

Treatment Strategies for Lower Third Molars Following Orthodontic Care

M. L. JONES B.D.S., M.SC., PH.D., F.D.S., D.ORTH. R.C.S.¹

R. ARMSTRONG E.D.H.²

O. KOSTOPOULOU PH.D.²

M. R. BRICKLEY PH.D., B.D.S., M.SC.D.²

¹Division of Dental Health & Development and

²Clinical Decisions Research Group, Dental Hospital and School, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XY, U.K.

Abstract. *To examine the extent to which orthodontists consider it appropriate to refer post-orthodontic patients for lower third molar treatment. The subjects were 10 orthodontists from hospital, specialist practice and community setting. Participants were presented with 15 case histories of orthodontic patients (including full records) and asked to state whether they would have referred the case to an oral surgeon for management of their third molars. Data were analysed using multi-kappa measure. Two-hundred-and-sixty-two decisions of 300 (88%) were not to refer cases for third molar management. Little consensus was seen between orthodontists on which post-orthodontic cases referral ($\kappa = 0.14$). Some clinicians referred cases much more frequently than others (Kruskall-Wallis = 46.84, $P < 0.001$) and some clinicians referred much more frequently (Kruskall-Wallis = 85.57, $P < 0.001$).*

This group of orthodontists did not refer post-orthodontic cases for third molar management and little consensus was observed regarding which cases did warrant removal.

Index words: Lower Incisor Imbrication, Orthodontics, Third molars.

Refereed Paper

Introduction

Surgery for impacted third molars has achieved pre-eminent status in oral surgery care, occupying substantial resources and manpower in this sector. Perhaps the most important document to date concerning the indications for third molar surgery has been the report of the U.S. National Institutes of Health consensus development conference (Guralnick and Laskin, 1980). Consensus for removal was agreed only where there was evidence of pathology. These criteria included teeth associated with present or past pericoronitis, severe periodontitis, cystic change and unrestorable carious third molars. It was agreed also that the removal of third molars to alleviate or prevent future crowding of lower incisor teeth was unsupported by current evidence.

The practice of removing lower third molars to prevent or alleviate crowding in the lower arch has been a controversial subject for a considerable time. In 1941, Broadbent concluded that pressure from erupting third molars was not responsible for 'buckling' of the lower arch and that the discrepancy was due to failure of the facial skeleton to attain full size and proportions. Cryer (1967) felt that the third molars had been 'unfairly blamed' in causing lower incisor crowding and suggested that both incisor crowding and impaction of the third molars were symptoms related to the same cause: a shortness of the mandibular arch. In 1970, Woodside suggested that if a third molar was not present, then distal settling of the lower arch could occur in response to growth and soft tissue pressure, thus implying a passive role for the third molar in providing an obstacle

to the settling of the dentition rather than actively applying pressure to anteriorly positioned teeth.

Extraction of teeth for orthodontic purposes may improve the potential for eruption of lower third molars and a number of studies have been undertaken to examine this question. Bjork *et al.* (1956) investigated the relationship of mandibular growth to third molar impaction. They employed radiographic evidence from 243 Swedish males, aged between 12 and 20 years, and reported a decreased likelihood of impaction when teeth had been lost anterior to the third molar (9 per cent compared to 13 per cent), and also that any impaction was more likely to be unilateral in these cases. Faubion (1968) compared patients treated by the removal of four first premolar teeth with an equal number of patients treated without extraction. The extraction group were found to have almost four times the number of functional, fully erupted lower third molars. In 1975, Richardson reported the results of a longitudinal study of 160 patients. Of those cases which had lower first premolars removed, 28 per cent had impacted lower third molars in comparison with 34 per cent of the non-extraction cases. However, in a study including 45 cases where first premolar extraction had taken place and a further 47 cases with no extractions, Graber and Kaineg (1981) concluded that the incidence of third molar impaction is probably not reduced by extraction of the first premolars. Williams and Hosila (1976) compared the extraction patterns of 260 orthodontic patients and concluded that 'in this investigation, removal of four premolar teeth, either first or second premolars, did not have nearly so great an effect on the successful eruption of third molars as might have been

expected, particularly when compared with non extraction orthodontic treatment'.

These controversial research findings have led to considerable debate, particularly amongst oral surgeons, as to the most appropriate strategy for managing lower third molar teeth following orthodontic treatment (Shepherd and Brickley, 1994).

