Journal of Orthodontics, Vol. 29, 2002, 287–291

SCIENTIFIC SECTION

Perception of orthodontic treatment need: opinion comparisons of orthodontists, pediatric dentists, and general practitioners

N. W. Berk, H. Dukich Bush, J. Cavalier, R. Kapur, D. Studen-Pavlovich, J. Sciote, & R. J. Weyant
University of Pittsburgh, USA

Abstract

Aim: To determine the relationship between treatment need assessment scores of orthodontists, general practitioners, and pediatric dentists.

Study design: Observational.

Sample: Ten general dental practitioners, 18 orthodontists and 15 pediatric dentists reviewed 137 dental casts and recorded their opinion on whether orthodontic treatment was needed.

Results: We found a high level of agreement between pediatric dentists, orthodontists and general practitioners (Kappa range 0.86–0.95). Between the groups, the amount of agreement was lower.

Conclusions: Orthodontists, general dental practitioners, and pediatric dentists in this sample exhibit high levels of agreement on orthodontic treatment need.

Index words: Orthodontic, treatment need.

Received 2 May 2002; accepted 13 June 2002

Introduction

A person's decision to seek orthodontic treatment is often based on multiple factors. While treatment seeking may be solely patient initiated, it is most often professionally initiated or influenced by referral from general dentists, pediatric dentists, or orthodontists.[1] Given possible differences in educational training, perception of orthodontic treatment need may differ by dentist group or specialty affiliation. Although the gold standard for assessment of orthodontic treatment need is assessment by an orthodontic specialist, it is important to understand the perceptions of other dental professionals as they can impact, directly and indirectly, the utilization and success of orthodontic treatment.

The purpose of this study was to determine the relationship between treatment need assessment scores of orthodontists, general practitioners, and pediatric dentists. This study compares the orthodontic treatment need opinions of orthodontists to those of non-orthodontic dentists (pediatric dentists and general practitioners) who tend to most often refer patients for orthodontic treatment.

Materials and methods

Participants

Three panels of dentists were recruited to serve as raters. Recruitment was based on dental clinical training, with the aim of creating three panels composed of:

- general dental practitioners (no specialty training);
- pediatric dentists;
- orthodontists.

For practical reasons, the selection was limited to Western Pennsylvania, primarily Allegheny County, where the city of Pittsburgh and the University of Pittsburgh are located. Participation was solicited through letters sent to dentists' offices. Interested individuals were asked to contact one of the investigators (either HD or JC) and details of the study were then provided.

Address for correspondence: Nancy W. Berk, Department of Behavioral Sciences, School of Dental Medicine, University of Pittsburgh, 3501 Terrace Street, Pittsburgh, PA 15210, USA. Email: nwb@sdmgenetics.pitt.edu

Once each panel was assembled, the participants were scheduled to visit the University of Pittsburgh's School of Dental Medicine, where they participated in the rating procedure similar to that utilized by Beglin *et al.*[2]

Procedure

The members of each panel were asked to score 137 study models with respect to their need for orthodontic treatment. These models represented the full range of severity of malocclusion based on the Dental Health Component of the Index of Orthodontic Treatment Need.[3] Most of the cases were in the permanent dentition, but there was a small percentage of late mixed dentition cases that contained over-retained deciduous teeth.

For each study model, each rater was given the following instructions: '*You are the dental consultant for a private corporation for which a fund has been established to provide orthodontic treatment for personnel. You are to assess these 137 models of personnel and answer the following question: In your opinion, to what extent does this occlusion need orthodontic treatment? Please circle the number of your score (None/minimal = 1, Great = 7).*' These instructions were designed to minimize extraneous factors that might influence a rater's opinion of treatment need. For example, to alleviate cost concerns, no amount of treatment funds was specified to the participants. This judgment allowed each participant to concentrate on the determination of orthodontic treatment need and not the amount of money that was available. Financial responsibility for the treatment was stated as coming from a private corporation not associated with the consultant. Additionally, there was no mention of the treatment provider or beneficiary of the fee.[4,5]

Following the rating of all 137 casts, each participating dentist was asked to indicate on the same 7-point scale, the cut-off point or score above, which he/she believed treatment would be indicated for any patient. This data point was utilized to tap indicated treatment need (ITN) as a dichotomous variable. In this way, treatment need could be assessed as either (1) indicated or (2) not indicated for each participant's rating for all casts.

