# Variations in orthodontic treatment planning decisions of Class II patients between virtual 3-dimensional models and traditional plaster study models

**Joshua L. Whetten,[a] Philip C. Williamson,[b] Giseon Heo,[c] Connie Varnhagen,[d] and Paul W. Major[e]**
*Edmonton, Alberta, Canada*

**Introduction:** Study models provide invaluable information in treatment planning. Digital models have proved to be an effective measurement tool, but their use in treatment planning has not been studied. **Methods:** Ten sets of records of Class II malocclusion subjects (dental study models, lateral cephalograms/tracings, panoramic radiographs, intraoral and extraoral photographs) were used for treatment planning by 20 orthodontists on 2 separate occasions. Digital models were used to evaluate the patients at 1 session and plaster models were used at the other session. Treatment recommendations were scored and compared for agreement. Eleven orthodontists served as the control group, looking at the records on 2 occasions with plaster models for agreement. **Results:** Good agreement was noted for surgery ($P$ = 1.00, $\kappa$ = 0.549), extractions ($P$ = .360, $\kappa$ = 0.570), and auxiliary appliances ($P$ = 1.00, $\kappa$ = 0.539) for the digital/plaster group. Agreement in the plaster/plaster group for surgery ($P$ = 1.00, $\kappa$ = 0.671), extractions ($P$ = 1.00, $\kappa$ = 0.626), and auxiliary appliances ($P$ = .791, $\kappa$ = 0.672) was also good. Overall proportions of agreement ranged between 0.777 and 0.870 for digital/plaster and 0.818 and 0.873 for plaster/plaster. **Conclusions:** There was no statistical difference in intrarater treatment-planning agreement for Class II malocclusions based on the use of digital models in place of traditional plaster models. Digital orthodontic study models (e-models) are a valid alternative to traditional plaster study models in treatment planning for Class II malocclusion patients. (Am J Orthod Dentofacial Orthop 2006;130:485-91)

Computers have become a mainstay in almost every workplace. In the 1990s, digital radiographs, photographs, and electronic charts were introduced in orthodontic practice. Recently, computer models have been developed. OrthoCAD (Cadent, Carlstadt, NJ, www.orthocad.com), e-models (Geodigm Corporation, Chanhassen, Minn, www.geodigmcorp.com), and Orthographics (Ortho Cast, High Bridge, NJ, www.orthocast.com) provide 3-dimensional computer images of a patient's teeth within a few days for customers who send an existing model or a dental alginate impression. The image is then downloaded and viewed by using free software available from the supplier's website.

The reasons for incorporating or ignoring digital technology in orthodontic practice are varied. Many orthodontists are comfortable using the traditional records they have relied on for years. Others integrate plaster models with digital photographs and radiographs in their offices. In doing so, these orthodontists are already using software that uses digital photographs and radiographs to plan treatments, write referral letters, and predict growth, leaving models as the last element missing from a fully digital treatment record.

The literature includes several articles dealing with the dimensional precision of digital model fabrication and laser surface scanning.[1-4] Other researchers considered the accuracy of digital and plaster models in measurements.[5-8] The digital model format has not yet been used in a clinical trial to test the reliability of decision making.

Orthodontists' reliability in patient classification and treatment-planning decisions using traditional tools has been previously reported. Lee et al[9] evaluated intrarater

From the University of Alberta, Edmonton, Alberta, Canada.
[a]Orthodontics resident, Department of Dentistry, Faculty of Medicine and Dentistry.
[b]Clinical assistant professor, Department of Dentistry, Faculty of Medicine and Dentistry.
[c]Assistant professor, Department of Dentistry, Faculty of Medicine and Dentistry.
[d]Professor, Department of Psychology, Faculty of Science.
[e]Professor, Director of Orthodontic Graduate Program, Department of Dentistry, Faculty of Medicine and Dentistry.
Supported by the University of Alberta Fund for Dentistry Grant # 2003-05; GeoDigm Corporation provided the digital e-models.
Reprint requests to: Dr Paul W. Major, Faculty of Medicine and Dentistry, Room 4051b, Dentistry/Pharmacy Centre, University of Alberta, Edmonton, Alberta, Canada T6G 2N8; e-mail, major@ualberta.ca.
Submitted, November 2004; revised and accepted, February 2005.
0889-5406/$32.00

