Rating the preventive indication for mandibular third-molar surgery

The appropriateness of the visual analogue scale

Leif Lysell, Berndt Brehmer, Kerstin Knutsson and Madeleine Rohlin

Department of Oral Surgery, Central Hospital, Kristianstad; Department of Psychology, Uppsala University, Uppsala; and Department of Oral Radiology, Center for Oral Health Sciences, Lund University, Lund; Sweden

Lysell L, Brehmer B, Knutsson K, Rohlin M. Rating the preventive indication for mandibular third-molar surgery. The appropriateness of the visual analogue scale. Acta Odontol Scand 1995;53:60–64. Oslo. ISSN 0001–6357.

The aim was to determine whether a visual analogue scale (VAS) is an appropriate way for general dental practitioners and oral surgeons to rate their judgement of the indication for therapy of asymptomatic mandibular third molars. Thirty general dental practitioners and 10 oral surgeons had to judge the need for removal of 36 third molars. They were also asked to estimate the strength of the indication for extraction on a VAS. To assess the reliability of the judgement, the 36 cases were duplicated. For each participant, two mean indication indices with 95% confidence limits were calculated, one index for molars proposed to be extracted and one index for molars proposed not to be extracted. Pearson's correlation coefficient was used to estimate the intra-examiner reliability and the correlation between the number of molars proposed for extraction and the mean indication index. The correlation between the number of molars proposed to be extracted and the mean indication index was high ($p < 0.001$). The intra-examiner reliability was also high, with a mean correlation coefficient of 0.72 for the general practitioners and 0.84 for the oral surgeons. No single judge presented any overlap for the 95% confidence limits for the mean indication index of teeth proposed to be extracted versus teeth proposed not to be extracted. These results indicate that the VAS seems to be an appropriate method for analyzing the judgements on a therapeutic strategy like extraction versus no intervention for asymptomatic mandibular third molars. □ *Decision-making; oral surgery; tooth extraction; visual analogue scale*

Leif Lysell, Department of Oral Surgery, Central Hospital, S-291 85 Kristianstad, Sweden

Previous studies (1, 2) have indicated that about one-fourth of mandibular third-molar removals were on strictly preventive indications and that approximately half of the extracted third molars were asymptomatic and without any associated pathologic condition. Although the preventive indication is still proclaimed in modern textbooks of oral and maxillofacial surgery (3), the validity of this indication has been questioned (4–6). The current controversy is reflected in an equivocal attitude among clinicians towards removal of asymptomatic mandibular third molars. When we studied how general dental practitioners and oral surgeons judged the need for extraction of 36 asymptomatic mandibular third molars, the results showed a great diversity of opinion (7, 8). The number of mandibular third molars proposed to be removed ranged from 0 to 26 among the general dental practitioners and from 3 to 21 among the oral surgeons.

Although most decisions are ultimately dichotomous, the role of the dentist is to act as a quantitative judge. In diagnosis, prognosis, and treatment the judgements are often probabilistic. Therefore, we also asked each general dental practitioner and each oral surgeon to rate the need for extraction—that is, to express a degree of confidence in their judgement on each of the 36 cases. A tool commonly used for measuring subjective phenomena in clinical and research settings since the 1920s is the visual analogue scale (VAS). A review of the applications of VAS, its history, reliability, and validity, and a discussion on its strength and limitations was published by Wewers & Lowe in 1990 (9). In relation to third-molar surgery the VAS has been applied in studies on postoperative pain control (10–13) and patients' assessment of postoperative swelling and trismus (13–15).

In judgement analysis applied to medicine, a 100-mm VAS was used to study diagnostic strategies when rheumatologists were asked to assess current disease activity in rheumatoid arthritis (16). With regard to therapeutic decisions, physicians' judgement concerning estrogen replacement therapy for menopausal women (17) and the decision whether or not to use tube feeding in seriously ill patients (18) were analyzed with similar scales. As far as we know, the VAS has so far not been used to rate either diagnostic or therapeutic decisions in dentistry.

