European Journal of Orthodontics 24 (2002) 99–106

Fuzzy modelling for selecting headgear types

M. Okan Akçam* and Kenji Takada**
Departments of Orthodontics, Faculties of Dentistry, *Ankara University, Turkey, and **Osaka University, Japan

SUMMARY The purpose of this study was to develop a computer-assisted inference model for selecting appropriate types of headgear appliance for orthodontic patients and to investigate its clinical versatility as a decision-making aid for inexperienced clinicians. Fuzzy rule bases were created for degrees of overjet, overbite, and mandibular plane angle variables, respectively, according to subjective criteria based on the clinical experience and knowledge of the authors. The rules were then transformed into membership functions and the geometric mean aggregation was performed to develop the inference model. The resultant fuzzy logic was then tested on 85 cases in which the patients had been diagnosed as requiring headgear appliances. Eight experienced orthodontists judged each of the cases, and decided if they 'agreed', 'accepted', or 'disagreed' with the recommendations of the computer system. Intra-examiner agreements were investigated using repeated judgements of a set of 30 orthodontic cases and the kappa statistic. All of the examiners exceeded a kappa score of 0.7, allowing them to participate in the test run of the validity of the proposed inference model.

The examiners' agreement with the system's recommendations was evaluated statistically. The average satisfaction rate of the examiners was 95.6 per cent and, for 83 out of the 85 cases, 97.6 per cent. The majority of the examiners (i.e. six or more out of the eight) were satisfied with the recommendations of the system. Thus, the usefulness of the proposed inference logic was confirmed.

Introduction

Headgear is mainly used in orthodontic practice to deliver extra-oral forces to the upper dental arch for anchorage purposes, distalizing teeth and/or inhibiting forward maxillary growth. It has three main types, i.e. low, medium, and high-pull describing the direction of force applied to the upper molar teeth in the sagittal plane (Williams *et al*., 1996). The choice of the precise type of headgear may not be difficult when considering its application in 'typical' cases, such as those exhibiting a Class II malocclusion with a deep overbite, large overjet, and a low mandibular plane angle. A problem may arise, however, particularly for orthodontists who have less clinical experience, with 'borderline' or 'marginal' subjects, such as those having a deep overbite, a moderate to severe overjet, and a high mandibular plane angle. This is because decision making in choosing an appropriate headgear type cannot be dealt with in a discrete, but rather a continuous manner, i.e. fuzzy logic.

Generally, in a medical or dental expert system, a set of knowledge base is derived from experienced clinicians and represents their knowledge, which can be used for clinical consultations (Stheeman *et al*., 1992; Takada *et al*., 1998). With this type of system, uncertainty is a major problem in decision making because non-evidence-based knowledge has to be represented mathematically (Hudson and Cohen, 1994). Fuzzy logic (Zadeh, 1965) has been applied to dental and medical sciences (Sims-Williams *et al*., 1987) in order to construct systems that can infer precise recommendations for solving problems that have uncertain properties (Sims-Williams *et al*., 1987; Mackin *et al*., 1991; Hudson and Cohen, 1994; Stephens *et al*., 1996; Tanaka *et al*., 1997).

Brown *et al*. (1991) applied fuzzy logic to solve orthodontic problems in an expert system, designed to provide advice for treatment planning of Class II division 1 malocclusions. They reported that their system produced more acceptable treatment plans than those used by general dental practitioners. Similarly, Tanaka *et al*. (1997) applied fuzzy reasoning to their computer-assisted diagnostic system for ultrasonography for the purpose of providing a diagnostic aid for unskilled clinicians.

The purpose of the present research was to develop an expert system based on fuzzy reasoning, which could infer an optimum selection of headgear type for orthodontic patients and then to test its reliability. The construction of such a system would broaden understanding of the rational basis of headgear selection and provide the prospect of assisting the less experienced in the delivery of clinical care.

Materials and methods

Materials

Pre-treatment orthodontic records of 85 Japanese patients (33 males, 52 females, mean age 12.9 years, SD 4.6 years, range 8.1–31.1 years) who had been or were under treatment by means of headgear appliances at the university dental hospital were used. The records consisted of dental casts, intra- and extra-oral photographs, a panoramic radiograph and a lateral cephalogram, and accompanying analysis, which included the means and standard deviations of the cephalometric measurements.