The current study was undertaken to determine the level of consensus between orthodontists regarding the need to refer post-orthodontic patients to an oral surgeon for third molar management, and to determine the degree to which orthodontists were consistent with one another regarding which cases should be referred for this procedure.

Methods

A series of 15 completed orthodontic cases were selected from those treated in the Orthodontic Clinic of the Dental Hospital in Cardiff between 1991 and 1994. The inclusion criteria for the study were that patients were aged at least 14 years, but under 18 years of age at the completion of their orthodontic treatment. Thirteen of the cases had four premolar extractions during their orthodontic care (five cases first premolars, two cases second premolars, and six cases upper first premolars and lower second premolars), while two cases had upper second molars only removed with no lower arch extractions. This was to ensure that the lower third molars were present and radiographically detectable. Case notes, photographic slides, radiographs, and models were collected and presented in booklet form to orthodontists with a cross section of experience and clinical practice (three senior academic staff, four senior hospital staff and three 'specialist' staff). However, a common factor was that, they all had completed their respective specialty training programmes. The booklet contained anterior-posterior and lateral colour photographs together with black and white photographs of study models at the beginning and end of treatment. For each case, copies of lateral cephalometric radiographs obtained at the beginning and at the end of treatment were also included, as was a pan-oral radiograph obtained at the end of treatment. An evaluation form was provided for each case.

Respondents were asked to consider which of the following management options they would select regarding each case's third molars:

(1) refer the patient for removal of the lower third molars;

(2) refer for a specialist oral surgery opinion regarding the third molars;

(3) do not refer the patient at all at this time.

The degree of agreement between observers was measured by kappa indices. Kappa indicates the degree of inter-observer agreement over and above that which would be expected by chance alone (Fleiss, 1981; Beck, 1985). Kappa values (using the appropriate methodology for multiple observers rating a case series) were calculated separately for the decision not to refer, the decision to refer for removal, the decision to refer for opinion and for the overall agreement between observers.

Kappa values between 0.40 and 0.75 may be taken to represent fair to good agreement on those cases that do require referral, whereas values below 0.40 represent poor agreement beyond that expected by chance (Landis & Koch, 1977).

The Kruskal-Wallis test was used to assess whether some observers were more inclined to suggest referral than others and whether some cases were referred more often than others.

Results

Overall, there was a high level of consensus not to refer cases, either for removal or treatment planning. Out of 300 treatment decisions, 263 were decisions not to refer (88 per cent), 10 were decisions to refer for removal (3%) and 27 were decisions to refer for opinion (9%). On average, each tooth was referred by a mean of 1.7 observers. Seventeen teeth would have been referred by no observer while a further seven teeth would have been referred by two or less observers. Ten observers judging 30 cases produced a total of 300 judgments (Table 1). To characterise the degree of agreement between any pairs of observers on which cases did not warrant referrals, a symmetrical agreement matrix was formed (Beck, 1985; Ismail *et al.*, 1989). This had a total of 2700 possible pairs ($30 \times 10 \times 9$). Table 1 shows that 263 out of 300 judgments (88 per cent) were decisions not to refer. Each of these were contrasted with each of the nine judgments by other observers on the same cases, resulting in a total of 2367 judgments (263×9). In 2134 instances the other observers also decided not to refer the same third molars, in 56 instances they decided to refer them for removal and in 177 instances they decided to refer them for opinion.

The agreement between observers is presented in the main diagonal of Table 1: $2134 + 8 + 40 = 2182$ out of 2700 paired judgments (81 per cent). However, some degree of

TABLE 1 Agreement between the 10 observers in judging need for referral of 30 third molars

	Number of individual observations	Do not refer	Refer for removal	Refer for opinion	Total	Kappa	Standard error of kappa	<i>z</i>
Do not refer	263	2134	56	177	2367	0.20	0.027	6.67
Refer for removal	10	56	8	26	90	0.06	0.027	2.22
Refer for opinion	27	177	26	40	243	0.08	0.027	2.96
Total	300	2367	90	243	2700	0.14	0.022	5.45

Bold type within the body of the table indicates the number of observer pairs in agreement.

agreement is to be expected by chance alone, therefore, chance agreement was incorporated into the assessment of inter-observer reliability. Minimum chance agreement occurs when all categories are used by all observers with equal frequency. Any deviation from a rectangular distribution of frequencies across categories will increase the agreement expected by chance and consequently reduce the index of agreement (Scott, 1955). In this study, chance agreement was estimated as 78 per cent, on the basis of the marginal frequencies. The kappa for multiple ratings was 0.14. Following the same procedure, kappa indices for the individual categories were also obtained and are presented in Table 1. The small size of kappas indicates that although the vast majority of observers agreed not to refer most of the cases, they did not agree on which cases to refer. This suggests that opinions regarding referral are personal rather than based on some underlying rationale.