doi:10.1016/j.ajodo.2005.02.022

and interrater reliability of treatment-planning decisions, specifically noting treatment need, extractions, functional appliances, and surgical need. They reported that orthodontists' reliability was generally substantial[10] when compared with themselves but only fair when compared with each other. Other researchers examined extraction decisions between orthodontists, but their results were conflicting, with 1 study showing substantial agreement and the other fair.[11,12] This inconsistency is possibly due to how orthodontists define commonplace terms. It has been suggested that, unless all orthodontists are trained in the same verbiage, this will continue to be a problem.[13]

Studies of malocclusion,[14,15] study-cast mounting,[16] and determination of treatment need[10] showed varying degrees of concurrence. When orthodontists were scored against each other, they had considerably lower levels of agreement than when scored against themselves. Study models were thought to be the most important component in treatment decision making in some studies.[17,18]

The effectiveness of digital decision making has not been evaluated. Consistency in orthodontic decision making between traditional plaster models and the virtual format should be examined. If there is consistency between the 2 (especially when compared with similar trials with controls), then the technology of digital models could reliably replace the traditional counterpart.

Our primary objective was to determine whether there is a difference in intrarater agreement measurements for surgery, extractions, and auxiliary appliances based on the study model format.

## MATERIAL AND METHODS

The Human Research Ethics Board at the University of Alberta granted approval of this study.

Pretreatment records for all patients in active orthodontic treatment at the University of Alberta's Orthodontic Graduate Clinic during the spring of 2003 were examined. A total of 107 patients met the initial inclusion criteria of at least an end-to-end molar relationship on 1 side. The following additional selection criteria were then applied: (1) ANB angle between 4° and 9°, (2) positive overjet of at least 4 mm, (3) at least 13 years of age at the time of records, and (4) at least a half step Class II molar relation on 1 side. Twenty-four patients fit these criteria, and 15 were randomly chosen.

A focus group of 3 orthodontic instructors at the University of Alberta evaluated the 15 sets of pretreatment records. They ranked the patients according to treatment difficulty and selected 10 to make up the patient record pool. Two sets of records from the extreme ends were described as "almost surely surgery" and "almost surely not surgery." Six more patients were called "truly borderline" cases. The pretreatment records, including study models, extraoral photographs, panoramic radiographs, lateral cephalograms (and tracings), were duplicated and constituted the patient information given to the orthodontist sample.

To obtain the digital model, a duplicated set of plaster models and a wax bite wafer for each patient were sent to Geodigm Corporation for registration and model fabrication. The duplicate model was sent for scanning to digital format so that the plaster and digital models were more closely comparable (both duplicates of the original). Additionally, another patient (Class I malocclusion) was selected as a test case to familiarize the practitioners with the model technology on the day of treatment planning.

The examiners and the records were each assigned a code number for blinding, to maintain doctor anonymity, and to protect patient confidentiality. Both the orthodontists and the patients consented to the use of their information in this study.

Lists of practicing orthodontists were obtained from the Alberta Society of Orthodontists and the Nevada State Board of Dental Examiners and randomized, and 26 orthodontists were contacted to participate. For the experimental (digital/plaster) group, our goal was 10 orthodontists from Las Vegas and 10 from Edmonton, so 13 from each city were contacted to allow for dropout. If a selected orthodontist practiced with a partner, then all practitioners in that office were asked to participate. Twenty-three of the 26 contacted orthodontists agreed to participate in the study, and 20 of the 23 who agreed to participate completed both study sessions.

To be included in the experimental group, the orthodontists could not have used digital models in treatment planning previously. Each orthodontist was questioned about familiarity with digital models. Those who had used digital models were considered for inclusion in the control group.