The aim of this investigation was to determine whether a VAS is an appropriate way to rate the judgement of extraction or non-extraction of a sample of asymptomatic mandibular third molars. Therefore, dentists were asked to establish the degree of confidence with which they judged the need for the extraction (an indication index). If the VAS can be considered appropriate for rating an indication index for mandibular third-molar surgery, the following assumptions can be made:

1. Judges more prone to suggest preventive extractions will give a higher indication index than those with an ambiguous or negative attitude towards extraction.
2. There should be a high reliability when duplicate cases are rated.
3. The indication index proposed for extraction cases should differ clearly from ratings proposed for non-extraction cases.

Materials and methods

Materials

Thirty-six asymptomatic mandibular third molars (M3s) were selected. The cases were distributed equally among men and women, age groups, angular positions, and degrees of impaction. Thus, half of the subjects were women. Three age-groups were selected: 19–25 years, 26–40 years, and 41–60 years. The position was classified as vertical, mesioangular, distoangular, or horizontal, in accordance with Winter (19). The degree of impaction consisted of one of three categories: partially covered by soft tissue, completely covered by soft tissue, and completely covered by bone tissue.

Participants

Thirty general dental practitioners (GDPs) and 10 oral surgeons were asked to participate in this study. The selection of GDPs was presented by Knutsson et al. (7). The length of their professional experience varied from 1 to 30 years, with a mean of 12 years. The selection of oral surgeons was presented in another study (8). They had been certified for from 3 to 30 years, with a mean of 17 years, and had from 2 to 14 years' experience as GDPs before their oral surgery training.

Case scenario

Each of the cases was presented to the 40 participants as shown in Fig. 1, with a radiograph and a brief text informing them about the gender and age of the patient and the degree of impaction of the M3. The participants were also informed that the M3s were asymptomatic. To assess the intra-examiner reliability of the judgements, the 36 cases were duplicated. When re-presented, they were reversed so that they appeared to come from the contalateral side. The 72 cases were presented in random order in relation to age group, angular position, degree of impaction, and original versus duplicate case.

As presented in Fig. 1, the participants were asked to judge whether to extract the M3. Then they were also asked to establish the degree of confidence with which they judged the need for extraction (indication index) on a 100-mm VAS that ranged from 'very weak' (0 mm) to 'very strong' (100 mm).

CASE 19.
Woman, 29 years.
The lower third molar completely covered by bone tissue.

A. Do you consider that this molar should be removed? ☐ YES ☐ NO

B. How strong is the indication for removing this third molar?

Very weak _____ Very strong

Fig. 1. One of the 36 cases of asymptomatic mandibular third molars and the questions presented to the participants. A radiograph showed the angular position, and a brief text gave the gender and age of the patient and the degree of impaction of the molar.

