Consistency of orthodontic treatment decisions relative to diagnostic records

Unae Kim Han, DMD, MPH, MS,[a] Katherine W. L. Vig, BDS, D. Orth., FDSRCS(Eng.), MS,[b] Jane A. Weintraub, DDS, MPH,[c] Peter S. Vig, BDS, PhD, D. Orth., FDSRCS(Eng.),[d] and Charles J. Kowalski, PhD[e]

Farmington Hills and Ann Arbor, Mich., and Chapel Hill, N.C.

The purpose of this study was to evaluate how incremental information obtained from different types of diagnostic records contributes to the determination of orthodontic treatment decisions. Pretreatment records of 57 orthodontic patients were assessed by five orthodontists who were part-time faculty members and also in private practice. This sample consisted of dental school orthodontic patients who had Class II malocclusions and included patients at three different dental developmental stages. The following diagnostic records were used: study models (S), facial photographs (F), a panoramic radiograph (P), a lateral cephalogram (C), and its tracing (T). Five combinations of diagnostic records were presented to the orthodontists in the following sequence: (1) S; (2) S + F; (3) S + F + P; (4) S + F + P + C; and (5) S + F + P + C + T. The simultaneous interpretation of all diagnostic records (S + F + P + C + T) was used as the "diagnostic standard." There was a diagnostic standard for each of the patients and for each of the orthodontists. The diagnostic standard was achieved: (1) S = 54.9%, (2) S + F = 54.2%, (3) S + F + P = 60.9%, and (4) S + F + P + C = 59.9%. Thus, in a majority of cases (55%), study models alone provided adequate information for treatment planning, and incremental addition of information from other types of diagnostic records made small differences. (Am J Orthod Dentofac Orthop 1991;100:212-9.)

There has recently been a proliferation in the diagnostic and treatment regimens available in orthodontics. The battery of diagnostic tests include study models mounted on semiadjustable articulators, types of jaw tracking, TMJ tomograms, and other presumably necessary devices. Most of these modalities are suggested and used on the basis of clinical experience and personal preference. Their clinical efficacy has not been validated, as yet by the orthodontic specialty. Likewise the probabilities of the expected treatment outcomes have not been assessed.

In medicine the escalating cost of the nation's health care has resulted in cost containment and methods to assess the efficacy of treatment alternatives.[1,2] The field of clinical epidemiology or "clinimetrics" has been developed to evaluate the efficacy of treatments and diagnostic tests. Clinical epidemiology combines biostatistics, epidemiology, clinical decision analysis, risk-benefit analysis, cost-benefit analysis, and cost-effectiveness analysis. It is concerned with the study of groups of patients to provide evidence on which to base clinical decisions in health care.[3] Although these techniques have been applied in medicine[3-5] and in some areas of dentistry,[6-8] they have not been widely applied in orthodontics.[9]

Diagnosis in orthodontic practice includes the classification of malocclusion, and currently emphasis is being directed toward a comprehensive synthesis of information. Proffit and Ackerman[10] advocate the "problem-oriented" approach to orthodontic diagnosis that was originally developed in medicine to provide a rational approach to diagnosis.[11] The decision-making process in orthodontic diagnosis and treatment planning involves (1) the recognition of the characteristics of malocclusion and dentofacial deformity, (2) the definition of the nature or cause of the problem, and (3) the design of a treatment strategy based on the specific needs of the patient. Fundamental to this process is the acquisition of relevant information to form a data base.

This work was supported by USPHS, NIH/NIDR grants DE 06881 and DE 08714.

[a]In private practice, Farmington Hills, Mich.
[b]Associate Professor and Chairman, Department of Orthodontics, University of Pittsburgh School of Dental Medicine, Pittsburgh, Pa.
[c]Assistant Professor, Department of Dental Ecology, School of Dentistry, University of North Carolina, Chapel Hill.
[d]Associate Dean of Research, University of Pittsburgh School of Dental Medicine, Pittsburgh, Pa.
[e]Professor, Department of Biologic and Materials Sciences, School of Dentistry, The University of Michigan, Ann Arbor.

