A comparison of the ability of experts and non-experts to differentiate between adult and child human bite marks using receiver operating characteristic (ROC) analysis

D.K. Whittaker¹*, M.R. Brickley¹, L. Evans¹

¹Dept. of Basic Dental Science, Dental School, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XY, UK
²Clinical Decision Research Group, Dept. of Oral Surgery, Medicine and Pathology, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XY, UK

Received 21 January 1997; accepted 9 December 1997

Abstract

Fifty colour prints of human bite marks were sent to 109 observers who were asked to decide using a six point rating scale, whether the marks had been produced by the teeth of an adult or a child. The observers consisted of accredited senior forensic dentists, accredited junior forensic dentists, general dental practitioners, final year dental students, police officers and social workers. The results were compared against a "gold standard" which was the actual verdict from the case. Comparison of the results between the groups of observers and the standard was made using Receiver Operating Characteristics (ROC) methodology. The best decisions were made by senior/junior experts or final year dental students. General dental practitioners and police officers were least able to differentiate correctly between adult and child bite marks. The effect of training is important and its effects need to be assessed in more detail in future studies. © 1998 Elsevier Science Ireland Ltd.

Keywords: Bite marks; Adult; Child; Forensic dentistry

1. Introduction

The term "bite mark" is used rather loosely to describe a mark caused by the teeth
alone, or the teeth in combination with other mouth parts. Bite marks can be found in flesh, foodstuffs and less frequently in a variety of other materials [1].

The procedure for comparison of bite marks on the skin of victims to the dentition of possible assailants is well established [2] and has been reported in historical cases [3,4], and in many cases to the present day. Since 1966 there have been more than 100 papers written on the subject, the majority of which are case reports or descriptions of technique. Bite mark evidence has recently become increasingly important in the investigation of non-accidental injuries to children [5].

Some well published court proceedings have highlighted disagreement between opposing experts not only as to methodology but in the fundamental question as to whether a defendant was responsible for the bite. Those with long experience in the courts will testify that dissent amongst so-called bite mark experts is commonplace [6]. Some odontologists are of the opinion that bite mark analysis should never be used to convict a suspect but only to eliminate him and some of the judiciary have recommended that bite mark evidence should not be acceptable in court.

Using an artificial model of post mortem bites in pig skin it has been shown [7] that even under standardised laboratory conditions, photographs of bites could only with difficulty be matched to the dentitions making them even when the exercise was carried out immediately. The possibility of errors of judgement have been emphasised [8] and it has been urged that more experimental work be carried out. The process has been criticised on a number of occasions in relation to its reliability [9] and the statistical problems associated with it [10,11]. Assessment of the probability of a bite having been made by a particular individual is a difficult subjective judgement requiring substantial experience and knowledge on the part of the expert. Clearly, this judgement is likely to be subjected to rigorous examination in court [11]. Bite mark investigation starts with examination to determine if the wound can be positively identified as a bite mark. If the wound can be orientated in such a way that it is possible to say which teeth in the mouth have caused each element of the mark, then it is appropriate to make a firm statement that the wound is a bite mark. Frequently however, an individual wound will show limited detail and it will be appropriate to identify it only as a possible bite mark.

There is sufficient disquiet in the ranks of practising forensic dentists to warrant further research being carried out and in this study we have tested the reliability of decision making in the interpretation of bite marks and compared established "experts" (forensic odontologists) with non-qualified lay personnel. This study has focused on the ability of these groups to differentiate between human bite marks by adults and children. This decision has important legal implications in terms of cases of non-accidental injury to children.

ROC analysis has not previously been applied to dental forensic decision making but is a well validated method of assessing treatment decisions. The study reported here examined the quality of the decision making abilities, in relation to the analysis and characteristics of bite marks as being of adult/child in origin. In particular, the aim of the study was to assess the degree to which it can be determined whether a bite mark was made by an adult or a child and the effect on the variability of the status of the observer.
2. Materials and methods

A total of 109 observers from England and Wales were included in the study from the following groups: Accredited Senior forensic dental experts (11), accredited Junior forensic dental experts (18), general practitioners with no experience in Forensic Dentistry (12), final year dental students (who had recently undertaken a short course in forensic dentistry including basic bite mark analysis (30), police officers with occasional involvement in child abuse cases (28), and social workers who see "bites" on their clients (10).

