxP/MnP	1.91	–1.07	4.90	–10.5	13.1	7.23	1.32	0.20
LAFH%	–0.45	–0.85	–0.04	–2.4	2.2	0.98	–5.19	0.00
LAFH	–2.51	–3.51	–1.51	–8.7	1.4	2.42	–2.29	0.03
LPFH	–2.22	–4.54	0.09	–12.4	7.3	5.60	–1.99	0.06
Wits	–5.89	–6.97	–4.81	–11.3	–0.8	2.62	–11.25	0.00
OJ	–0.28	–0.72	0.15	–3.5	1.2	1.05	–1.36	0.19
OB	–0.34	–0.95	0.26	–3.9	2.2	1.47	–1.17	0.25
U1/MxP	1.05	–2.42	4.51	–17	23	8.39	0.62	0.54
L1/MnP	0.82	–2.11	3.75	–15.5	13	7.10	0.58	0.57

the mean differences in these variables (e.g. 1.5 degrees in ANB) as being clinically important given the small treatment changes involved. Interestingly, in both the present and this Quick Ceph Image Pro study, the most pronounced differences between the predicted and actual hard tissue changes occurred in the Wits measurements. This is unsurprising given that the Wits value involves determination of the functional occlusal plane, which in turn relies on the accurate identification of six landmarks. As such, whilst the Wits analysis[16] may be valid for diagnosis, clearly it is not reliable for orthognathic planning usage.

Assessment of TIOPS (Total Interactive Orthodontic Planning System) software in several consecutive studies also led to the observation that there were no statistically significant mean differences between the predictions and post-operative results.[8,9,10] Initially, the authors concluded that the surgical results were 'acceptably predictable' and, in common with other studies,[6,7,12] they failed to discuss the large variability evident in their results.[8] Subsequently, when large standard deviations were noted in a study of 40 consecutive cases it was opined that the errors occurred during the model surgery and surgical stages rather than the prediction process *per se*.[9]

An analysis of Dentofacial Planner pre-surgical predictions in mandibular setback cases revealed systematic errors in the OJ, mandibular angle and LAFH, but the distribution of the predicted versus actual differences was not discussed.[3] The relatively large standard variations evident with both Dentofacial Planner and Quick Ceph was noted in a study of 28 cases, but this investigated soft tissues only.[1] The authors concluded that there was 'generally good accuracy in the overall sample but with marked variability in the prediction consistency'. This comparative study of Dentofacial Planner and Quick Ceph also found evidence of wide variations in the accuracy of specific (soft tissue) landmark predictions.[1] Interestingly, these errors appeared to be related to the reproducibility/prediction of specific landmarks rather than any discernible difference between the two software packages. These wide sample variations are similar to that observed here with OPAL predictions, especially where MxP/MnP, LPFH, Wits, L1/MnP, and U1/Man exhibited large standard deviations and confidence intervals (Table 2). These wide individual deviations tended to be camouflaged by a regression to the mean. As such, it should not be assumed that individual patients will follow the mean pattern of small differences between their predicted and actual values, especially for vertical skeletal and incisor inclination measurements. In reality, it is very difficult to determine how much of the discrepancy between a predicted and actual result is due to inaccuracies in one or more of the following preparatory and surgical stages: landmark identification, digitizing, cephalometric software manipulation, transfer of movements from the prediction to articulated models, model surgery, surgical technique, and early settling/relapse. It is also worth remembering that predictions are two-dimensional representations of 3D features.

Aside from significant individual variability, systematic prediction error was seen in several measurements: MxP/MnP, LAFH and LPFH, with the differences in the latter two being statistically significant. As the mandible was advanced it underwent a backward rotation (Table 1), but this increase in the MxP/MnP was

over-estimated by the pre-treatment OPAL prediction (Table 2). This is consistent with the subjective observation of an exaggerated mandibular plane angle on individual prediction tracings and may be associated with closure of molar open bites by the OPAL software. Conversely, when TIOPS was tested in nine mandibular advancement cases an under-prediction in the MxP/MnP increase was observed.[10] A Dentofacial Planner study of 18 mandibular setbacks also revealed a mean under-prediction of the MxP/MnP increase by at least 3 degrees, although the source of this discrepancy was not evaluated.[3] The relatively greater inaccuracy of vertical change has also been noted in soft tissue analyses.[19] It is likely that these discrepancies in mandibular height and inclination may be due to inaccuracies in simulation of the mandibular rotational effects and angle remodelling, especially given the alteration in gonion landmark identification.