8/1/23631

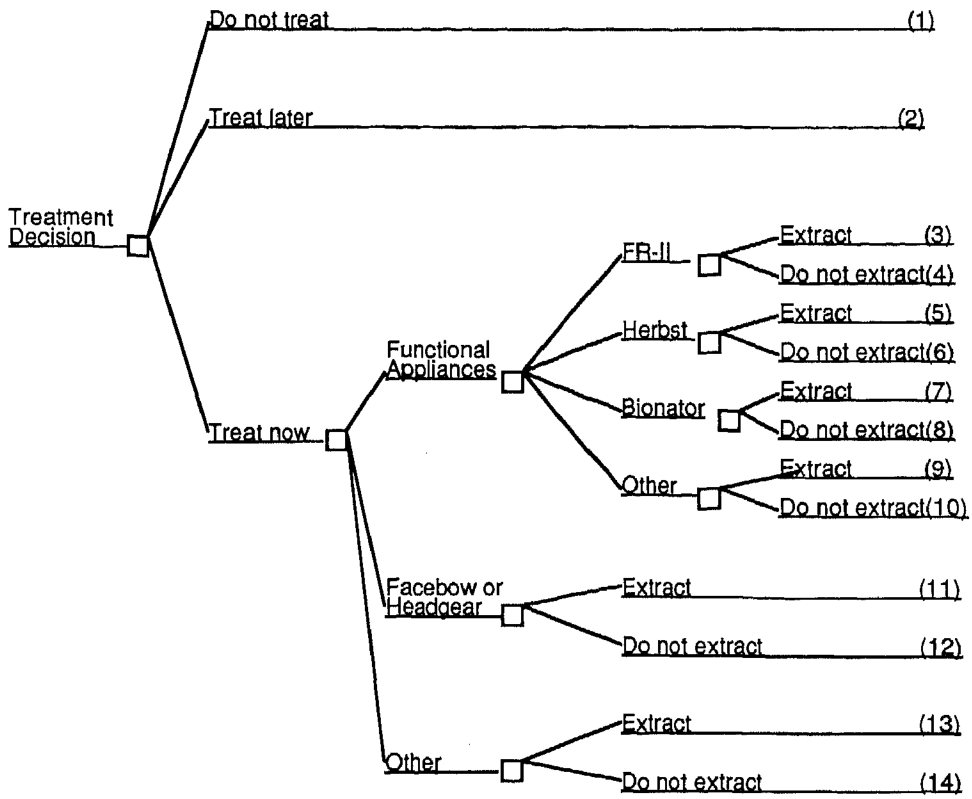

Fig. 1. Decision tree for patients with late mixed dentition.

The purpose of this study was to evaluate the relative usefulness of routine diagnostic records obtained by orthodontists.

The synthesis of information from diagnostic procedures yields degrees of diagnostic value. Diagnostic utility can be assessed and is based on the probability that either doing or not doing something will influence all or some of the following outcomes: (1) the diagnosis and its accuracy, (2) the treatment process, and (3) the outcome of treatment. The selection of tests should be rational and based on a comparison for sensitivity, specificity, and predictive values, both positive and negative. The problem is to determine the "gold standard" for evaluating such tests and thus for determining the prevalence of the disease or condition. The purpose of this study was to compare the diagnostic utility of different incremental combinations of diagnostic information to a "gold standard."

MATERIALS AND METHODS

Selection of cases

Pretreatment records of 57 orthodontic patients were selected from the graduate orthodontic clinic at the University of Michigan according to specified eligibility criteria. This study was limited to Class II, Division 1 malocclusion, and the cases were chosen by a stratified random selection process. The three strata were determined by the dentition status: (a) late mixed dentition ($n = 20$), (b) early permanent dentition ($n = 20$), and (c) adult dentition ($n = 17$). The pretreatment records, including study models, extraoral photographs, panoramic radiographs, and lateral cephalograms, were duplicated, and these records were used for the experimental diagnostic and treatment planning sessions.

Selection of orthodontists

Five orthodontists who practice in the state of Michigan and teach part-time in the graduate orthodontic clinic participated in the study. The orthodontists were selected on the basis of their availability and commitment to this study.

