

## TECHNICAL NOTE

Iain A. Pretty,<sup>1</sup> *B.D.S.(Hons), M.Sc., Ph.D., M.F.D.S.R.C.S.(Ed)*

# Development and Validation of a Human Bitemark Severity and Significance Scale

**ABSTRACT:** Numerous efforts have been made to develop a consistent manner to describe bite injuries. Some have been related to the type of injury, others to the manner in which it was caused or simply its anatomical location. Bitemark severity is related to forensic significance and hence the ability to use a common means of injury description would be of benefit to odontologists and those who commission their services. A novel index, relating severity to forensic significance, was developed. A text version and accompanying visual index were produced and distributed (via the web) to three groups: odontologists, forensic pathologists, and police officers. A total of 35 bitemarks were assessed and rated using the new index. Weighted  $\kappa$  analyses were used to determine the agreement data both between and within groups and individuals.  $\kappa$  demonstrated a high level of intraoperator and interoperator reliability, particularly in the police officer group. The index shows promise as a universal means of describing bite injuries between professionals concerned with their detection and analysis.

**KEYWORDS:** forensic science, bitemarks, index, scale, reliability

Numerous methods for describing bitemarks have been suggested (1–3). Common methods include the use of descriptors for the severity of the injury, the location of the injury, or the presence of gross, class, or unique characteristics (4). However, there is no universally accepted means of describing bite injuries and hence communication between professionals dealing with such injuries is complicated. A review of the literature found no studies that examined the use of such an index, or any attempts to characterize any bitemark descriptors by means of reliability and validity testing (4). In 1980s a group at the Northwestern Dental School led by Smith developed a scale of bitemark severity. This was aimed at enabling forensic dentists to communicate with each other on the nature of an injury. There was no link within the scale to comment on significance and the index was never published.

There is a clear link between the severity of a bite injury at presentation and its forensic significance (5). For example, a bite injury that presents as a diffuse, nondiscrete bruise is unlikely to possess unique characteristics suitable for analysis resulting in the positive identification of the perpetrator. However, on the other end of the severity spectrum, very aggressive, avulsive injuries are frequently poor candidates for analysis (5). A combination of factors including the loss of tissue, tearing, and distortion of wound margins and the need for urgent medical treatment generally render such injuries poor candidates for analysis. Bite injuries that present in the middle of these extremes, i.e., injuries made up of discrete, individual bruises, small abrasions, and lacerations frequently and considered by odontologists to present the highest level of significance will enable the exclusion and inclusion of potential suspects.

The forensic significance of bite injuries is intimately related to an increasing number of class and unique characteristics of teeth that can be observed, measured and compared. Injuries limited to only gross or a few class characteristics can be regarded as of low forensic significance and those with numerous unique characteristics of high significance.

Bitemark injuries are not uncommon (6). The author receives over 60 requests per year to examine injuries on deceased and living individuals that investigators believe may be bitemarks. However, despite this high number of cases, only an average of 15 are suitable for further work and 10 that hold sufficient unique details for a bitemark analysis to be undertaken. However, the workload and expense involved in the assessment of these injuries could be reduced if an effective means of characterizing these injuries could be performed by either pathologists, police officers, coroner's officers, or social workers and others requesting odontological opinions.

An ideal bitemark index would be as follows:

- (a) be easy and quick to use,
- (b) be reproducible,
- (c) provide information on appearance and likely significance,
- (d) be used on living and deceased individuals,
- (e) not be reliant on subjective measures, such as color or texture,
- (f) easy to communicate between professionals,
- (g) multicultural and universally applicable,
- (h) integral scale to allow for possible statistical analysis.

Many indices within the medical field are both text and visually based. A good example of this in the Index of Orthodontic Treatment Need (IOTN), which has been successfully used by orthodontists for many years to describe the nature of an individual's dental appearance and relate this to the treatment required (7). The index was developed and subsequent research assessed its validity and reliability in the hands of a number of different

<sup>1</sup>The Dental School, Manchester M15 6SH, U.K.

Received 10 June 2006; and in revised form 10 Oct. 2006; accepted 29 Oct. 2006; published 5 April 2007.

professional groups (8,9). The index is now universally established and an IOTN number will accompany every orthodontic assessment.