A series of 50 actual photographed bite mark cases (combined into a book) were distributed to each observer. The bite mark cases included a mix of non-accidental injuries inflicted by adults and accidental injuries inflicted by children. A self completed questionnaire pro forma (Appendix A) was devised in which the observer was asked to assign a level of certainty regarding whether the bite mark was made by an adult or a child. A pilot study was carried out in the University hospital with a random selection of staff and students, to test the clarity of the questionnaire. The books were then sent out to each observer with a letter explaining the purpose of the study and instructions on how to complete the questionnaire. Each observer was assured that participation would be anonymous and were asked to rate all of the cases to the best of their ability. The data were collected over a period of 6 months.

The results of the decisions made from each group of observers were compared against a "gold standard," which was the actual case verdict from the court. It was acknowledged that the court verdict could always be questionable, although this is unlikely. In addition one of the authors (DW) specifically selected bite mark cases for inclusion where other evidence available strongly corroborated the courts' decision. The experts involved in the original court case may have had access to general information surrounding the circumstances of the case. The information given to the observers was limited to an actual photograph of the injury. This was to ensure that the decision made by the observers was based solely on their ability to identify the actual bite mark and was not influenced by any other information about the case.

When evaluating a decision-making system, it is important to measure both the sensitivity, (in this study, the degree to which observers correctly identified a bite mark as being inflicted by an adult in a non-accidental injury), and specificity (in this study, the degree to which an observer was able to limit the classification of bite marks as being inflicted by adults to those cases in which the bite marks were actually made by adults). In any decision making system there is the risk of cases being incorrectly identified. The measures may be quantified as the false positive rate (when the observer states a child made the bite but in fact it was inflicted by an adult).

It is possible to describe these functions numerically by:

(i) \[ P(TP) = \text{Sensitivity} = \frac{M_{TP}}{M_D} \]

where:
In a series of cases, the expert attempting to classify the bite mark injury will be more certain about classifying some bite marks as being made by adults than others. To actually make the decision the observer will set a "cut off" point or "threshold" of certainty. If the case exhibits enough characteristics of an adult bite to exceed the observer's threshold, then the bite mark will be identified as being of adult origin. If this is not so then the case will be classified as being a child's bite mark (representing a "non-case" of non-accidental injury). As an observer alters the degree of certainty at which he/she identifies an injury as being inflicted by an adult (i.e., uses more or less stringent criteria), the sensitivity and false positive rate will alter. At a given "cut off," particular sensitivity and false positive rates will be achieved. There are two extreme limits to this process: if no matter how certain the observer, no bite mark is identified as being made by an adult, both the true positive rate and false positive rate would be zero indicating that while the observer would not correctly identify any of the adult bite marks, neither would any injuries caused by children be incorrectly misdiagnosed. At the other extreme limit, if no matter how uncertain the observer is that the bite mark was made by an adult, all bite marks are nevertheless identified as inflicted by adults, both the true positive rate and the false positive rate would then be one. This indicates that while the observer would correctly identify all the bite marks made by adults inflicting non-accidental injuries, all of the bite marks inflicted by children would also incorrectly be identified as adult assault cases.

Holding a "cut off" point or threshold at some intermediate level between these two extremes results in both the true positive rate and false positive rate lying between zero and one. The relationship between the two rates varies with changing "cut off" points.
ROC analysis provides a graphical representation of how the true positive and false positive rates vary with varying "cut-off" points [12].

Such an experiment may be undertaken by asking the observer to rate a series of cases on a rating scale (such as that used in this study, shown in Appendix A) to indicate how certain he/she is that the bite mark was made by an adult. True positive and false positive rates can then be calculated at each level of certainty (i.e. threshold) according to a gold standard (in this report, the court verdict). An ROC curve is then constructed by plotting these false positive rates against the corresponding true positive rates, at each level of certainty (i.e. at each action threshold).

The area beneath the ROC curve thus provides a graphical representation of the ability of an observer to discriminate between adult bite marks and child bite marks, at all possible levels of uncertainty. ROC curve analysis provides an objective measure of the observers' ability to identify adult/child bite marks, over all levels of certainty. If the observer were to make random treatment decisions the ROC curve would be a straight line equi-distant from both axes. The ROC curve for a perfect observer given perfect information would be a horizontal line from (0,1) to (1,1) and would contain the maximum area. It follows that the further above the random line, a particular observer's curve lies, the better his discrimination ability. Statistical tests exist to determine whether there is any significant difference between two ROC curves [13]. Changing the observer while keeping the clinical information identical, and comparing the areas under the curves, provides a statistically robust method of comparing two or more observers' decision making ability [14].