Description of decision trees

A decision tree was constructed for each of the three groups (Figs. 1-3). These trees were prepared to allow the orthodontists to follow a logical sequence of treatment decision making, which was initiated with a strategy choice, followed by a treatment decision, and finally an "extract or do not extract" decision.

Data collection

For each of the orthodontists, five diagnostic and treatment-planning sessions for every case were scheduled

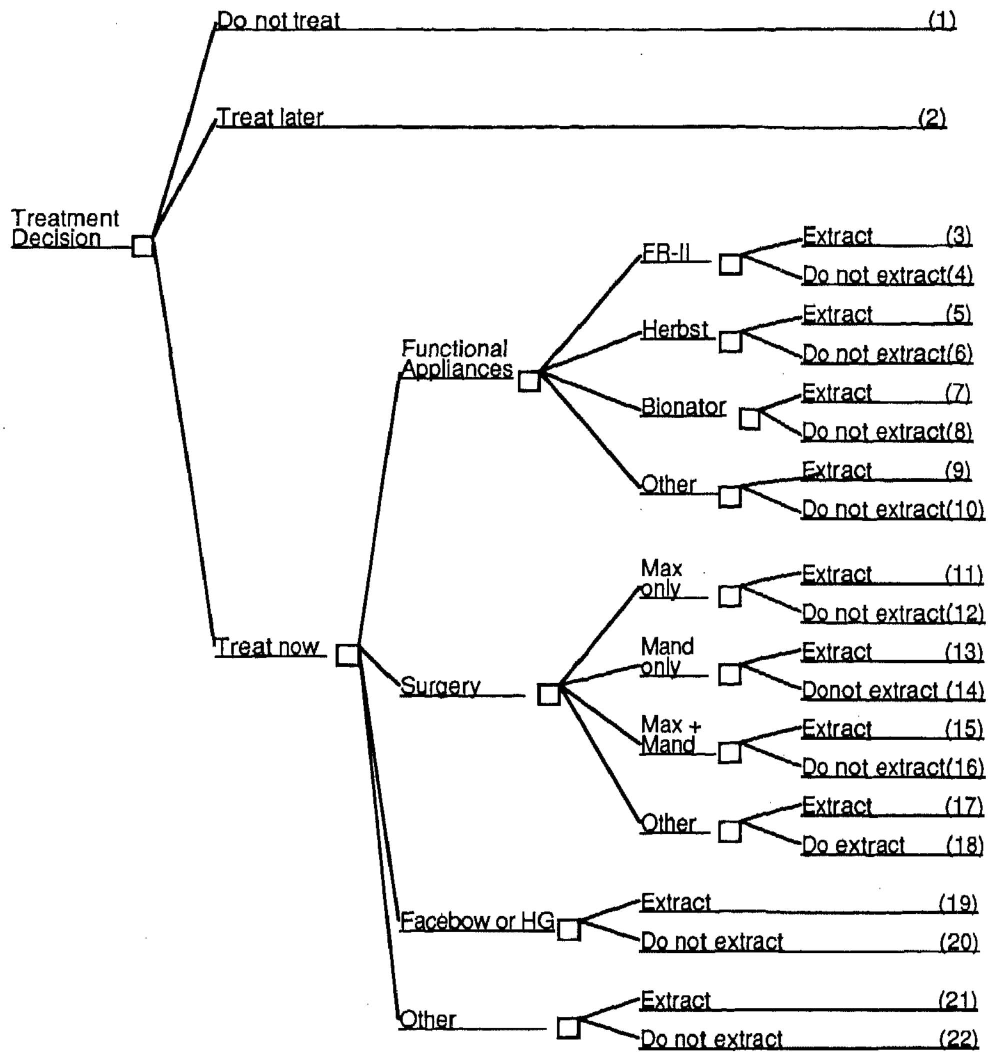

Fig. 2. Decision tree for patients with early permanent dentition.

and conducted according to a standardized data-gathering protocol with an approximately 1-month interval separating each session. Five combinations of diagnostic information, or records, were presented to the clinician at each session in the following order: (I) study models only (S); (II) S + facial photographs (F); (III) S + F + panoramic radiograph (P); (IV) S + F + P + lateral cephalogram (C); and (V) S + F + P + C + tracing (T). During each session the orthodontists were asked to select a decision pathway from a decision tree, which was constructed for each of the three patient categories (Figs. 1-3).