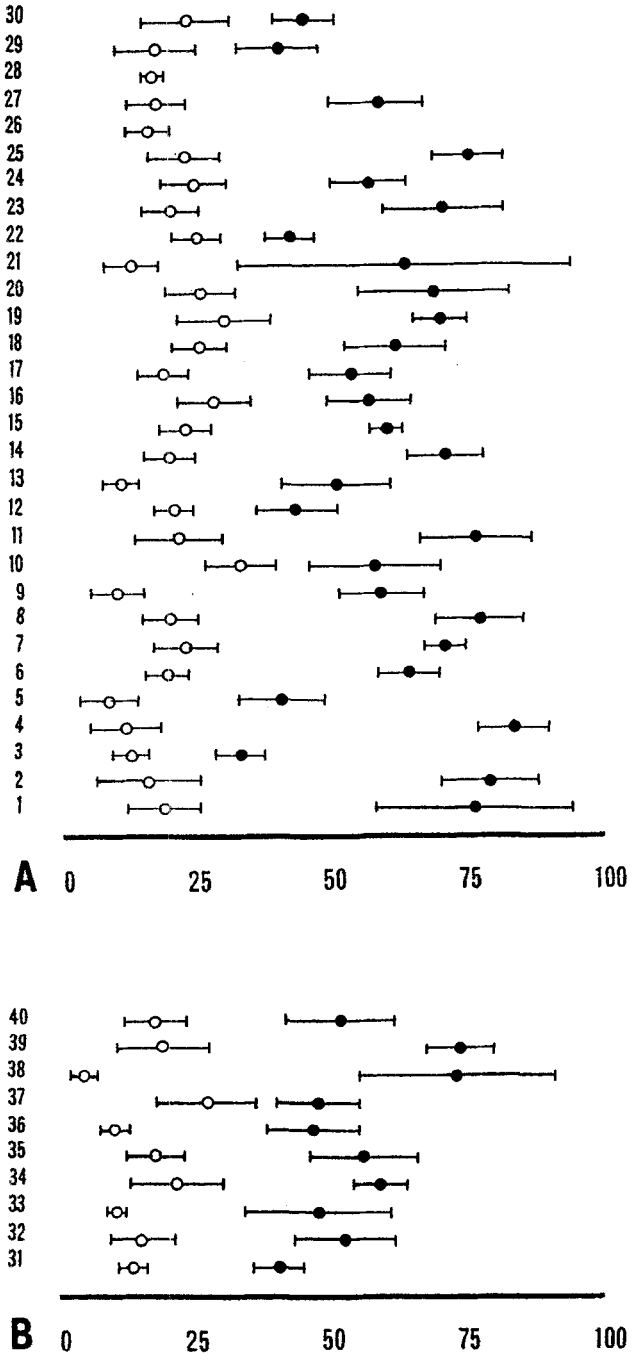

Fig. 2. Mean indication index with 95% confidence limits for asymptomatic mandibular third molars proposed not to be extracted (—○—) and to be extracted (—●—) by each participant of (A) 1–30 general dental practitioners and (B) 31–40 oral surgeons.

Analysis

The data were coded by measuring the response on each VAS to the nearest millimeter. This measure of the degree of confidence the participants had in their judgement on the need for removal of each M3 was called indication index for extraction. For each participant two mean indication indices with 95% confidence limits were calculated, one index for M3s proposed to be extracted and one index for M3s proposed not to be extracted. The analysis was performed on the original 36 cases.

The intra-examiner reliability and the correlation between the number of M3s proposed for extraction and the mean indication index was estimated with Pearson's correlation coefficient.

Results

The correlation between the number of M3s proposed for extraction and the indication index was high, with a correlation coefficient for GDPs of 0.70 ($p < 0.001$) and for oral surgeons of 0.93 ($p < 0.001$). The intra-examiner reliability was also high, with a mean correlation coefficient of 0.72 for GDPs and of 0.84 for oral surgeons.

Fig. 2 shows the mean indication indices with 95% confidence limits for each GDP and for each oral surgeon. No single participant presented any overlap for the confidence limits of teeth proposed to be extracted versus teeth not to be extracted. For all participants, 0–35 appeared to be a non-extraction interval and above 35 to be an extraction interval. At 35 only 20% of the participants presented a confidence limit for either decision crossing over to the alternative therapy.

For the GDPs and the oral surgeons as groups, the indication index was 61 and 57, respectively, for M3s proposed for extraction. The corresponding figures for M3s proposed not to be extracted were 20 and 15, respectively. There was a marked interindividual variation of the mean indication index among both GDPs and oral surgeons. The mean indication index for M3s proposed for extraction ranged from 33 to 83 for the GDPs and from 40 to 73 for the oral surgeons. For M3s proposed not to be extracted the corresponding figures ranged from 8 to 28 for the GDPs and from 4 to 27 for the oral surgeons.

Discussion

The results in our study support the appropriateness of using the VAS to analyze judgements on a therapeutic strategy like intervention versus no intervention in the case of asymptomatic M3s. First, there was an equally high correlation between the number of cases proposed for extraction and the mean indication index of the individual judges. Thus, the participants being more prone to suggest preventive M3 surgery gave a higher indication index than participants with a more reluctant attitude. We are well aware that these results could be expected, since the judges had access to their yes/no

responses when rating the indication index. Second, however, no participant presented any overlap between the confidence limits for M3s proposed to be extracted compared with M3s proposed not to be extracted. This result supports the validity of the VAS as applied in our study. There seemed to be transition points between indices for the decision to extract and not to extract for both groups of participants. The results in our study suggest the following rule: extract if the index is above 40; do not extract if the index is below 30; and treat the index of 30–40 as equivocal. The participants evidently judged the value of preventive surgery to outweigh the risks of surgical complications or withholding the treatment at an index above 40. The decision of whether to extract the asymptomatic M3 seems to be a threshold phenomenon. We interpret the results of this study as indicating that the participants perceived the 100-mm line as a continuous unit and were able to rate a binary decision within this unit. This approach has points of similarity with the theory suggested by Kvist et al. (20) in the field of endodontic treatment. They found that clinicians seem to regard different periapical conditions as different stages on a continuous axis and that the interindividual variation in decision-making may be considered the result of choosing different cut-off points on this continuum for prescription of retreatment.