Intra-rater agreement. To assess intra-rater reliability, a second rating session took place 4 weeks following the first session and employed the same rater instructions. However, in the second session, participants were asked to assess the need for orthodontic treatment on only 37 casts. Intra-rater reliability was determined using two statistical techniques and comparing each rater's treatment need scores between the first and second ratings. First, an intra-class correlation coefficient (ICC) was used to calculate the intra-rater reliability of each dentist rater. The intra-class correlation coefficient provides a measure of the reliability of the rating process, with values greater than 0.75 considered to be 'highly reliable'.[6]

Second, to account for chance agreement, a weighted kappa statistic[7] was calculated using the weights suggested by Fleiss and Cohen.[8]

Inter-rater agreement. Inter-rater reliability was evaluated by calculating a weighted kappa statistic for a randomly selected set of 20 pairwise comparisons of raters within each group of dentists. Similar to the method used for intra-rater agreement, the weights suggested by Fleiss and Cohen [8] were used.

Comparison of ratings among dentist groups. To compare group intra-rater scores, a mean 'treatment need' score was calculated for each of the 137 casts for each group of dentists. For example, for cast number 1, a mean score across the 18 orthodontists raters was calculated, similarly, the mean score for cast number 2 was calculated and so on. Likewise, mean scores for each cast were calculated for the panels of pediatric dentists and general practitioners.

Using each panel's mean score, all pairwise comparisons between dentist groups were calculated using Pearson's correlation coefficient as a measure of agreement on treatment need. Additionally, all pairwise associations between groups were done using a paired *t*-test to assess for systematic scoring differences (bias) that would not be detected by the correlation coefficient.

Calculation of TI variable. The ITN (the cut-off point given for indicated treatment need) for each dentist was compared with each of that dentist's numerical ratings on the 137 casts in order to create a new/resultant variable. This two-level nominal variable indicated whether each cast should receive either (1) treatment or (2) no treatment. This resultant variable was called 'Treatment Indicated' (TI).

Calculation of PIT variable. Using the TI variable for each dentist, a new variable, 'percentage indicating treatment' (PIT), was calculated for each cast for each panel of dentists. Thus for cast number 1, the percent of

orthodontists indicating treatment needed was calculated by dividing the number of orthodontists indicating treatment need for that cast by the total number of orthodontist raters ($n = 18$). A similar approach was used for pediatric dentists ($n = 15$) and general practitioners ($n = 10$), using the appropriate denominator. Thus, three new continuous 'percentage indicating treatment' (PIT) variables were created, one for each panel of dentists.

Once again, all pairwise association between panels were done using correlations and a paired *t*-test to assess for systematic differences in dentists' patterns of opinion of orthodontic treatment need.

Calculation of TC variable. The final analysis was done with the use of the 'percent indicating treatment' variable, by creating a new three-level categorical variable for each panel of dentists. The new variable, 'treatment concordance' (TC) had the following values: (1) treatment agreement, (2) treatment uncertain and (3) no treatment agreement. The variable was created for each panel of dentists as follows: for a given cast, if the percentage indicating treatment (PIT) was at least 75 per cent, then the cast was scored as 'treatment agreement'; if the PIT was 25 per cent or less, then the cast was scored as 'no treatment agreement', and casts with PIT between 25 and 75 per cent were scored as 'treatment uncertain'.

The resultant three-level categorical variables were compared in pairwise comparisons using chi-square and percent agreement.

Results

Pediatric dentists' reliability estimates

Intra-rater reliability. Intra-rater reliability using ICC and weighted kappa was high for all dentist groups (Table 1). With the exception of one rater among the pediatric dentists, whose ICC score was 0.638, all ICC measures of intra-rater reliability for all dentists were in the 'highly reliable' range.[6] Furthermore, all dentists' weighted kappa scores were in the ranges defined by Landis and Koch as either 'substantial' (kappa between 0.61 and 0.80) or 'almost perfect agreement' (kappa greater than >0.8).[9]

Inter-rater reliability. The results for all three panels of dentists show that inter-rater agreement was generally in the 'substantial' or 'almost perfect' range (Table 2). However, one general practitioner in one pairwise comparison had only a 'fair' (kappa between 0.21 and 0.40) strength of agreement. In all other comparisons, this rater demonstrated substantial to almost perfect agreement.