Analytical methods

The simultaneous interpretation of all diagnostic materials was used as the diagnostic standard. A diagnostic standard for each of the cases and for each of the orthodontists was therefore established. With the presentation of combinations of diagnostic records to the participating orthodontists, a proportion of the treatment plans equivalent to the diagnostic standard was calculated, and the changes in the treatment decision attributable to each additional incremental record were estimated. To assess the orthodontists' consistency in treatment decision making, given the same records over a period of time, session V was repeated for 15 of the cases. The proportions of agreement between the two sessions were then analyzed.

In the clinical decision analysis, "pruning" of a decision tree refers to elimination of those decision branches that would never be followed.[4] In this study the decision trees were modified by combining some of the decision branches. After the initial analysis of the proportions of agreement to the diagnostic standard, the decision trees were modified in the following three steps to reduce the number of treatment options:

Step 1. In group 1, the "headgear/facebow" and "other" options were combined and the choice of different

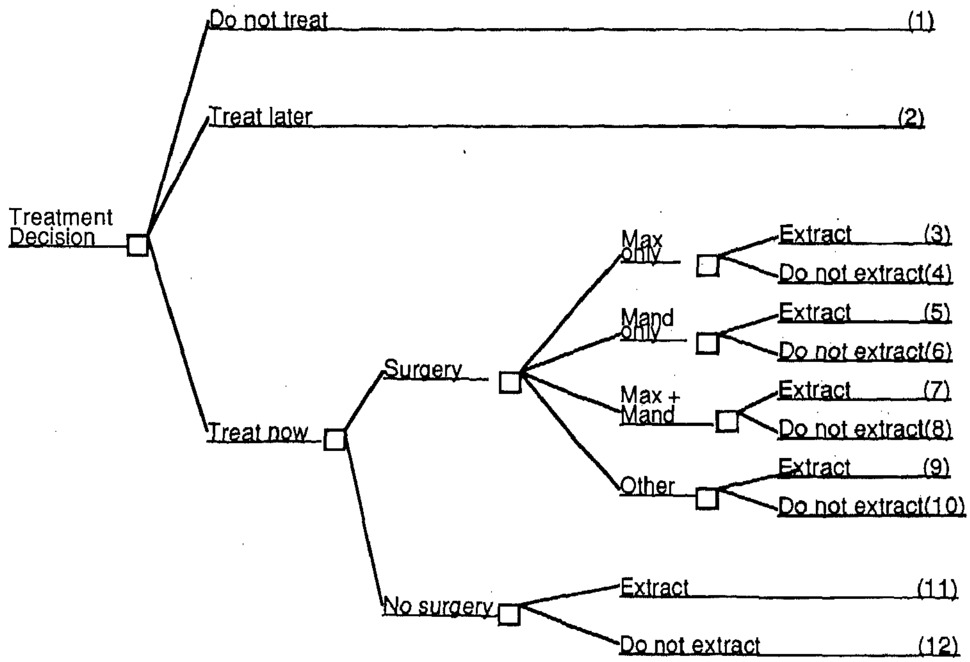

Fig. 3. Decision tree for patients with adult dentition.