Inference modelling

Fuzzy logic is concerned with quantifying and reasoning with vague or 'fuzzy' terms, which appear in our natural language. It allows intermediate values of certainty to be defined between conventional deterministic two-valued logic, such as yes/no, high/low, or true/false. Intermediate values can be expressed mathematically in terms of the degree to which an element belongs to a 'fuzzy set', where 0 represents non-membership and 1 complete membership of that set. According to Zadeh (1965), a fuzzy set A in the universal space X is characterized by a membership function fA(x), which associates with each point in X a real number in the interval [0,1], with the value of fA(x) at X representing the 'membership grade' of x in A. Thus, the nearer the value of fA(x) to unity, the higher the grade of membership of x in A. One of the essential steps that influence the versatility of the inference system is the creation of membership functions, i.e. fuzzy sets that assume values in the range from 0 to 1 defined in universal space. The mathematical forms of each membership function are usually selected subjectively (Stachowicz and Beall, 1995) and may be expressed in linear, trapezoid, or Gaussian forms.

In the present study, three variables, i.e. overjet, overbite, and mandibular plane angle, were used as input variables to the system. The mandibular plane angle was defined as the angle formed by the SN and mandibular planes. These variables were obtained from the lateral cephalograms. For each input variable, three fuzzy sets for the low, medium, and high-pull types of headgear were defined on the basis of the authors' subjective judgement, which included their clinical experience and knowledge of the normative means and standard deviations for each variable (Susami, 1967; Riolo *et al*., 1974). The J hook type headgear was not included in the current paradigm. For each fuzzy set, the fuzzy trapezoid function was employed to construct membership functions. The universal spaces for each of the three input variables were divided into the aforementioned fuzzy sets. The fuzzy sets for each variable were determined with an assumption that the remaining two variables took normative values. For ease of understanding and simplicity, a graphic interpretation of the element and membership grade pairs which were created for the low, medium and high-pull types using each input variable is provided in Figures 1–3.

With regard to overjet (Figure 1), an assumption was made that there would be no chance of any type of headgear being used where the overjet was equal to or smaller than 2 mm. If the overjet was between 3 and 4 mm, the degree of certainty for using the high-pull type was assumed as 1. However, the membership grade for using the

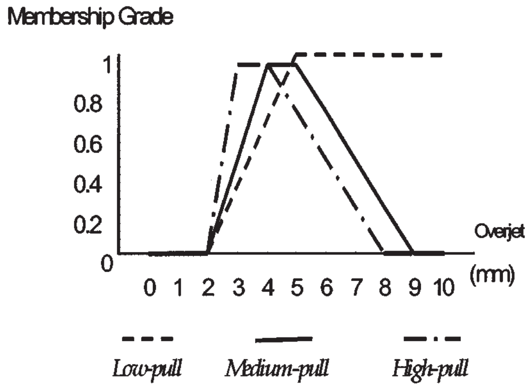

Figure 1 Plot of membership functions for the input of overjet for each of three sets, i.e. the low-, medium-, and high-pull types of headgear.

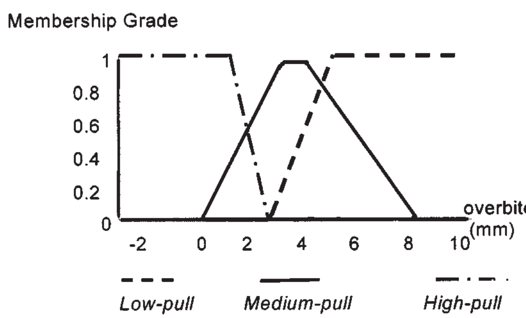

Figure 2 Plot of membership functions for the input of overbite for each of three sets, i.e. the low-, medium-, and high-pull types of headgear.

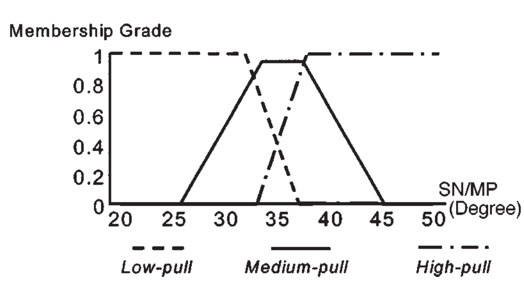

Figure 3 Plot of membership functions for the input of mandibular plane angle for each of three sets, i.e. the low-, medium-, and high-pull types of headgear.

high-pull type was assumed to decrease to 0 if the overjet was equal to or larger than 8 mm. The membership grade for use of the medium-pull type increased from 0 with an overjet of 2 mm to 1 at an overjet of 4 mm, from which it decreased linearly reaching 0 again with an overjet of 9 mm. The membership grade for use of the low-pull type was 1 if the overjet was larger than 5 mm and 0 if the overjet was 2 mm or less. Within the range 2–5 mm there was a linear increase.