Finally, it would be interesting to assess the correlation between the amount of orthodontic/skeletal changes and prediction accuracy, but the sample size here was insufficient for robust statistical analysis. Previously, Aharon *et al.*[1] found 'a linear relationship between the prediction error and the surgical change' for the majority of soft tissues landmarks, i.e. the greater the surgical movement the greater the prediction error. Indeed, in some cases the error exceeded the actual change observed. Unfortunately, their sample size of 28 cases prevented any further analysis of this relationship. Similarly, analysis of soft tissue changes in 25 cases of mandibular advancement showed a correlation between the magnitude of movement and error.[4]

Conclusions

1. On average, pre-treatment OPAL predictions were accurate for many values, especially SNA, ANB, LAFH%, OJ and OB. However, wide individual variations occurred between the actual and predicted changes for many variables, and the Wits measurement was clearly inaccurate.
2. Systematic error occurred in the prediction of MxP/MnP, LPFH and LAFH changes, such that the prediction over-estimated the amount of surgically induced backward mandibular rotation.

Acknowledgements

The authors wish to thank their surgical and orthodontic colleagues who were involved in the care of these patients: Miss Pat Jenkins, and Messrs Peter Cove, Julian Crabb and David Roberts. We are also grateful to Jane Baxter, from the York Clinical Effectiveness Unit, for her help with the data processing.

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. Eales E, Newton C, Jones ML, Sugar A. The accuracy of computerised prediction of the soft tissue profile: a study of 25 patients treated by means of the Le Fort I osteotomy. *Int J Adult Orthod Orthognath Surg* 1994; **9**: 141–52.
3. Kolokitha OG, Athanasiou AE, Tuncay OC. Validity of computerized predictions of dentoskeletal and soft tissue profile changes after mandibular setback and maxillary impaction osteotomies. *Int J Adult Orthod Orthognath Surg* 1996; **11**: 137–54.
4. Schlutes G, Gaggl A, Karcher H. Accuracy of cephalometric and video imaging program Dentofacial Planner Plus in orthognathic surgical planning. *Comput Aided Surg* 1998; **3**: 108–14.
5. Gerbo LR, Poulton DR, Covell DA, Russell CA. A comparison of a computer-based orthognathic surgery prediction system to postsurgical results. *Int J Adult Orthod Orthognath Surg* 1997; **12**: 55–63.
6. Hing, NR. The accuracy of computer generated prediction tracings. *Int J Oral Maxillofac Surg* 1989; **18**: 148–51.
7. Loh S, Heng JK, Ward-Booth P, Winchester L, McDonald F. A radiographic analysis of computer prediction in conjunction with orthognathic surgery. *Int J Oral Maxillofac Surg* 2001; **30**: 259–63.
8. Donatsky O, Hillerup S, Bjorn-Jorgensen J, Jacobsen PU. Computerized cephalometric orthognathic surgical simulation, prediction and postoperative evaluation of precision. *Int J Oral Maxillofac Surg* 1992; **21**: 199–203.
9. Donatsky O, Bjorn-Jorgensen J, Holmqvist-Larsen M, Hillerup S. Computerized cephalometric evaluation of orthognathic surgical precision and stability in relation to maxillary superior repositioning combined with mandibular advancement or setback. *J Oral Maxillofac Surg* 1997; **55**: 1071–9.
10. Hillerup S, Bjorn-Jorgensen J, Donatsky O, Jacobsen PU. Precision of orthognathic surgery. *Int J Oral Maxillofac Surg* 1994; **23**: 255–61.
11. Harradine NWT, Birnie DJ. Computerized prediction of the results of orthognathic surgery. *J Maxillofac Surg* 1985; **13**: 245–9.
12. Eales EA, Jones ML, Newton C, Sugar AW. A study of the accuracy of predicted soft tissue changes produced by a computer software package (COG 3.4) in a series of patients treated by the Le Fort I osteotomy. *Br J Oral Maxillofac Surg* 1995; **33**: 363–9.

13. Bryan DC, Hunt NP. Surgical accuracy in orthognathic surgery. *Br J Oral Maxillofac Surg* 1993; **31**: 343–50.
14. Houston WJB. The analysis of errors in orthodontic measurements. *Am J Orthod* 1983; **83**: 382–90.
15. Harris M, Reynolds IR. Cephalometric analysis. In: *Fundamentals of Orthognathic Surgery*. London: WB Saunders Company, 1991: 38–48.
16. Jacobson A. The 'Wits' appraisal of jaw disharmony. *Am J Orthod* 1975; **67**: 125–38.
17. Battagel JM. A comparative assessment of cephalometric errors. *Eur J Orthod* 1993; **15**: 305–14.
18. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: a meta analysis. *Am J Orthod Dentofac Orthop* 1997; **112**: 165–70.
19. Upton PM, Sadowsky PL, Sarver DM, Heaven TJ. Evaluation of video imaging prediction in combined maxillary and mandibular orthognathic surgery. *Am J Orthod Dentofac Orthop* 1997; **112**: 656–65.