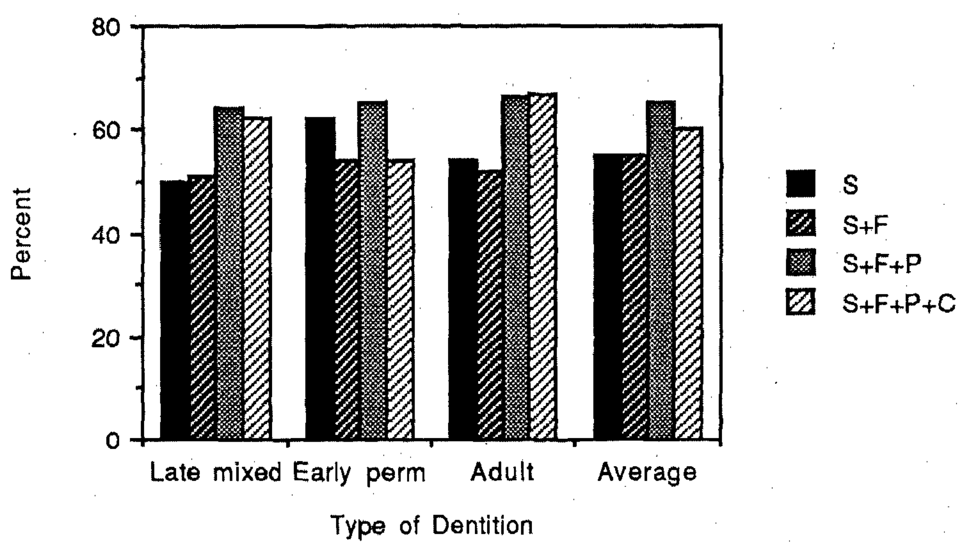

Fig. 4. Proportion of agreement with "diagnostic standard" for all five orthodontists by dentition status of patient and combination of diagnostic procedures.

types of functional appliances was eliminated. The number of treatment options decreased to six. In group 2, the "headgear/facebow" and "other" options were combined and the choice of different types of functional appliances and types of surgery was eliminated. There were eight treatment options. In group 3, elimination of the types of surgery reduced the number of treatment options to six.

Step 2. The "extract or do not extract" decision was eliminated. There were four treatment options for group 1, five for group 2, and four for group 3.

Step 3. The "functional appliance - no functional appliance" or "surgery - no surgery" decision was eliminated.

After step 3, the trees were left with the basic strategy decision—"do not treat," "treat later," or "treat now."

Table I. Proportion of agreements with tree modifications for all five orthodonists

Session	Modifications (%)			
	No modification	Step 1	Step 2	Step 3
Group 1				
I	50	57	66	80
II	51	64	74	81
III	64	69	77	82
IV	62	71	80	87
Group 2				
I	62	74	85	99
II	54	68	80	98
III	65	69	80	96
IV	54	63	75	99
Group 3				
I	54	55	75	76
II	51	71	86	95
III	66	73	85	95
IV	67	68	81	96
Average				
I	55	62	75	92
II	55	68	80	92
III	65	70	80	92
IV	60	67	79	94

Table IIA. Proportion of agreements between sessions

Group	Sessions (%)			
	I	II	III	IV
1	50	51	64	62
2	62	54	65	54
3	54	51	66	67
Average	55	55	65	60

Analysis of variance (ANOVA) was conducted to determine whether the variation in outcome was because of the differences in the combinations of diagnostic records employed (variable = sessions) or the dentition status of the patients (variable = group). The ANOVA was also conducted to estimate the differences in the means of the proportions of agreement between the modification of decision trees. The analysis of proportions of agreement was repeated with session IV (minus tracing) as the diagnostic standard.

RESULTS

Intraclinician reliability

On the basis of 15 repeated cases, the average proportions of consistent treatment plan decisions made over a period of 4 to 6 weeks ranged from 53% to 73%, with an average of 65% for the five clinicians. Reliability was higher for the adult-dentition group (76%) than for groups in the late mixed (56%) or early permanent (64%) dentition.

Table IIB. Analysis of variance

Source of variation	F	P
G	0.13	0.88
S	1.50	0.26
G × S	1.66	0.17

G = Group; S = session.

Table IIIA. Proportion of agreements between tree modifications

Group	Modifications (%)			
	No modification	Step 1	Step 2	Step 3
1	57	65	74	83
2	59	69	80	98
3	60	67	82	96

Table IIIB. Analysis of variance

Group	Source of variation	F	P
1	T	26.06	0.00
	S	1.42	0.29
	T × S	0.93	0.52
2	T	56.58	0.00
	S	0.72	0.56
	T × S	1.19	0.33
3	T	84.38	0.00
	S	2.43	0.12
	T × S	2.54	0.02

T = Modification; S = sessions; T × S = interaction.