For the control (plaster/plaster) group, letters were sent to 13 orthodontists in Edmonton and Calgary, Alberta, Canada. Eleven orthodontists responded positively, and all planned the treatment of the subjects twice. This group was the control group for the digital model tool.

### Decision flow chart

A decision tree (Fig 1), adopted from Han et al,[17] was modified to fit our subjects. This tree was given to both the experimental and control groups. "No treatment" was not an option on the chart because we were looking for treatment consistencies, and, for the surgical patients, fixed banding was assumed. Three major

- Treatment Options
  - Surgery
    - Maxilla Only
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
    - Mandible Only
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
    - Maxilla + Mandible
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
    - Other (i.e., SARPE)
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
  - Non-Surgical
    - Fixed Banding
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
    - Functional Appliances (Fränkel, Herbst, Bionator, Other) or Facebow or Headgear _____________(please specify appliance)
      - Extract _______________(indicate teeth here or below)
      - Do Not Extract
    - Other ( RPE etc.) _________________(please specify appliance)

| **Extractions:** | Right | 8 7 6 5 4 3 2 1 | 1 2 3 4 5 6 7 8 | Left |
|---|---|---|---|---|
| | | 8 7 6 5 4 3 2 1 | 1 2 3 4 5 6 7 8 | |

**Fig 1.** Decision flow chart for treatment planning.

categories were targeted in this study for agreement: (1) surgery or no surgery, (2) extraction or nonextraction, and (3) auxiliary appliance or no auxiliary appliance (rapid palatal expansion, headgear/facebow, Fränkel, Herbst, Bionator, other).

## Data collection

For the experimental group, 2 treatment-planning sessions were scheduled, with at least a 1 month between sessions. The orthodontists were given either the digital model or the plaster model version of the records at the first meeting, with the other format at the second session. For the digital model sessions, the models were brought to a location chosen by each orthodontist and shown on a laptop. A trial model was used to familiarize the orthodontists with the software and model manipulation. The principal investigator (J.L.W.) was present to answer technical questions about software use. No auxiliary help was needed for the plaster session. At each session, the orthodontists marked the treatment that they would recommend to the patient, assuming no precluding factors to treatment. Each orthodontist was asked to arrive at a final treatment plan based on what he or she thought was the treatment of choice in his or her office, with patient options not affected by financial restraints. There was no "correct" treatment plan, because each orthodontist's treatment plans were tested only against themselves. There were no time limits, thus eliminating each orthodontist's variance in methods of deriving a treatment plan.

For the control group, 2 treatment-planning sessions were also scheduled at least a month apart. At both sessions, the plaster record format was used, with the subjects presented in random order. No one provided technical support, because those records were considered standard. Each orthodontist's selected treatment was based on the same assumptions as previously described, and the scoring was handled in the same manner.

### Statistical analysis

All measurements were recorded in a Microsoft Excel 2000 spreadsheet (Microsoft, Redmond, Wash) and analyzed with SPSS version 11.5 (SPSS, Chicago, Ill). The traditional set of plaster models, in conjunction with the other materials, was considered the gold standard by which diagnostic and treatment decisions made with the digital model should be compared. The McNemar test was used to generate a *P* value for each decision grouping. This test provided a method of evaluating where discrepancies arose between the first and the second treatment plans; a nonsignificant test result indicated agreement between the 2 decisions, and a significant result indicated lack of agreement. The data were examined by both group and case in this manner. The total number of observations in this group was 400 (200 digital and 200 plaster), leading to 200 comparisons of the new tool versus the standard. Data for the control group were analyzed similarly. The target number of observations was 200 for this group, and a total of 220 observations were obtained (110 plaster at each session). Thus, 110 treatment-planning session comparisons were made. Previous articles did not provide suitable data for sample size prediction. To evaluate for a possible type II statistical error, a post-hoc power analysis for McNemar *P* values was performed.[19]

A simple kappa statistic was also generated for surgery, extractions, and auxiliary appliance need. A proportion of agreement was calculated for each main treatment decision (surgery, extraction, auxiliary appliance) as a whole and by case.