A further advantage of a standardized bitemark scale is that of research consistency. A number of studies have examined the ability of forensic dentists to assess bitemarks and then link them to a suspect (10,11). However, it is impossible to compare these studies as there is little indication within the reports as to the severity or forensic significance of the injuries under study. The ability to utilize a common scale of reference for bite injuries will permit the comparison of future bitemark studies and enable a meta-analysis of their results (12,13).

The purpose of the current study was to assess the reliability of a novel bitemark index that would provide information on the severity and forensic significance of a bite injury.

## Materials and Methods

The proposed index is shown in Fig. 1, with the allied visual component shown in Fig. 2. A selection of 35 bitemark photographs were obtained from the case collection of the author. Cases were selected to represent a spread of bitemark presentations and image quality. An example is shown in Fig. 3. A total of 10 individuals were recruited from each of the following groups:

- (a) Forensic odontologists—via a website;
- (b) Police officers—via 2-week residential courses; and
- (c) Forensic pathologists—via a website.

A minimum level of experience was required for the odontologists (>3 bitemark cases), pathologists were board certified, and the police officers were all attending a forensic refresher course, a prerequisite of which was at least 7 years in post as a crime scene examiner. Each individual was provided with a copy of the text and visual scale and was then asked to grade a total of 35 bitemark images. Each image was a professionally developed photograph,

sized to 8 × 11 inches. Following a minimum washout period of 24 h, the examiners were asked to repeat the exercise. An area was made available at the end of the data-entry sheet to enable comments on the scale to be made should the examiners wish to do so.

The data were collected anonymously, save only for an indicator of the occupation group. Data were entered into SPSS (Version 11, SPSS Inc., Chicago, IL) and a range of statistical tests were undertaken. Weighted  $\kappa$  scores were applied between each examiner's first and second attempt (intraexaminer reliability), between each of the examiners within a group (interexaminer reliability), and between the mean score for each image from each occupation group (intraexaminer reliability). Analysis of variance was undertaken between each of the three groups to determine whether any statistical differences existed between the scores. A frequency distribution of the scores obtained was produced to determine whether the image set represented a fair spread of bitemark severity and significance. As the data were not normally distributed, Spearman's rank correlations between each of the groups in terms of mean score for each image were also performed. Comments at the end of the data sheet were analyzed qualitatively.

## Results

Ten odontologists, average years of experience 16.7 years ( $\pm 7$ ), 10 pathologists, average years of experience 18.1 years ( $\pm 8$ ), and 10 scenes of crime examiners (police) with 14.6 years of experience ( $\pm 9$ ) correctly completed the exercise.

The mean weighted  $\kappa$  scores for the comparison between the first and second attempt and the gold standard are shown in Fig. 4. The police and the pathologist examiners scored significantly higher ( $p < 0.05$ ) than the odontologists, although the agreement for all three groups was rated as substantial. Weighted  $\kappa$  scores for intraexaminer agreements are shown in Fig. 5, with the police scoring significantly better than the odontologists and pathologists, and the pathologists scoring significantly better than the