To test the hypothesis that the decisions to refer or not are independent, so that the underlying kappa value is zero, the standard errors of the overall kappa and of the individual kappas were estimated (Fleiss *et al.*, 1979) and are presented in Table 1. The *z*-values were sufficiently large to reject the null hypothesis. Therefore, overall agreement and agreement on individual categories was significantly better than chance agreement, but the magnitude of kappas indicated that the extent of the agreement beyond chance was small.

The high chance agreement (0.78), due to the fact that observations tended to concentrate in the 'don't refer' category rather than spreading more evenly across categories, explains why the index of inter-observer agreement was so small. As Table 2 shows, the category 'referral for removal' was used only by three observers (30 per cent) and two of them used it only once or twice. The category 'refer for opinion' was used by five observers (50 per cent) with a frequency of 1–12 times out of 30 possible times.

As Table 2 shows, some observers were more prone to refer for removal or opinion than others. The Kruskal–Wallis test indicated highly significant variation amongst observers ($\chi^2 = 46.84$, $P < 0.001$) in their propensity to refer third molars. Observer No. 5 decided to refer for opinion 12 third molars (40 per cent) and observer No. 9 decided to refer for removal seven third molars (23.3 per cent). On the other hand, four observers (Nos 1, 2, 7 and 8) decided not to refer any third molar (100 per cent).

Table 3 shows that some cases were referred for removal or opinion more frequently than others. The Kruskal–Wallis test indicated highly significant within case variation ($\chi^2 = 85.57$, $P < 0.001$). The highest rate of decisions to refer for opinion or removal was found in case No. 10 for both left and right molars. On the other hand, eight cases (Nos 2, 3, 4, 5, 6, 8, 9 and 11) were not referred by any observer.

In 11 cases (73 per cent), a decision to refer or not applied to both third molars. This suggests that observers did not assess third molars individually, but in relation to the contralateral molars.

TABLE 3 Variation among cases in relation to the referral decision of the 10 observers. Values are numbers of observers

Case no.	Do not refer	Refer for removal	Refer for opinion	Total
1 left	8	0	2	10
1 right	8	0	2	10
2 left	10	0	0	10
2 right	10	0	0	10
3 left	10	0	0	10
3 right	10	0	0	10
4 left	10	0	0	10
4 right	10	0	0	10
5 left	10	0	0	10
5 right	10	0	0	10
6 left	10	0	0	10
6 right	10	0	0	10
7 left	8	0	2	10
7 right	10	0	0	10
8 left	10	0	0	10
8 right	10	0	0	10
9 left	10	0	0	10
9 right	10	0	0	10
10 left	4	2	4	10
10 right	4	3	3	10
11 left	10	0	0	10
11 right	10	0	0	10
12 left	7	1	2	10
12 right	7	1	2	10
13 left	8	1	1	10
13 right	6	1	3	10
14 left	9	0	1	10
14 right	9	0	1	10
15 left	7	0	3	10
15 right	8	1	1	10
Total	263	10	27	300

TABLE 2 Variation among observers in judging need for referral of 30 third molars. Values are numbers of observations

Observer no.	Do not refer	Refer for removal	Refer for opinion	Total
1	30	0	0	30
2	30	0	0	30
3	26	0	4	30
4	23	0	7	30
5	18	0	12	30
6	26	1	3	30
7	30	0	0	30
8	30	0	0	30
9	22	7	1	30
10	28	2	0	30
Total	263	10	27	300

Discussion

In the current study, all cases except one were referred by only a minority of orthodontists to their oral surgery colleagues, whether for removal or for an opinion. Indeed, in eight cases no orthodontist would have referred, while a further case would only have been referred by a single clinician. Overall, in 88 per cent of treatment decisions, a decision not to refer was made. This suggests that in general, orthodontists may not consider the removal of third molars to be an important adjunct to orthodontic treatment. This is an important finding, especially since one of the reasons given by surgeons for early removal of third molars is to avoid the risk of late lower incisor crowding. The low kappa values reported reflect the fact that there was little consensus between observers on which cases should be referred. This strongly suggests that these decisions were based on personal preference and not on a consistent set of criteria applied by orthodontists as a group.