## RESULTS

Changes in treatment recommendations based on the plaster and digital models are shown in Figure 2, and statistical measures are given in Table I. There was overall agreement for surgery/no surgery of 0.775, with matching decisions made 155 of 200 times. When discrepancies arose, 22 times the digital model session gave a positive response for surgery. The other 23 positives for surgery were found with the plaster models. There was as close to an exact split in the discrepancies as possible (22 of 45 vs 23 of 45). Neither the digital model nor the plaster model skewed the orthodontists to make treatment decisions regarding surgery ($P = 1.00$).

There was overall agreement for extraction decisions of 0.785, with the same outcome 157 of 200 times. Differences were seen in 43 instances. Positives for extractions with the digital model occurred only 18 times, whereas 25 positives for plaster only were noted. This shows that the orthodontists were slightly more likely to suggest extractions with the plaster than the digital models (25 of 43 vs 18 of 43). However, this was not a statistically significant discrepancy ($P = .36$).

There was overall agreement for auxiliary-appliance need of 0.870, and total agreement was seen 174 of 200 times. There was an exact split between the digital and plaster positives. Thirteen positives were reported with digital only and 13 positives with plaster only. Orthodontists were not influenced to recommend auxiliary appliances more by either model format ($P = 1.00$).

The kappa statistics for the 3 scored decisions ranged from 0.539 to 0.570 (Table I) and are provided for comparison with the control group's outcomes. According to the interpretation guidelines for kappa scores by Richmond et al[10] (Table II), the experimental group's kappa scores indicate moderate agreement.

There were 104 total surgeries recommended for the digital model format and 105 surgeries for the plaster. In each case, maxilla only was suggested 4 times; mandible only 88 (plaster) and 84 (e-models) times; a combination of maxilla and mandible 13 (plaster) and 15 (e-models) times; other surgeries were recommended 0 (plaster) and 1 (e-models) time.

When extractions were indicated with e-models (98 times), there was a mean of 2.77 teeth. There was a mean of 2.70 teeth when extractions were recom-

**Surgery Recommendations, Experimental Group**

| | | Plaster | |
|---|---|---|---|
| | | No | Yes |
| emodels | No | 73 | 23 |
| | Yes | 22 | 82 |

**Extraction Recommendations, Experimental Group**

| | | Plaster | |
|---|---|---|---|
| | | No | Yes |
| emodels | No | 75 | 25 |
| | Yes | 18 | 82 |

**Auxiliary Recommendations, Experimental Group**

| | | Plaster | |
|---|---|---|---|
| | | No | Yes |
| emodels | No | 153 | 13 |
| | Yes | 13 | 21 |

**Surgery Recommendations, Control Group**

| | | Plaster 2 | |
|---|---|---|---|
| | | No | Yes |
| Plaster | No | 42 | 9 |
| 1 | Yes | 9 | 50 |

**Extraction Recommendations, Control Group**

| | | Plaster 2 | |
|---|---|---|---|
| | | No | Yes |
| Plaster | No | 36 | 10 |
| 1 | Yes | 10 | 54 |

**Auxiliary Recommendations, Control Group**

| | | Plaster 2 | |
|---|---|---|---|
| | | No | Yes |
| Plaster | No | 74 | 6 |
| 1 | Yes | 8 | 22 |

**Fig 2.** Contingency tables for treatment recommendation agreement (based on session/group).

mended with the plaster models (106 times). Frequencies of other treatment recommendations are given in Table III.

The 3 treatment decision groups all reported nonsignificant *P* values with the McNemar test. Surgery and extraction had the highest value (1.000), whereas auxiliary appliance listed a *P* of .791. The overall proportions of agreement were 0.836 for surgery, 0.818 for extractions, and 0.873 for auxiliary-appliance need (Table I).