For overbite (Figure 2), the premise was applied that the simultaneous choice of low- and high-pull headgear types could not co-exist. Thus, the universal spaces of the overbite for each headgear type were designed so that there was no chance of the system recommending all three headgear types concurrently. The membership grade for use of the low-pull type was assumed to increase linearly from 0 to 1 in the overbite range from 2 to 5 mm, and be 1 where the overbite was equal to or greater than 5 mm. For the medium-pull type, the membership grade increased linearly from 0 to 1 as the overbite increased from 0 to 3 mm. Above 4 mm it decreased again from 1 to 0 at an overbite of 8 mm or more. Membership grade for high-pull headgear was assumed to decrease from 1 to 0 in the range from 1 to 3mm.

With respect to the mandibular plane angle (Figure 3), the high-pull type headgear was assumed to show a linear increase in membership grade from 0 to 1 in the range from 33 to 38 degrees, while that for the low-pull type was assumed to decrease linearly from 1 to 0 in the range of 32–37 degrees. The medium-pull type headgear showed a linear increase from 0 to 1 in the range 25–34 degrees, a constant value of 1 between 34 and 37 degrees, and a decrease from 1 to 0 from 37 to 45 degrees. Because there were three fuzzy sets, i.e. membership grades for each of the three input variables, for a given headgear type, geometric mean aggregation was used for inference (Klir and Folger, 1988). Geometric mean aggregation was the operation by which multiple fuzzy sets were combined to produce a single fuzzy set. Thus, for a given input data set of a patient, the inference system was designed so that it could calculate degrees of certainty for the use of each respective headgear

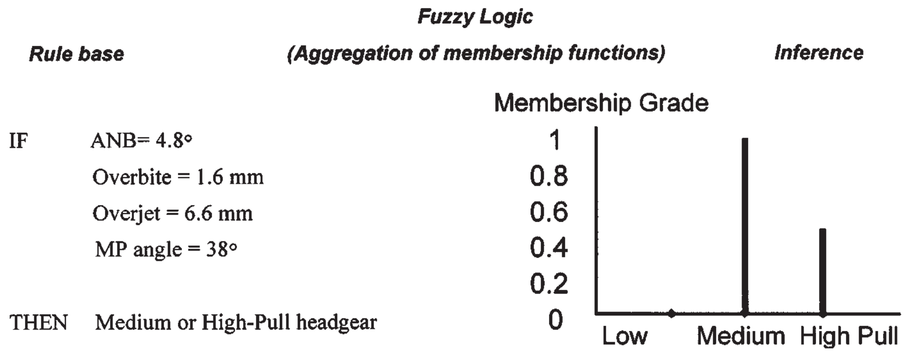

Figure 4 The computer provides a selection of headgear types in which each choice is accompanied by a membership grade.

type by means of membership grades (Figure 4). Calculations were made using fuzzy logic software (Mathematica 2.2, Fuzzy Logic Pack, Wolfram Research, USA).

Experimental implementation

Eight orthodontists (six men and two women) from different centres whose clinical experience ranged between 10.1 and 20.9 years (mean 14.7 years, SD 3.7 years) took part in this study as independent validators. They had not been provided with any information on the experimental design and purpose of the investigation. To determine whether or not they could reliably choose the type of headgear for a given patient in a reproducible manner, they were asked to make this choice in 30 cases that were randomly selected from the study sample of 85 cases. This was carried out twice with a minimum interval of one week. The strength of agreement between the two judgements was measured using the kappa statistic (SPSS Version 6.1. Base System, Illinois, USA) and classified into five categories according to the method of Landis and Koch (1977; Table 1). Examiners whose kappa statistics exceeded 0.6 were accepted to participate in the testing of the inference model as described below.