Association between pediatric dentists, orthodontists and general practitioners opinion of orthodontic treatment need. Systematic differences could occur whereby one panel of dentists consistently rated treatment need as higher than another group (Table 3). Therefore, this was evaluated by carrying out a paired *t*-test between the pediatric dentists and each of the other panels. In both cases, that is, the comparison between pediatric dentists and orthodontists ($p = 0.18$) and between pediatric dentists and general practitioners ($p = 0.89$), no significant differences were found.

Table 1 Results of ICC and weighted kappa statistics for the three panels of dentist raters

	n	Mean ICC (range)	Mean Kappa (range)
General practitioners	10	0.891 (0.768–0.983)	0.865 (0.730–0.942)
Orthodontists	18	0.872 (0.801–0.934)	0.896 (0.785–0.969)
Pediatric dentists	15	0.873 (0.637–0.935)	0.951 (0.918–0.974)

Table 2 Results of a weighted kappa statistic for 20 randomly selected pairwise comparisons for the three panels of dentist raters

	n	Mean Kappa (range)
General practitioners	10	0.733 (0.259–0.931)
Orthodontists	18	0.835 (0.707–0.932)
Pediatric dentists	15	0.808 (0.509–0.955)

Table 3 Results of correlation and paired *t*-test among the three groups of dentists.

Correlation (*p*-value for paired *t*-test of bias)	General practitioners	Orthodontists	Paediatric dentists
General practitioners	—	0.968 ($p = 0.066$)	0.952 ($p = 0.891$)
Orthodontists	—	—	0.929 ($p = 0.175$)
Pediatric dentists	—	—	—

Table 4 Correlations and paired *t*-test results (expressed as p-values) among the three panels of dentists using the PIT variable for each panel

Correlation (*P*-value for paired *t*-test of bias)	General practitioners	Orthodontists	Pediatric dentists
General practitioners	—	0.943 ($p = 0.0004$)	0.937 ($p = 0.0001$)
Orthodontists	—	—	0.940 ($p = 0.133$)
Pediatric dentists	—	—	—

Results from the ITNP calculations

The results of the pairwise correlations and ANOVA for the 'percentage indicating treatment' (PIT) demonstrated strong agreement between the groups (Table 4).

Analysis of the three-level 'treatment agreement' variable showed moderately high concordance between general practitioners and orthodontists (85 per cent agreement, $p < 0.0001$) and pediatric dentists (84 per cent agreement, $p < 0.0001$), and very high agreement between pediatric dentists and orthodontists (92 per cent agreement, $p < 0.0001$). When there was lack of agreement on the need for treatment in pairwise analysis, orthodontists more often (16 cases) indicated treatment need than general practitioners (four cases) when the other group did not. Pediatric dentists were even more likely to indicate treatment (18 cases) than general dentists (four cases). Finally, orthodontists and pediatric dentists agreed more often, with discordance about treatment being higher for the pediatric dentists (seven cases) than the orthodontists (four cases).

Discussion

The results of this study indicate high agreement between pediatric dentists, orthodontists, and general practitioners with respect to their assessment of orthodontic treatment need based upon cast assessment. These findings lend further support to previous research indicating that orthodontists and general practitioners strongly agree in evaluation of orthodontic treatment need.[5,10,11] Findings from this study further suggest that pediatric dentists concur with these other professional groups.

It is of interest that despite training differences between the panels, high agreement existed between the groups. Given that rationale for orthodontic treatment need was not assessed, it is unclear as to whether the agreement noted is based upon identical rationale for treatment. For example, an orthodontist may have ranked a pair of dental casts a 'five' due to the overjet and overbite, while a pediatric dentist ranked the same set a 'five', but as a result of the overjet and the amount of crowding. Thus, the same score does not necessarily indicate that the panels utilized the same parameters in determining treatment need.