The kappa statistic values for the orthodontists as a group ranged from 0.626 to 0.672 (Table I). These values were substantial (Table II).[10]

The control group also showed consistency in the frequency and type of surgery recommended. At the first plaster session, there were 59 total surgeries: 1 maxilla only, 48 mandible only, 10 combined maxilla and mandible, and 0 other. At the second plaster session, there were 58 total surgeries: 0 maxilla only, 49 mandible only, 8 combined maxilla and mandible, and 1 other surgery.

## DISCUSSION

The digital model tool did not significantly affect treatment-planning decisions in Class II malocclusions. When there was a discrepancy whether to recommend surgery, the discrepancies were split almost evenly (22 and 23) as to model format used to make that decision. The type of surgery planned did not significantly vary according to model type. The same can be said for the need for auxiliary appliances. There was an exact split between the model formats when treatment modality changed (13 and 13). The area of greatest deviation was for extraction decisions; the distribution was greater (18 and 25). Use of the plaster models tended to result in recommended extractions, whereas the digital models did not. However, this was not a statistically significant difference. Model format did not alter the number of extractions when the orthodontist suggested extraction treatment.

Although the kappa statistic measures the level of agreement between 2 models (digital and plaster), the McNemar test is a nonparametric method that uses matched pair labels (A,B). It is used when the data consists of paired observations or outcomes A and B; the hypothesis tested whether AB pairs are as likely as BA pairs (Table IV).[20]

An example of how these data can be interpreted is shown by using the experimental group's surgical decisions from Figure 2. There were 200 total comparisons between the 200 digital model recommendations and the 200 suggested treatments based on the plaster models. There was complete agreement (yes/yes, no/no) 155 times, and, in 45 instances, there were conflicting treatment plans. The proportion (percentage) of agreement is then .775 (155 of 200). The 45 disagreements of treatment recommendations between the 2 sessions ($\kappa$ = 0.549) cannot be attributed to the change in model format. The *P* value reported is nonsignificant because there was almost complete agreement in where

**Table I.** Tool's effect on agreement

| | Digital vs plaster | | | Plaster vs plaster | | |
|---|---|---|---|---|---|---|
| *Procedure* | *McNemar P value (power)* | *Kappa* | *Proportion of agreement* | *McNemar P value (power)* | *Kappa* | *Proportion of agreement* |
| Surgery | 1.000 (.965) | 0.549 | 0.775 | 1.000 (.975) | 0.671 | 0.836 |
| Extractions | .36 (.815) | 0.570 | 0.785 | 1.000 (.975) | 0.626 | 0.818 |
| Auxiliary appliances | 1.000 (.975) | 0.539 | 0.870 | .791 (.923) | 0.672 | 0.873 |

**Table II.** Guidelines for kappa statistic interpretation (Richmond et al[10])

| *Kappa statistic* | *Strength of agreement* |
|---|---|
| <0 | Poor |
| 0-0.2 | Slight |
| 0.21-0.4 | Fair |
| 0.41-0.6 | Moderate |
| 0.61-0.8 | Substantial |
| 0.81-1.00 | Almost perfect |

the discrepancy arose (22 and 23). Based on this information, the fact that the orthodontists were slightly more consistent in the control group cannot be because the tool remained the same for them.

In relation to previous studies that examined rater reliability in orthodontics, the kappa values in this study were consistent with those reported elsewhere. Lee et al[9] described ranges in intrarater kappa values for 10 orthodontists' evaluations of 60 vignettes based on 4 distinct categories: treatment needed (0.24-0.90), orthognathic surgery required (−0.02-1.00), extractions required (0.51-0.80), and functional appliance treatment (0.14-0.81). Ribarevski et al[12] studied the consistency of orthodontic extraction decisions and reported intrarater kappa values ranging from 0.54 to 0.96, with a percentage/proportion of agreement of 80% to 98% in these same cases. More recently, Mandall[20] aimed to evaluate 8 orthodontists' reliability for accepting referrals based on clinical photographs. Scores ranged from 0.34 to 0.90. In all studies mentioned, intrarater reliability was higher than interrater reliability.