The selected examiners were presented with full orthodontic records of 85 cases, which had been diagnosed by a group of orthodontists who were independent of the present study, as those in need of using headgear appliance. Membership grades for each of three headgear types were computed by the inference system for each case and presented to the examiners. The system's output consisted of two parameters, i.e. the names of the headgear types and their corresponding membership grades. The examiner evaluated each case and stated whether they would 'agree', 'accept', or 'disagree' with the recommendations given by the system. If an examiner completely agreed with the system's output, their judgement was assigned as 'agreed'. If the examiner agreed with the type of headgear that had been recommended by the system, but did not necessarily consent to the degree of certainty, the system's outputs were assigned as 'acceptable'. Finally, if the examiner did not agree with the type of headgear or degree of membership grade proposed by the system, the judgement was recorded as

Table 1 Kappa score evaluation according to Landis and Koch (1977).

Kappa statistic	Strength of agreement
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

'disagreed'. The categories 'agreed' and 'acceptable' were compounded into a single category—'satisfactory'. The recommendations for headgear choice provided by the inference model were examined using the simplified categories of 'satisfactory' and 'disagreed'. The proportion of 'satisfactory' judgements confirmed by each examiner in the total sample examined was calculated. In addition, the proportion of cases to which a significant majority of the orthodontic experts (i.e. equal to or more than six examiners out of the eight) assigned 'satisfactory' was computed. The percentages of cases to which (1) three, four, or five out of the eight experts; and (2) two or less experts assigned 'satisfactory' with or all of them rejected the recommendations given by the system were also calculated.

Results

The results for intra-examiner agreement are given in Figure 5.

From the kappa scores, it was concluded that five examiners had 'almost perfect' agreement and the remaining three examiners showed 'substantial' agreement between their first and second judgements. Accordingly, all examiners participated in the test of the inference model, the results of which are provided in Table 2.

The percentages of the system's recommendations for headgear choice that were considered 'satisfactory' by orthodontic experts ranged from 91.8 to 98.8 per cent with an average of 95.6 per cent. The examiners who were in the lowest two ranks for the intra-individual agreement exhibited the lowest proportion of the category 'satisfactory'. The proportion of the system's recommendations to which six or more of the eight orthodontic experts assigned 'satisfactory' was 97.7 per cent, while for the remaining 2.3 per cent of cases nearly half of the examiners were satisfied with the system's responses.

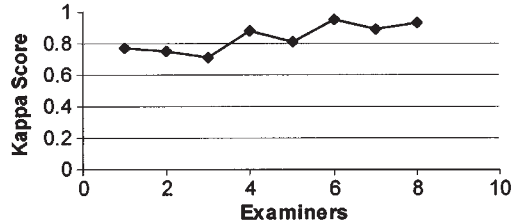

Figure 5 Kappa scores between the first and the second judgement sessions.

Discussion

In the process of diagnosing and treatment planning a malocclusion, patients are first observed to identify the various traits exhibited by them. These traits are then considered in order to elucidate the nature of the orthodontic problems and choose an optimum approach for solving them. Decision making through these processes may be performed on the basis of scientifically proven evidence, but also on speculations or beliefs that have been assumed

Table 2 Proportions of system's recommendations for headgear choice, which were considered 'satisfactory' by eight orthodontic experts. Evaluation was made for 85 cases.

	Examiners							
	1	2	3	4	5	6	7	8
Sample size	85	85	85	85	85	85	85	85
Agreed	71	70	67	81	77	78	82	78
Acceptable	9	8	11	3	5	3	2	5
Disagreed	5	7	7	1	3	4	1	2
Satisfactory (%)	94.1	91.8	91.8	98.8	96.5	95.3	98.8	97.6

Average satisfaction (%) 95.6 (SD 2.6).

to be valid from the clinician's past experience and knowledge. Such subjectivity in the evaluation of the membership grade may be characterized by one or more fuzzy sets. Subjectivity has the advantage that it enables one to take into account human experience and intelligence by translating vague human reasoning into a mathematical model (Stachowicz and Beall, 1995).

In the present study, those variables that were considered to characterize developmental stages or the sagittal jaw relationships of the patient were not incorporated. Although these traits are important factors for orthodontists when deciding whether or not they should use headgear, they do not play a deterministic role in the choice of headgear type. For the latter, it is the direction of extra-oral force to be applied that is the key factor. Similarly, the degree of crowding was not used as an input variable. This variable is important in determining whether the treatment can be performed with or without the extraction of teeth. Thus, if a patient shows mild to severe crowding of teeth in the upper arch, then either the upper molar teeth may be distalized or the premolar teeth extracted to acquire space for the remaining teeth anterior to the molars to be aligned. In either situation, headgear may be applied for anchorage purposes in order to prevent molar teeth being tipped mesially, but this cannot be a determining factor in the choice of headgear type.