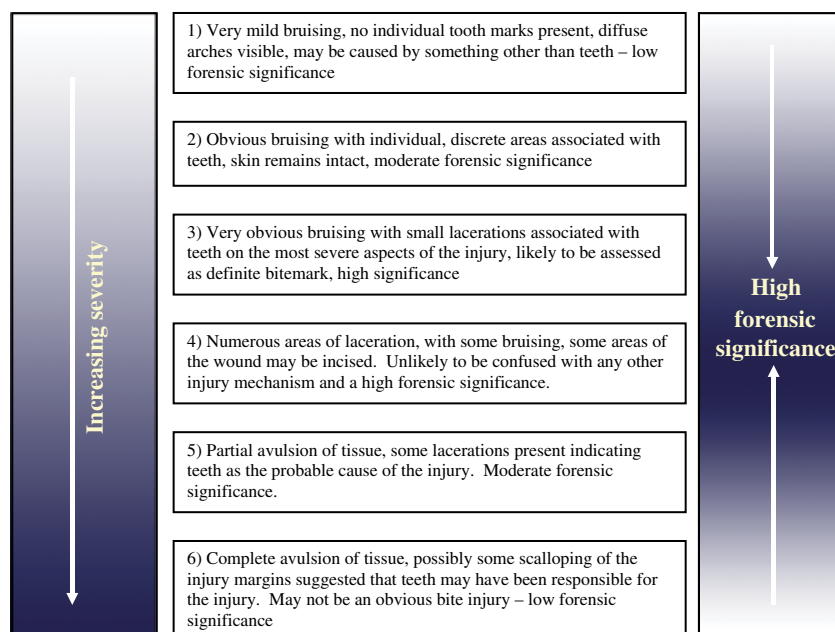


FIG. 1—Range of bitemark severity—the bitemark severity and significance scale.

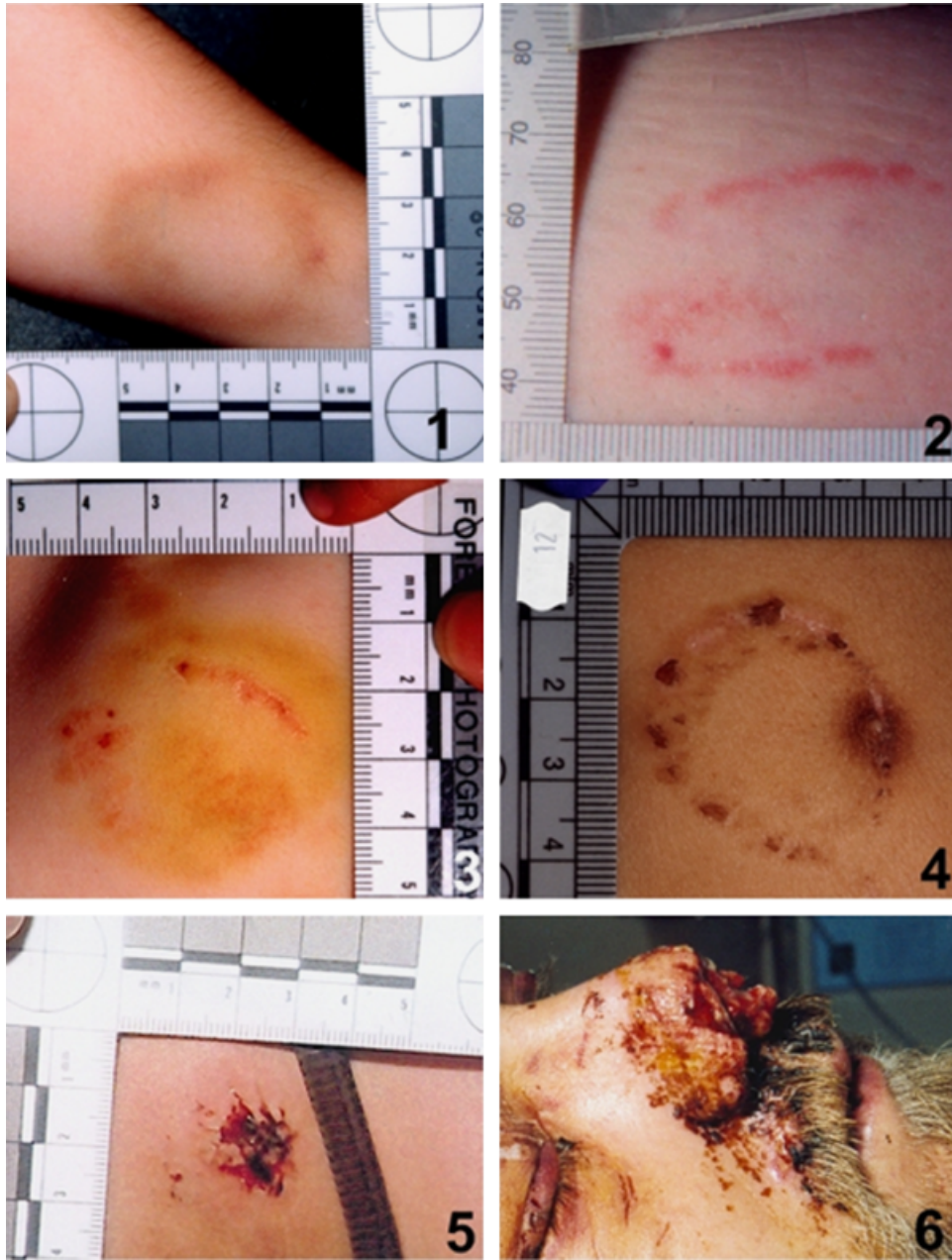


FIG. 2—Visual index of the bitemark severity and significance scale.