Because we measured only perceived need for treatment, we can make no conclusion regarding intention to refer or referral characteristics. This is in contrast to the work of Bearn *et al.*[12], where intention to refer was directly assessed. Their findings suggest that the treatment decisions of general dentists in the UK are variable, rather than consistent as we noted and that these treatment decisions were not consistent with the current orthodontic guidelines. These contrasting findings may reflect methodological differences and/or differences within the health services environments of Britain and the USA.

The examination of mean treatment need scores in this study provides only a peripheral view of professional opinion. The further examination of indicated treatment need cut-off points by dentist panel suggests that while *perceived treatment* need may not differ by educational training, the cut-off points for *indicated treatment* need do differ by educational training. Specifically, pediatric dentists and orthodontists have lower cut-off points for indicated treatment need than general dentists and thus may be more apt to refer a patient for orthodontic treatment.

Because only subjective opinion of orthodontic treatment need for pairs of dental casts was assessed, it is not known if these findings are generalizable to the clinical setting. No information from photography, radiographs, and patient health histories was presented to the raters. Nor were any interactional, demographic, or other patient variables presented. While diagnosis of malocclusion can be determined by assessing dental casts, this is rarely the only factor examined in the clinical setting. The examination of all available pieces of patient information enables the practitioner to make a more informed assessment of orthodontic treatment need. For example, if photographs were included in the process of determining orthodontic treatment need, the level of severity may appear to be different from when only dental casts were used thus affecting the orthodontic treatment need score. An open bite seen on dental casts with moderate malocclusion would warrant orthodontic treatment, but if that same open bite was presented on a full-face photograph the overall assessment may steer towards a severe malocclusion.

Additional limitations of the present study include the representativeness of the sample. All participants were orthodontists, pediatric dentists and general practitioners located in Western Pennsylvania. While there is no reason to believe that orthodontic opinions would be unique to this area, findings may reflect geographic differences in orthodontic opinion. For this study, however, it appears that regardless of specialty, having a dental degree is associated with an ability to determine orthodontic treatment need. The implications of these findings with respect to referral practices, patterns, and treatment emphasis between dentist groups are not known. The need for further research is essential in order to determine not only the bases of assessment decisions, but if referral decisions and outcomes are consistent across these groups and throughout dental communities.

References

1. O'Brien KD, McComb JL, Wright J. Do dentists refer orthodontic patients inappropriately? *Br Dent J* 2001; **181:** 132–6.
2. Beglin FM, Firestone AR, Vig KWL, Beck FM, Kuthy RA, Wade D. A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need. *Am J Orthod Dentofacial Orthop* 2001; **120:** 240–6.
3. Brook PH, Shaw WC. The development of an orthodontic treatment priority index. *Eur J Orthod* 1989; **11:** 309–20.
4. Younis, J. *Validation of the Index of Orthodontic Treatment Need in the United States*. MDS Thesis, Pittsburgh: University of Pittsburgh; 1995.
5. Dukich, H. Determination of orthodontic treatment need using orthodontists, general practitioners and occlusal indices. MDS Thesis, Pittsburgh: University of Pittsburgh; 1996.
6. Fleiss JL, Slakter MJ, Fischman SL, Park MH, Chilton NW. Inter-examiner reliability in caries trials. *J Dent Res* 1979; **58:** 604–9.
7. Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; **70:** 213–20.
8. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33:** 613–19.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33:** 159–74.
10. Prahl-Anderson B, Boersma H, Van der Linden FPGM, Moore AM. Perception of dentofacial morphology by laypersons, general dentists, and orthodontists. *J Am Dent Ass* 1979; **98:** 209–12.
11. Pietila T, Pietila I. Early screening for orthodontic treatment. *Comm Dent Oral Epidemiol* 1992; **20:** 208–13.
12. Bearn D, Wright J, Kay E, O'Brien K. Perceptions of orthodontic treatment need: Receiver operating characteristic analysis. *Comm Dent Oral Epidemiol* 1996; **24:** 303–6.