The mandibular plane angle was used as one of the input variables. However, this does not exclude the possibility of other variables being incorporated. For instance, one could choose the maxillo-mandibular plane angle or the lower anterior face height as an alternative to the S–N/mandibular plane angle. There is no absolute criterion for selection of variables. Those used in the current paradigm were selected on the assumption that, if the result produced by the inference model with the chosen set of variables showed a high proportion of coincidence with judgements given by a majority of orthodontic experts, it could be concluded that the selection of those variables was appropriate or validated. The ranges of the values for the mandibular plane angle that were employed in the current paradigm for the three types of headgear appliances may differ from those appropriate for other ethnic groups; this is because the Japanese are known to exhibit a steeper mandibular plane angle when compared with, for example, North Europeans. An assumption was followed that the steepness of the mandibular plane in relation to the selection of the headgear appliances is understood with respect to the statistical distribution pattern of this variable in each individual ethnic group and does not necessarily correspond with force directions of the headgear appliances. If the inference model was applied to other ethnic groups, this difference, however, could easily be adjusted mathematically by referring to a conversion table, which specifies the relationships between normative values of this variable in the two populations, e.g. 36.4 degrees (SD 5.1 degrees) for Japanese male adults and their corresponding values for Caucasians, 32.6 degrees (SD 5.2 degrees; Susami, 1967; Riolo *et al.*, 1974).

The performance reliability of the inference model described was tested by showing the system's recommendations to a group of experienced orthodontists, and asking them whether they would accept them or not (Weiss and Kulikowski, 1987). This method is employed frequently where there are no definite criteria that can be used as a basis for decision making. It should be noted, however, that an earlier exposure to the opinion generated by the model may influence the opinions of examiners testing such a system and invalidate their opinions.

Previous reports have shown that a large amount of data can be integrated and processed correctly by an expert system (Erdman, 1987; Kinney, 1987). The use of such a system is, therefore, considered to be capable of improving the quality and accuracy of the clinical decision-making process (De Dombal *et al.*, 1972). Brown *et al.* (1991) reported, from examination of 200 cases, the incidence of incorrect treatment planning by dental practitioners to be 41–46 per cent and they suggested that a computer-based expert system might even provide more acceptable treatment plans than those of specialists. The use of an orthodontic expert system in general dental practice has been suggested as being capable of reducing the incidence of incorrect diagnosis and treatment (Stephens *et al.*, 1996).

The results of the test run in this study confirmed the usefulness of the model under investigation. On average, the orthodontic experts were satisfied with the system's recommendations in almost 96 per cent of the cases. The remarkably high proportion of the system's recommendations to which the majority of the orthodontic experts assigned 'satisfactory' also confirmed the effectiveness of the system.

All examiners showed high scores for intra-examiner agreement between the two judgement sessions in evaluating and choosing headgear types. This means that the task for choosing headgear types can be performed in a consistent manner if clinicians have sufficient knowledge and experience. It is interesting, however, to note that the five examiners who had kappa scores above 0.8 showed a relatively higher 'satisfaction' rate with the system's recommendations, while the two examiners with the lowest kappa scores also showed the lowest satisfaction rates. In other words, the results suggest that, even among experienced clinicians, those having the ability to make more consistent judgements are likely to give opinions that are consistent with the system's advice. It is possible though that those who are more unreliable may have decision criteria that are different from those shared by the majority of the examiners and the system. In order to evaluate whether the cause of the relatively lower rate of satisfaction found in examiners 2 and 3 was due to the ineffectiveness of the system or the difference in criteria between the examiners for choosing the appropriate headgear types, those cases that were judged as 'disagreed' between the system and examiners 2 and 3 were re-examined. It was found that for all of these cases, the remaining examiners had judgements that were consistent with the system's recommendations. In other words, although these two examiners exhibited essentially high rates of 'satisfaction', their rationale for choosing headgear appliances seemed to differ both from those held by the remaining examiners and the authors of the inference model.

In summary, the system's recommendations were found to be consistent with the opinions of the majority of the examiners and it was concluded that the system's advice closely matched that given by the experts. Because headgear appliances have been used worldwide by dental practitioners with various levels of knowledge and experience in orthodontics, the use of an advisory system as proposed in the current report would help inexperienced clinicians reduce the occasions on which they choose inappropriate headgear types.