odontologists ( $p < 0.05$ ). Figure 6 presents the qualitative assessment of the scale where each of the three groups was questioned as to whether or not then would use the scale again, whether they found the scale easy to use, whether the scale was portable enough, and their confidence in using it correctly.

Spearman's rank correlations for the mean scores for each image within each group were determined with the results shown in Table 1.

## Discussion

The results demonstrate that, even at the lowest  $\kappa$  levels, the index demonstrates a substantial degree of reliability (14,15). The observation that the odontologists scored the lowest in each of the assessments may be an indication of past experience; i.e., they were not simply looking at the sample bitemarks and

comparing with the scale, but applying their own experience as to whether or not an injury contained sufficient forensic significance. It is interesting to note that in other assessments of bitemarks where nonodontologists have taken part, they will often score higher, or similar, to the experienced bitemark examiners (11).

Another aspect of the study that must be considered is the manner in which the index, and the sample bitemarks, was provided to the examiners. Both the odontologists and the pathologists accessed all the materials online via the study website. The police were each issued with a paper version of the scale, double sided (one with the pictorial index, the other side with text), which had been professionally printed onto a card and laminated. This is the intended mechanism of distribution for the scale and each officer was able to retain the index for their personal use. The police officers also viewed their sample bitemarks as photographic re-