Proportion of agreement

There was a marked variation among the orthodontists in the pattern of agreement with the diagnostic standard. Overall, in a majority of cases (55%), the study models alone yielded treatment strategies that were equivalent to the diagnostic standard. The rate at which each combination of records was sufficient to achieve the diagnostic standard was as follows: (1) S = 54.9%; (2) S + F = 54.2%; (3) S + F + P = 60.9%; and (4) S + F + P + C = 59.9% (Fig. 4).

Modification of trees

The proportion of consistent treatment decisions increased significantly with each step in the decision tree

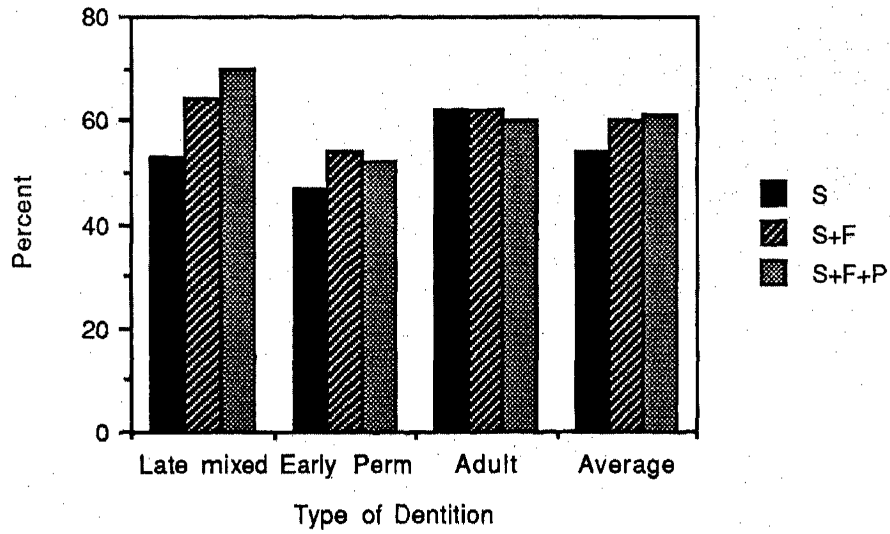

Fig. 5. Proportion of agreement with "diagnostic standard" for all five orthodontists with session IV as the "diagnostic standard."

modifications reducing the number of treatment options. Overall, this increased from 65% for the original tree to 71% for step 1, 79% for step 2, and 93% for step 3 modifications (Table I). With the tree modifications, the proportions of agreement, for each session, with the diagnostic standard also increased significantly. However, the pattern of agreement going from session I to session IV remained the same. In other words, the use of study models alone yielded treatment strategies that were equivalent to the diagnostic standard in a majority of cases, and additional information made small differences for all tree modifications (Table III).

Analysis of Variance

The differences in means (among five orthodontists) of their proportion of agreement between sessions I, II, III, and IV were not statistically significant, $p > 0.05$ (Table II). This finding indicates that the proportion of agreement with the diagnostic standard was not influenced by the number of diagnostic records, which varied in type. The differences in means of the proportion of agreements among the patients in groups 1, 2, and 3 were not statistically significant, $p > 0.05$ (Table II). This finding indicates that the pattern of decision making, based on the number of diagnostic records, was not different among the three patient categories. The proportion of agreement with the diagnostic standard was, however, significantly different among the tree modifications (Table III). This finding appears to be logical with each step of modification, reducing the number of treatment options, and increasing the proportion of agreement.

The use of session IV as the diagnostic standard

The proportions of agreement, when averaged over five orthodontists, were 54% for session I, 60% for session II, and 61% for session III when the groups were averaged (Fig. 5). This pattern of agreement when session IV was used as the diagnostic standard is similar to the pattern when session V was used. Study models alone yielded treatment decisions that were equivalent to the diagnostic standard and additional records made small differences.