These findings suggest that at least the orthodontists in this study do not consider that lower third molar surgery is appropriate for the majority of orthodontic cases. Therefore, they would not consider that this hypothetical risk of third molars contributing to post-treatment lower incisor crowding warrants surgical intervention with the associated risks and morbidity. Orthodontists are clearly in the best position to make such an assessment and as the findings of this study suggest, third molars should not be removed prophylactically in a misguided attempt to prevent crowding of the lower anterior teeth following orthodontic treatment, unless further evidence becomes available.

Conclusions

1. Orthodontists in this study did not, in general, consider referral of post-orthodontic patients for removal of lower third molars to prevent late incisor crowding, an appropriate strategy.
2. Considerable variations existed between orthodontists regarding both the number of cases they would refer and which should be referred.
3. Little consistency was observed between orthodontists as to whether a particular case warranted referral. This suggests that referral patterns are based on personal preference rather than some consistent referral criteria.

Acknowledgements

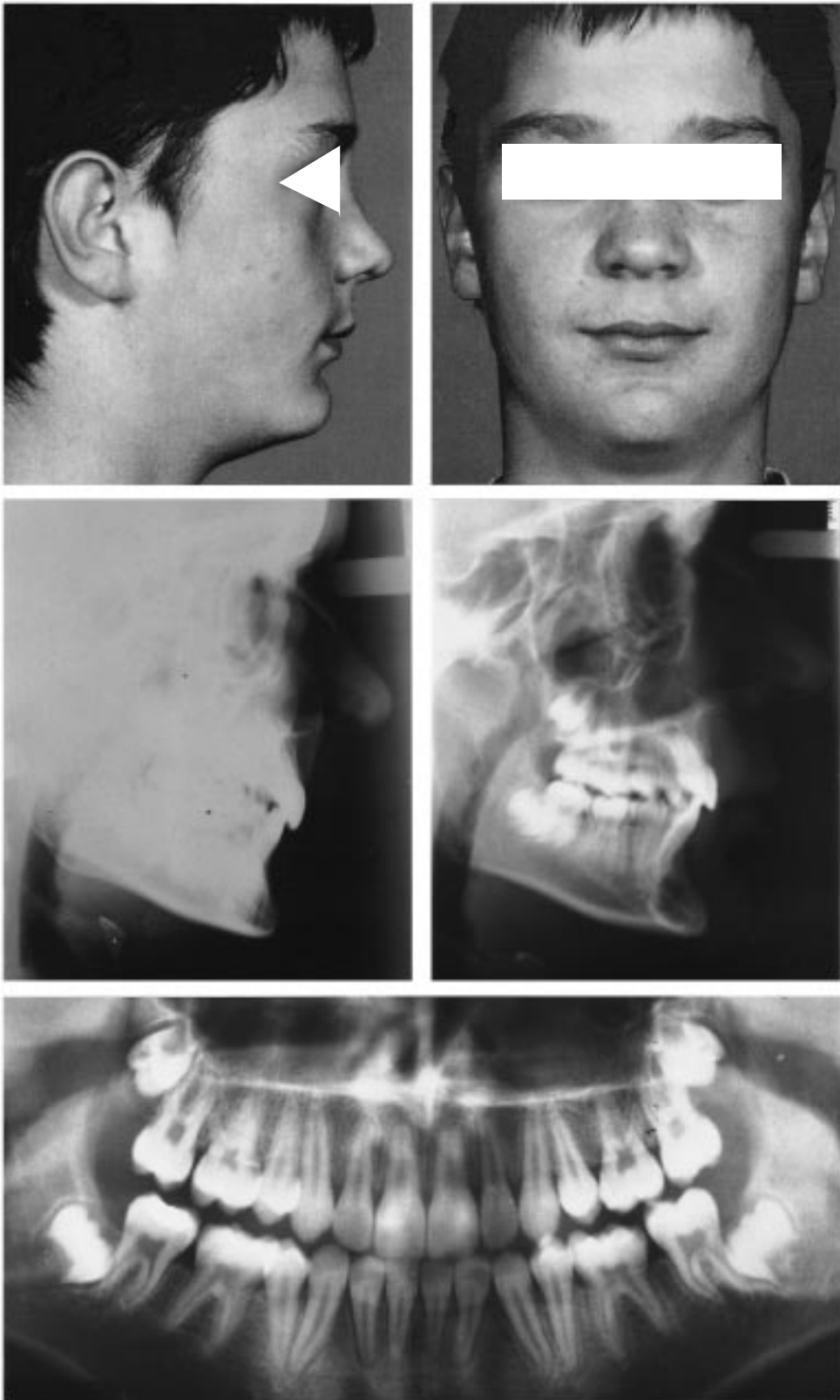
We would like to thank Dr Robert G. Newcombe of the Department of Medical Computing and Statistics, UWCM, for statistical advice.