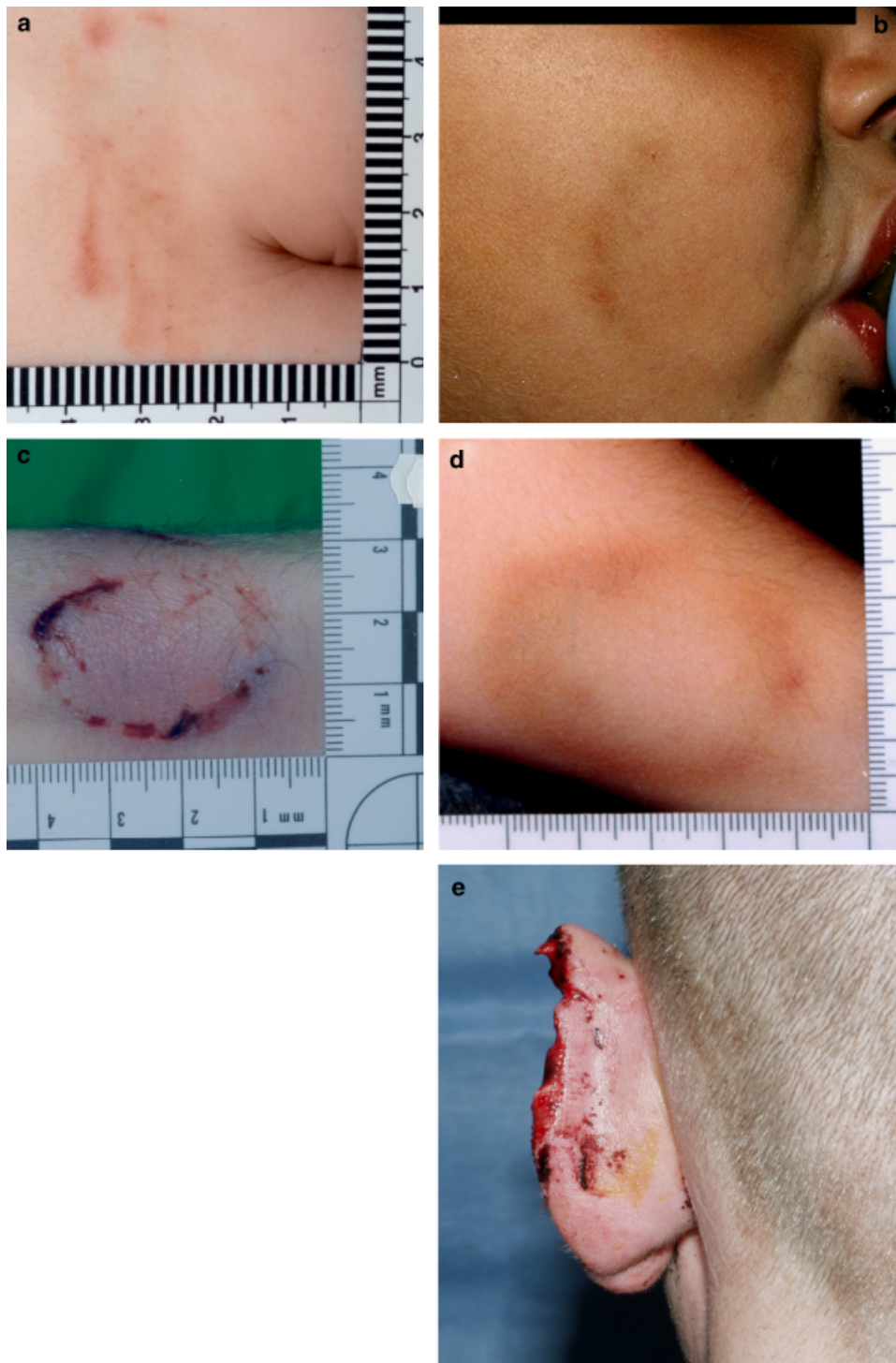


FIG. 3—Example of images provided to the examiners to rate.

productions rather than on-screen as did the pathologists and odontologists. This difference prevents direct extrapolation between the police officer group and the other examiners, but serves a useful purpose within this study as an example of use in practice. One should therefore be cautious in assuming that scenes of crime officers are superior at rating bite marks; these data should be interpreted as the police officers are at least as good as the other groups but had access to superior quality materials during their assessment.

The bite mark scale described within this study has the potential to enable forensic professionals to discuss the severity and significance of injuries before initiating complex and expensive evidence-collection procedures and may also enable future bite mark research studies to be compared. The success of any index is, however, based on its uptake by the profession. Further research is required to determine whether, once properly distributed, this index is correctly used by the forensic practitioners for whom it was designed.

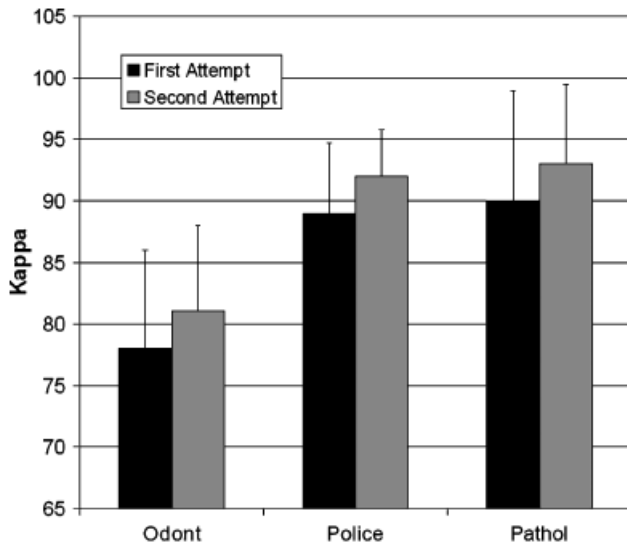
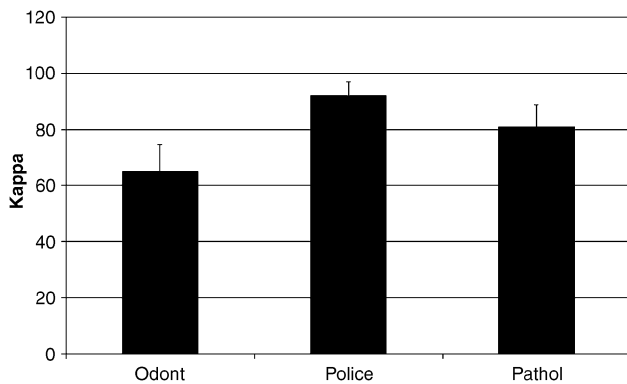
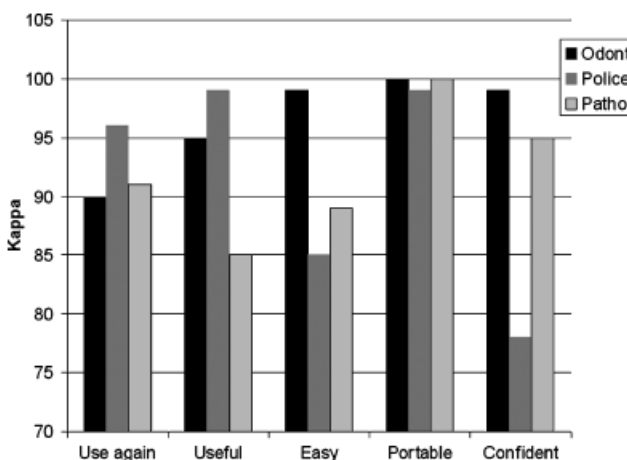