DISCUSSION

The five participating orthodontists in this study were similar in terms of their dental school and orthodontic education. When asked "which diagnostic records do you usually obtain at your office before making a diagnosis and treatment plan?" all five reported that they use study models, facial and intraoral photographs, a panoramic radiograph, a lateral cephalogram, and its tracing. According to the 1986 *Journal of Clinical Orthodontics* practice survey,[12] 86.3% of the surveyed orthodontists obtained pretreatment panoramic radiographs, 97.3% obtained lateral cephalograms, and 88.2% obtained trimmed study models.

In clinical epidemiology, the assessment of the efficacy of a diagnostic test depends on the availability of a "gold standard." In addition, the quality of the assessment is only as good as the quality of the gold standard. Unfortunately, in orthodontics there is not a single "positivity criterion" to distinguish patients with "normal" skeletodental relationship and facial form from those considered to be "diseased" or "abnormal" for purposes of subsequent decision making.[13] The ultimate goal of the diagnostic process is to benefit the patient by producing an outcome that would not have occurred without diagnosis and subsequent intervention. The focus of this study was the treatment-planning decision rather than the diagnosis. The gold standard of the orthodontic diagnostic procedures was termed the *diagnostic standard* and was the treatment decision of a particular case and for a particular orthodontist when all the records (S,F,P,C,T,) were provided. This was assumed to represent the "best" that the clinician can do for each of the given patients. Because of the heterogeneous nature of orthodontic practice,[14] it did not seem reasonable to have one diagnostic standard for each case for all five orthodontists.

The order in which the records were added to the diagnostic battery from session I to session V—study models, facial photographs, panoramic radiograph, lateral cephalogram, and, finally, tracing—seemed appropriate. In clinical practice, the study models and facial photographs are comparable to an initial clinical examination of the patient. The tracing usually follows the lateral cephalogram, although some orthodontists do not trace the cephalogram. It was the consensus of our research team that a panoramic radiograph should follow the photographs, since it provides a general survey of the dentition.

It was surprising to find that study models alone provided adequate information for treatment planning in 55% of the cases. This finding was in agreement with the article by Naccache et al.,[15] who found study models to be very useful in their computer-assisted orthodontic diagnosis.

We anticipated finding an incremental increase in the proportions of agreement with the diagnostic standard with each addition of diagnostic records. This was expected to lead to a convenient analysis of incremental effectiveness of the records. The ANOVA indicated that the proportions of agreement with the diagnostic records were not influenced by the number of records, which was confirmed by the variation within each session that was greater than the variation among the sessions. Some of the sources of variance within each session might be attributed to the variation among orthodontists and over time.

The intraclinician reliability, which included the proportion of consistent treatment plan decisions made over a period of time, was lower than expected, with an average of 65%. The design of this study allowed for a "washout period" of approximately 1 month between sessions so that possible memory bias in treatment decisions would be excluded. This was supported by the relatively low proportion of consistent decisions by the clinicians. The variability in treatment decisions made over time should also be taken into consideration when one is interpreting the consistency of treatment decisions as a function of diagnostic records. Although the significant variation among the orthodontists in terms of specific treatment decisions and the variation over time were interesting aspects of orthodontic diagnostic process, they might also have interfered with the contribution of incremental diagnostic records to the overall treatment planning process.

The overall proportion of agreement of 60% for session IV indicates that tracings made a large difference in treatment decisions. This could also be a function of our study design, which used the tracing plus all the other records as the diagnostic standard. When the session IV (minus tracing) was used as the diagnostic standard, the proportions of agreement were 54% for session I, 60% for session II, and 61% for session III. The pattern of agreement remained the same, with the study models alone providing adequate information for treatment planning and additional records making small differences. The overall proportions of agreement of 61% for the session III seemed to indicate the lateral cephalograms made a large difference. Inherent "noise" in the study design prevented the proportions of agreement to be greater than 60% without modification of the decision trees and reduced the number of possible treatment options.

Orthodontic cases of Class II, Division 1 malocclusion were selected for this study. This reduced the number of variables and decision trees and increased the potential for reaching levels of statistical significance by the number of cases in each developmental stage of the dentition and age category.