The reliability of the VAS has usually been evaluated by test–retest methods. This approach has been questioned for dynamic phenomena such as pain and mood (9). Such subjective conditions change over time; for example, postoperative pain rated 3 h after surgery cannot, owing to the dynamic character of postoperative pain, be retested the next day. On the other hand, a test–retest evaluation after a few minutes up to a few hours may be biased by a clear memory of the first marking (21). For reliability assessment of 'paper' cases like ours, a test–retest approach using duplicate cases seems appropriate and has been used by other researchers (16, 17, 22, 23) and has also been suggested by Stewart (24) and Brehmer (25) to be an important aspect of a judgement analysis. Very high test–retest correlations ($r > 0.95$) suggest that the cases are few and easy to remember or that the judgements are easily predictable. In our study the correlations between duplicate cases were 0.72 for GDPs and 0.84 for oral surgeons. A similar correlation between duplicate cases (0.81) was reported by Holzman et al. (17) when physicians rated the probability of prescribing estrogen therapy for menopausal women. Comparison with studies using binary judgements are not valid, as they often present their reliability tests as percentage of yes/no responses. We interpret our duplicate correlations as acceptably high, with the 36 cases all duplicated as not being too few and the task not too predictable.

The mean indication index of GDPs as a group was similar to that of the oral surgeons for both the extraction and the non-extraction cases. As the training and practice situation of oral surgeons differs from that of GDPs, it might have been anticipated that they would have different thresholds. This was found for a surgical panel, which rated more indications appropriate for carotid surgery than did a multi-speciality panel (26). The authors expected this 'user bias', as 'doers' are thought to be more committed to the value of the service they provide and would be expected to focus more on its strengths than on its limitations. We do not know whether this similarity between GDPs and oral surgeons is restricted to preventive third-molar surgery. Perhaps the bias of the doers would be more pronounced if the judgement were in the fields of temporomandibular joint or orthognathic surgery.

When it comes to the decision on a single M3, there was a marked interindividual variation among the participants. The wide range of indication indices for the extraction of asymptomatic M3s also indicates that both the GDPs and the oral surgeons differed from individual to individual in their judgement. Such a variation might be one of the explanations for the regional variations, ranging from 72 to 521 M3 operations per 100,000 inhabitants in Sweden (27), as the treatment strategy is probably influenced by a few oral surgeons in the different counties. Uncertainty about outcomes is an unavoidable part of decision-making in dentistry, particularly when scientific data are lacking, such as for the prediction of the outcome of treatment versus no treatment. One might agree with Eddy (28) that 'clinical research is extremely inefficient. Some questions are dissected beyond recognition; others are virtually ignored'. Several studies on the complications after removal of M3s have been performed; for a review see, for example, Kugelberg (29). However, systematic follow-up studies on the complications of *no* treatment of asymptomatic M3s are virtually lacking. The increase in preventive M3 surgery during the past 30 years took place without any decision analysis or cost-effectiveness analyses. Not until recently have cost-effectiveness analyses been performed by Tulloch et al. (6), demonstrating that the more conservative approach of waiting and extracting only M3s that develop pathologic conditions minimizes both morbidity and costs on a population basis.

References

1. Nordenram A, Hultin M, Kjellman O, Ramström G. Indications for surgical removal of the mandibular third molar. Swed Dent J 1987;11:23–9.
2. Lysell L, Rohlin M. A study of indications used for removal of the mandibular third molar. Int J Oral Maxillofac Surg 1988; 17:161–4.
3. Laskin DM. Oral and maxillofacial surgery. Vol 2. St Louis (MO): CV Mosby, 1985:57.
4. Friedman JW. Containing the cost of third-molar extractions: a dilemma for health insurance. Public Health Rep 1983;98:376–84.
5. Stephens HR, Kogon SL, Reid JA. The unerupted or impacted third molar—a critical appraisal of its pathologic potential. J Can Dent Assoc 1989;55:201–7.