 FIG. 4—Weighted  $\kappa$  scores compared with the gold standard.

 FIG. 5— $\kappa$  comparisons between the first and second attempts (intra-examiner agreement).


FIG. 6—Qualitative assessment of the index by each potential user group.

TABLE 1—Spearman rank correlations between mean scores and the first attempt (nonshaded) and the second attempt (shaded) for each group.

	Group		
	Odontologists	Police	Pathologists
Odontologists		0.85	0.74
Police	0.83		0.92
Pathologists	0.85	0.96	

## References

1. Atsu SS, Gokdemir K, Kedici PS, Ikyaz YY. Bitemarks in forensic odontology. *J Forensic Odontostomatol* 1998;16(2):30–4.
2. Wagner GN. Bitemark identification in child abuse cases. *Pediatr Dent* 1986;8(1 Spec. No.):96–100.
3. Hale A. The admissibility of bitemark evidence. *Southern Californian Law Rev.* 1978;51(3):309–34.
4. Sweet D, Pretty IA. A look at forensic dentistry—Part 2: teeth as weapons of violence—identification of bitemark perpetrators. *Br Dent J* 2001;190(8):415–8.
5. Pretty IA. The barriers to achieving an evidence base for bitemark analysis. *Forensic Sci Int* 2006;159(Suppl 1):S110–20.
6. Pretty IA, Sweet D. Anatomical location of bitemarks and associated findings in 101 cases from the United States. *J Forensic Sci* 2000;45(4):812–4.
7. Mandall NA, Wright J, Conboy F, Kay E, Harvey L, O'Brien KD. Index of orthodontic treatment need as a predictor of orthodontic treatment uptake. *Am J Orthod Dentofacial Orthop* 2005;128(6):703–7.
8. Louwerse TJ, Aartman IH, Kramer GJ, Pahl-Andersen B. The reliability and validity of the index of complexity, outcome and need for determining treatment need in Dutch orthodontic practice. *Eur J Orthod* 2006;28(1):58–64.
9. Hlongwa P, Beane RA, Seedat AK, Owen CP. Orthodontic treatment needs: comparison of two indices. *SADJ* 2004;59(10):421–4.
10. Arheart KL, Pretty IA. Results of the 4th ABFO Bitemark Workshop—1999. *Forensic Sci Int* 2001;124(2–3):104–11.
11. Pretty IA, Sweet D. Digital bitemark overlays—An analysis of effectiveness. *J Forensic Sci* 2001;46(6):1385–91.
12. Mullen PD, Ramirez G. The promise and pitfalls of systematic reviews. *Annu Rev Public Health* 2006;27:81–102.
13. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care* 2005;9(5):R575–82.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–74.
15. Koch GG, Landis JR, Freeman JL, Freeman DH Jr., Lehnen RC. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics* 1977;33(1):133–58.

Additional information and reprint requests:

Iain A. Pretty, B.D.S., M.Sc., Ph.D., M.F.D.S.R.C.S.

Dental Health Unit

3A Skelton House, Lloyd Street North

Manchester Science Park

Manchester M15 6SH

U.K.

E-mail: iain.pretty@manchester.ac.uk