References

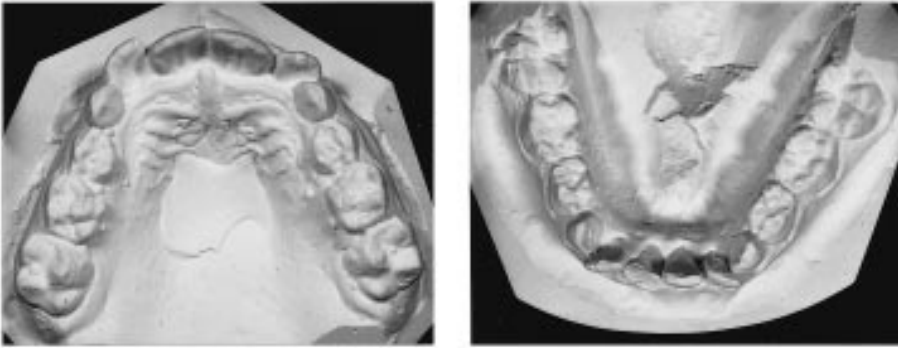
- Beck, J. S. (1985)**
Observer variability in reporting of breast lesions, *Journal of Clinical Pathology*, **38**, 1358–1365.
- Bjork, A., Jensen, E. and Palling, M. (1956)**
Mandibular growth and third molar impactions, *32nd congress of the European Orthodontic Society*.
- Broadbent, B. H. (1941)**
Ontogenic development of occlusion, *Angle Orthodontist*, **11**, 223–241.
- Cryer, B. S. (1967)**
Third molar eruption and the effect of extraction of adjacent teeth, *The Dental Practitioner*, **17**, 405–418.
- Faubion, B. H. (1968)**
Effect of extraction of premolars on eruption of mandibular third molars, *Journal of the American Dental Association*, **76**, 316–320.
- Fleiss, J. L. (1981)**
Statistical Methods for Rates and Proportions, 2nd edn, Wiley, New York, pp. 212–237.
- Fleiss, J. L., Nee, J. C. & Landis, J. R. (1979)**
The large sample variance of kappa in the case of different sets of raters, *Psychological Bulletin*, **86**, 974–977.
- Graber, T. M. and Kaineg, T. F. (1981)**
The mandibular third molar—its predictive status and role in lower incisor crowding, *Proceedings of the Finnish Dental Society*, **77**, 37–44.
- Guralnick, G. and Laskin, D. (1980)**
NIH Consensus Development Conference for removal of third molars, *Journal of Oral Surgery*, **38**, 235–236.
- Ismail, S. M., Colclough, A. B., Dinnen, J. S., Eakins, D., Evans, D. M. D., Grandwell, E., O'Sullivan, L. P., Summerell, J. M. & Newcombe, R. G. (1989)**
Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia, *British Medical Journal*, **298**, 707–710.
- Landis, J. R. & Koch, G. G. (1977)**
The measurement of observer agreement for categorical data, *Biometrics*, **33**, 159–174.
- Richardson, M. E. (1975)**
The relative effects of the extraction of various teeth on the development of mandibular third molars, *Transactions of the European Orthodontic Society*, **1975**, 79–85.
- Shepherd, J. P. and Brickley, M. R. (1994)**
Surgical removal of third molars (editorial), *British Medical Journal*, **309**, 620–621.
- Scott, W. A. (1955)**
Reliability of content analysis: the case of nominal scale coding, *Public Opinion Quarterly*, **19**, 321–325.
- Williams, R. and Hosila, F. (1976)**
Different extraction sites and incisor retraction, *American Journal of Orthodontics and Dental Orthopedics*, **69**, 388–410.
- Woodside, D. G. (1970)**
Round table: extra oral force, *Journal of Clinical Orthodontics*, **4**, 554–577.

Appendix

Example of records of one of the cases used in the study.



May 1985



May 1991