6. Tulloch JFC, Antczak-Bouckoms AA, Ung N. Evaluation of the costs and relative effectiveness of alternative strategies for the removal of mandibular third molars. Int J Technol Assess Health Care 1990;6:505–15.
7. Knutsson K, Brehmer B, Lysell L, Rohlin M. General dental practitioners' evaluation of the need for extraction of asymptomatic mandibular third molars. Community Dent Oral Epidemiol 1992;20:347–50.
8. Knutsson K, Brehmer B, Lysell L, Rohlin M. Asymptomatic mandibular third molars: oral surgeons' judgment of the need for extraction. J Oral Maxillofac Surg 1992;50:329–33.
9. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990;13:227–36.
10. Quiding H, Persson G, Ahlström U, Hellem S, Johansson G, Jönsson E, et al. Paracetamol plus supplementary doses of codeine. An analgesic study of repeated doses. Eur J Clin Pharmacol 1982;23:315–9.
11. Dahl E, Feldmann G, Jönsson E. Acetylsalicylic acid compared with acetylsalicylic acid plus codeine as postoperative analgesics after removal of impacted third molars. Swed Dent J 1985;9:207–12.
12. Habib S, Matthews RW, Scully C, Levers BGH, Shepherd JP. A study of the comparative efficacy of four common analgesics in the control of postsurgical dental pain. Oral Surg Oral Med Oral Pathol 1990;70:559–63.
13. Lysell L, Anzén B. Pain control after third molar surgery—a comparative study of Ibuprofen (Ibumetin) and a paracetamol/codeine combination (Citodon). Swed Dent J 1992;16:151–60.
14. Henrikson PÅ, Thilander H, Wåhlander LÅ. Voltaren as an analgesic after surgical removal of a lower wisdom tooth. Int J Oral Surg 1985;14:333–8.
15. Berge TI. Visual analogue scale assessment of postoperative swelling. A study of clinical inflammatory variables subsequent to third-molar surgery. Acta Odontol Scand 1988;46:233–40.
16. Kirwan JR, Chaput de Saintonge DM, Joyce CRB, Currey HLF. Clinical judgment in rheumatoid arthritis. II. Judging current disease activity in clinical practice. Ann Rheum Dis 1983;42:648–51.
17. Holzman GB, Ravitch MM, Metheny W, Rothert ML, Holmes M, Hoppe RB. Physicians judgments about estrogen replacement therapy for menopausal women. Obstet Gynecol 1984;63:303–11.
18. Smith DG, Wigton RS. Modeling decisions to use tube feeding in seriously ill patients. Arch Intern Med 1987;147:1242–5.
19. Winter GB. Principles of exodontia as applied to the impacted third molar. St Louis (MO): American Medical Books, 1926.
20. Kvist T, Reit C, Esposito M, Mileman P, Bianchi S, Pettersson K, et al. Prescribing endodontic retreatment: towards a theory of dentist behavior. Int Endodont J 1994;27:285–90.
21. Revill SI, Robinson JO, Rosen M, Hogg MIJ. The reliability of a linear analogue for evaluating pain. Anaesthesia 1976;31:1191–8.
22. Gillis JS, Lipkin JO, Moran TJ. Drug therapy decisions. A social judgment analysis. J Nerv Ment Dis 1981;169:439–47.
23. Rutkow IM. Surgical decision making. The reproducibility of clinical judgment. Arch Surg 1982;117:337–40.
24. Stewart TR. Judgment analysis: procedures. In: Brehmer B, Joyce CRB, editors. Human judgment. The SJT view. Amsterdam: Elsevier Science Publishers BV (North Holland), 1988:41–74.
25. Brehmer B. Models of diagnostic judgments. In: Rasmussen J, Duncan K, Leplat J, editors. New technology and human error. Chichester: John Wiley & Sons Ltd, 1987:87–95.
26. Leape LL, Park RE, Kahan JP, Brook RH. Group judgments of appropriateness: the effect of panel composition. Qual Assur Health Care 1992;4:151–9.
27. Lysell L. Visdomstandsoperationer vid landstingens oralkirurgiska kliniker. Tandlakartidningen 1993;85:108–13 (in Swedish).
28. Eddy DM. Clinical decision making. From theory to practice: the challenge. JAMA 1990;263:287–90.
29. Kugelberg CF. Third molar surgery. Curr Opin Dent 1992;2:9–16.

Received for publication 22 March 1994
Accepted 12 August 1994