Conclusions

A fuzzy model that can infer precise choice of headgear types appropriate to the treatment of an orthodontic case has been developed. The model was designed to calculate the degree of certainty for choosing low-, medium- or high-pull types of headgear. Eight orthodontic experts evaluated the decisions inferred by the system for 85 orthodontic cases. This group of clinicians was satisfied with the system's recommendations in 95.6 per cent of the cases. In addition, the majority of the examiners (i.e. equal to or more than six out of eight) were satisfied with the system's recommendations in 97.6 per cent of the cases examined. Thus, the inference system developed was confirmed as being reliable and effective for clinical use in orthodontics.

Address for correspondence

Dr M. Okan Akçam
Ankara Üniversitesi
Diş Hekimliği Fakültesi
Ortodonti AD
Beşevler,
Ankara 06500
Turkey

Acknowledgements

We would like to thank to Drs K. Nakagawa, S. Adachi, Yh. Yasuda, M. Kawakami, N. Kitai, Y. Sorihashi, Yt. Yasuda, Y. Kakiuchi, and S. Maeda for their kind contributions to this study as volunteer examiners. We are also indebted to Mrs Elizabeth Williams for grammatical correction of the manuscript.

References

Brown I D, Adams S R, Stephens C D, Erritt S J, Sims-Williams J H 1991 The initial use of a computer controlled expert system in the treatment planning of Class II division 1 malocclusion. British Journal of Orthodontics 18: 1–7

De Dombal F T, Leaper D J, Staniland J R, McCann A P, Horrocks J C 1972 Computer-aided diagnosis of acute abdominal pain. British Medical Journal 2: 9–13

Erdman H P 1987 A computer consultation program for primary care physicians. Impact of decision making model and explanation capability. Medical Care 25: 138–147

Hudson D L, Cohen M E 1994 Fuzzy logic in medical expert systems. IEEE Engineering in Medicine and Biology 3: 693–698

Kinney E L 1987 Medical expert systems. Who needs them? Chest 91: 3–4

Klir G J, Folger T A 1988 Fuzzy sets, uncertainty and information. State University of New York, Binghampton. Prentice Hall, Englewood Cliffs, New Jersey

Landis J R, Koch G G 1977 The measurement of observer agreement for categorical data. Biometrics 3: 59–74

Mackin N, Sims-Williams J H, Stephens C D 1991 Artificial intelligence in dental surgery: an orthodontic expert system, a dental tool of tomorrow. Dental Update 18: 341–343

Riolo M L, Moyers R E, McNamara J A Jr, Hunter W S 1974 An atlas of craniofacial growth, Monograph No. 2, Craniofacial Growth Series. Center for Human Growth and Development, University of Michigan, Ann Arbor

Sims-Williams J H, Brown I D, Matthewman A, Stephens C D 1987 A computer controlled expert system for orthodontic advice. British Dental Journal 163: 161–166

Stachowicz M S, Beall L 1995 Fuzzy logic pack, printed version of the notebook, Mathematica. Soft Control Inc., Minneapolis, Wolfram Research Inc., Champaign, Illinois

Stephens C D, Mackin N, Sims-Williams J H 1996 The development and validation of an orthodontic expert system. British Journal of Orthodontics 23: 1–9

Stheeman S E, Van der Stelt P F, Mileman P A 1992 Expert systems in dentistry past performance—future prospects. Journal of Dentistry 20: 68–73

Susami R 1967 A cephalometric study of dentofacial growth in mandibular prognathism. Journal of Japanese Orthodontic Society 26: 1–34

Takada K, Sorihashi Y, Akçam M O 1998 Orthodontic treatment planning: its rationale for inference. In: Carels C, Willems G (eds) The future of orthodontics. Leuven University Press, Belgium, pp. 203–211

Tanaka T, Miwa K, Kanda S 1997 Application of fuzzy reasoning in an expert system for ultrasonography. Dentomaxillofacial Radiology 26: 125–131

Weiss S M, Kulikowski C A 1987 A practical guide to designing expert systems, Japanese translation. Mori K Kindai Kagagusha Publishing, Tokyo

Williams J K, Cook P A, Isaacson K G, Thom A R 1996 Fixed orthodontic appliances—principles and practice. Wright, Oxford

Zadeh L A 1965 Fuzzy sets. Information and Control 8: 338–353