According to Thomas,[21] postpubertal patients with Class II malocclusions often have only 2 options for correcting the problem: mandibular surgery and camouflage. He also added that there is little disagreement in considering the patients at the extreme ends of the scale (as seen with our focus group's decisions). Problems arise when the patient is somewhere in the middle and could benefit from either option. Thomas[21] estimated that 15% to 20% of the teenage population in the United States has an overjet of 6 mm or more. Proffit et al[22] put the number of Class II patients who could benefit from surgical intervention at about 10% of the general population.

When the subjects were evaluated individually, it was evident that the focus group did an excellent job in choosing the ones to be examined in this study. The focus group's "almost surely surgery" and "almost surely not surgery" subjects were planned for treatment similarly by orthodontist sample. The subjects with the lowest proportion of agreement for all major decisions were deemed borderline in the focus group's initial case selection.

Initially, the focus group selected patients 1, 2, 4, 5, 8, and 10 as borderline. Analysis showed that the lowest overall proportions of agreement were for patients 2, 4, 8, and 10; all were girls between the ages of 13 years 2 months and 16 years 1 month. All had ANB angles greater than 6°, with low mandibular planes (mean 24.1°). Crowding was mild (0-2 mm) in each arch, yet overjet was moderate (5-7 mm). All patients had discrepances between bilateral molar and canine relationships, with a Class II relationship on 1 side and a Class I or an end-on relationship on the other.

The actual choice of treatment is most likely a function of which orthodontist the patient happens to contact. Weaver[23] studied decisions in surgical cases by Canadian orthodontists. She noted that such factors as years of experience (practice), part-time vs full-time practice, private practice vs academia, and level of training (masters vs diploma) did not significantly affect decisions for the latest possible time to intervene with surgical treatment. Beyond the bias based on the orthodontist's personal characteristics (eg, risk aversion), there is little hard evidence to justify 1 treatment as superior to another. Orthodontic treatment requires subjective judgment based on patient values, practitioner values and experiences, and many other factors.

This study was designed to simulate the real-life clinical situation when study models are part of the normal diagnostic records. The objective was to evaluate whether digital models can be used in place of plaster models in the context of full diagnostic records, without altering treatment-planning decisions. In some instances, an orthodontist may have relied almost entirely on records other than study models in formulating a treatment plan. The accuracy of measurements derived from digital study models

**Table III.** Frequency of treatment recommendations for each observation

| | *Experimental* | | *Control* | |
|---|---|---|---|---|
| | *E-model* | *Plaster* | *Plaster 1* | *Plaster 2* |
| Total surgeries | 104 | 105 | 59 | 58 |
| Average surgeries per doctor | 5.2 | 5.25 | 5.36 | 5.27 |
| Total extractions | 271 | 286 | 198 | 195 |
| Average extractions per doctor | 13.55 | 14.3 | 18.0 | 17.73 |
| Total auxiliary appliances | 34 | 34 | 30 | 28 |
| Average auxiliary appliances per doctor | 1.7 | 1.7 | 2.72 | 2.54 |

**Table IV.** McNemar test results

| | *No* | *Yes* | *Total* |
|---|---|---|---|
| No | AA (a) | AB (b) | AA + AB ($p_1$) |
| Yes | BA (c) | BB (d) | BA + BB ($q_1$) |
| Total | AA + BA ($p_2$) | AB + BB ($q_2$) | AA + AB + BA + BB |

McNemar calculation is based on following formula:

$$df = 1: \chi^2 = \frac{(|AB - BA| - 1)^2}{AB + BA} b + c.$$

has been previously reported and found to be suitable for clinical use.[5-8] It would be useful to investigate the reliability of extraction decisions with digital vs plaster models. This would involve a nonclinical situation in which treatment decisions were made without the benefit of photographs and radiographs.

## CONCLUSIONS

Digital orthodontic study models (e-models) are a valid alternative to traditional plaster study models in treatment planning for Class II malocclusions.