Clinicians require further information from personal history and clinical examination that is not available from the diagnostic records. We agree that much pertinent information regarding chief complaint, functional aspects of malocclusion, periodontal condition, and some other attributes of a patient can be obtained only from history and clinical examination. In this sense the diagnostic materials studied are adjuncts to the total evaluation but, in our view, provided the major information content of the data base on which therapeutic decisions are based.

The order in which the records are presented to the orthodontists may be varied, with the panoramic radiograph after the tracing of the lateral cephalogram. For those orthodontists whose diagnosis and treatment planning depends on cephalometric analyses, the tracing could be substituted as the first record before the study models.

Future investigation with a broader spectrum of malocclusion and more clinicians could develop a decision-making algorithm in diagnostic records for patients with various types of malocclusion at different ages. The utility of the diagnostic records may be assessed by evaluating the treatment outcome as a result of the diagnostic process. The utility of the records should be evaluated in terms of cost-benefit and risk-benefit ratios.

SUMMARY

Finally, the techniques and discipline of clinical epidemiology and clinical decision analysis should be rigorously applied to evaluate the efficacy of various diagnostic procedures as well as of the treatment efficacy in orthodontics. This preliminary study investigated the usefulness of orthodontic records and has provided a basis for ongoing clinical study to evaluate the efficacy of orthodontic treatments.

REFERENCES

1. Alper PR. Medical practice in the competitive market. N Engl J Med 1987;316:337-9.
2. Aaron H, Schwartz WB. Rationing health care: the choice before us. Science 1990;247:418-22.
3. Feinstein AR. Clinical epidemiology: the architecture of clinical research. Philadelphia: WB Saunders, 1985.
4. Weinstein MC, Feinberg HC, Elstein AS, et al. Clinical decision analysis. Philadelphia: WB Saunders, 1980.
5. Krischer JP. Utility structure of a medical decision making problem. Operations Res 1978;24:951-72.
6. Antczak-Bouckoms AA, Weinstein MC. Cost-effectiveness analysis of periodontal disease control. J Dent Res 1985;16:1630-5.
7. Douglass CW, Valachovic RW, Wijesinha A, Chauncey HH, Kapur KK, McNeil B. Clinical efficacy of periodontal disease. Oral Surg Oral Med Oral Pathol 1986;62:330-9.
8. Tulloch JF, Antczak-Bouckoms AA. Decision analysis in the evaluation of clinical strategies for the management of mandibular third molars. J Dent Educ 1987;51:652-60.
9. Vig PS. Reflections on the rationality of orthodontics: toward a new paradigm. In: Vig PS, Ribbens KA, eds. Science and clinical judgement in orthodontics. Monograph 19, Craniofacial Growth Series. Center for Human Growth and Development, The University of Michigan, Ann Arbor, 1986.
10. Proffit WR, Ackerman JL. Diagnosis and treatment planning in orthodontics. In: Graber TM, Swain SB, eds. Orthodontics: current principles and techniques. St. Louis: CV Mosby, 1985.
11. Weed LL. Medical Records, medical education and patient care: the problem-oriented record as a basic tool. Cleveland: Case Western Reserve Press, 1969.
12. Gottlieb EL, Nelson AH, Vogels DS. 1986 JCO study of orthodontic diagnosis and treatment procedures. Part I. Overall results. J Clin Orthod 1986;20(9):612-25.
13. Rowe KGT. The concordance of pre-treatment malocclusion assessment among orthodontic specialty practitioners [Master's thesis]. Ann Arbor, Michigan: Horace Rackham School of Graduate Studies, University of Michigan, 1989.
14. Weintraub JA, Vig PS, Brown C, Kowalski CJ. The prevalence of orthodontic extractions. AM J ORTHOD DENTOFAC ORTHOP 1989;96:462-6.
15. Naccache H, Bernard C, Brodeur JM, Pournier A. Epidemiological evaluation of a computerized diagnosis in orthodontics. J Dent Res 1989;68:Abst. 776.

Reprint requests to:
Dr. Katherine W. L. Vig
Department of Orthodontics
School of Dental Medicine
University of Pittsburgh
3501 Terrace St.
Pittsburgh, PA 15261