## REFERENCES


1. Kuroda T, Motohashi N, Tominaga R, Iwata K. Three-dimensional dental cast analyzing system using laser scanning. Am J Orthod Dentofacial Orthop 1996;110:365-9.
2. Kusnoto B, Evans CA. Reliability of a 3D surface laser scanner for orthodontic applications. Am J Orthod Dentofacial Orthop 2002;122:342-8.
3. Sohmura T, Kohjima T, Wakabayashi K, Takahashi J. Use of an ultrahigh-speed laser scanner for constructing three-dimesional shapes of dentition and occlusion. J Prosthet Dent 2000;84:345-52.
4. Motohashi N, Kuroda T. A 3-D computer-aided design system applied to diagnosis and treatment planning in orthodontics and orthognathic surgery. Eur J Orthod 1999;21:263-74.
5. Santoro M, Galkin S, Teredesai M, Nicolay O, Cangialosi T. Comparison of measurements made on digital and plaster models. Am J Orthod Dentofacial Orthop 2003;124:101-5.
6. Tomassetti JJ, Taloumis LJ, Denny JM, Fischer JR Jr. A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. Angle Orthod 2001;71:351-7.
7. Zilberman O, Huggare J, Parikakis K. Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. Angle Orthod 2003;73:301-6.
8. Stevens D. Clinical use of virtual study models versus traditional plaster study models for orthodontic treatment diagnosis [thesis]. Edmonton, Alberta, Canada: University of Alberta; 2004. p. 79.
9. Lee R, MacFarlane T, O'Brien K. Consistency of orthodontic treatment planning decisions. Clin Orthod Res 1999;2:79-84.
10. Richmond S, O'Brien KD, Roberts CT, Andrews M. Dentists variation in the determination of orthodontic treatment need. Br J Orthod 1994;21:65-8.
11. Baumrind S, Korn EL, Boyd RL, Maxwell R. The decision to extract: part 1—interclinician agreement. Am J Orthod Dentofacial Orthop 1996;109:297-309.
12. Ribarevski R, Vig P, Vig KD, Weyant R, O'Brien K. Consistency of orthodontic extraction decisions. Eur J Orthod 1996;18:77-80.
13. Pair J, Luke L, White S, Atchinson K, Englehart R, Brennan R. Variability of study cast assessment among orthodontists. Am J Orthod Dentofacial Orthop 2001;120:629-38.
14. Firestone A, Beck F, Beglin F, Vig K. Validity of the index of complexity, outcome, and need (ICON) in determining orthodontic treatment need. Angle Orthod 2002;72:15-20.
15. Keeling SD, McGorray S, Wheeler TT, King GJ. Imprecision in orthodontic diagnosis: reliability of clinical measures of malocclusion. Angle Orthod 1996;66:381-91.
16. Ellis PE, Benson PE. Does articulating study casts make a difference to treatment planning? J Orthod 2003;30:45-9.
17. Han UK, Vig KW, Weintraub JA, Vig PS, Kowalski CJ. Consistency of orthodontic treatment decisions relative to diagnostic records. Am J Orthod Dentofacial Orthop 1991;100:212-9.
18. Greenhill R, Basford K. The effects of varying diagnostic records on subjective orthodontic treatment priority assessments. Aust Orthod J 1979;6:55-63.
19. Zar J. Multisample hypothesis: paired-sample hypotheses. In: Zar J, editor. Biostatistical analysis. 3rd ed. Toronto: Prentice-Hall; 1996. p. 173.
20. Mandall NA. Are photographic records reliable for orthodontic screening? J Orthod 2002;29:125-7.
21. Thomas PM. Orthodontic camouflage versus orthognathic surgery in the treatment of mandibular deficiency. J Oral Maxillofac Surg 1995;53:579-87.
22. Proffit WR, Phillips C, Dann CT. Who seeks surgical-orthodontic treatment? Int J Adult Orthod Orthognath Surg 1990;5:153-60.
23. Weaver N. Age limitation on provision of orthopedic therapy and orthognathic surgery. Am J Orthod Dentofacial Orthop 1998;113:156-64.