The comparison of the results between data obtained from the groups of the observers as described above and the "gold standard" was made using this established ROC methodology. ROC curves were constructed for grouped data for each observer type and the area beneath the curves compared [13]. It has been shown that combining data between observers into one group in this way is statistically valid [15].

3. Results

The ROC curves of each group are shown graphically in Fig. 1 while the areas beneath the ROC curves developed for each group are shown in Table 1, together with standard error data for each. The areas beneath the curves vary between 0.618 and 0.693. As previously stated, the nearer this area approaches the value of 1, the better the decision making of the group.

Table 2 shows the significance of differences between areas beneath ROC curves of each group in pairwise comparison. As can be seen from this table, the respondents could be divided into three groups. The best decisions, judged against the gold standard, were made by senior/junior forensic experts or final year dental students. There were no significant differences in ability to correctly classify a bite mark as adult/child between any of these three groups on the basis of performance. A second slightly anomalous group were social workers who performed significantly worse than senior forensic experts or students but no differently from junior forensic experts. A third group
consisting of general dental practitioners and police officers were least able to
differentiate between adult and child bite marks.

4. Discussion

The literature abounds with reports of the assessment of bite marks in criminal injury
cases, and subsequent opinion as an expert in court [9]. Many of these have been
associated with N.A.I. to children [10,16,17]. In many jurisdictions no formal training as
a Forensic Odontologist is required and courts may choose to regard any dental surgeon

Table 1
Areas beneath ROC curves of different groups of observers classifying bite mark's child/adult in origin

<table>
<thead>
<tr>
<th>Observer group</th>
<th>Area beneath ROC curve</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior forensic expert</td>
<td>0.693</td>
<td>0.0248</td>
</tr>
<tr>
<td>Junior forensic expert</td>
<td>0.680</td>
<td>0.0206</td>
</tr>
<tr>
<td>General dental practitioner</td>
<td>0.618</td>
<td>0.0262</td>
</tr>
<tr>
<td>Student dentist</td>
<td>0.690</td>
<td>0.0157</td>
</tr>
<tr>
<td>Police officer</td>
<td>0.618</td>
<td>0.0171</td>
</tr>
<tr>
<td>Social worker</td>
<td>0.634</td>
<td>0.0305</td>
</tr>
</tbody>
</table>
as an expert in bite mark cases. Justification of an expert in relation to his training and experience is to be expected in court, and a dentist without considerable experience in bite mark analysis would be unwise to become involved in a criminal trial. Assessment of bite mark evidence can be difficult even for experienced forensic odontologists and no attempts appear to have been made to assess the importance of experience in this area. We have chosen to compare the ability of dentally trained and interested lay personnel to make an apparently simple decision as to whether a bite mark was produced by a child or an adult. This decision is frequently crucial in determining whether a case is proceeded with, and even whether an expert is called in to examine the victim [18]. The decision as to whether a bite mark was produced by a child or adult is dependant upon a number of factors which may include size, shape, size of individual tooth marks and recognition of individual teeth. The placement of an assailant in the "child" or "adult" group depends upon the definition of the terms. Dentists, and particularly forensically trained dentists may use different criteria to lay people less versed in the development of the occlusion. It was therefore decided to allow each experimental group to make their own decision as to what was meant by the terms child and adult. No guidance was given to any of the groups.

The groups of observers were selected using the following criteria:

The senior and junior forensic dentists were chosen at random from the list published by the British Association of Forensic Odontology. All the seniors would have at least some experience of decision making in relation to bite marks and all the juniors would have received theoretical training and may have dealt with a small number of cases. The general dental practitioners were randomly chosen and included a wide range of experience and undergraduate training. None had any experience in the forensic field and

---

Table 2

<table>
<thead>
<tr>
<th>Senior forensic expert</th>
<th>Junior forensic expert</th>
<th>Dentist</th>
<th>Student</th>
<th>Police officer</th>
<th>Social worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior forensic expert</td>
<td>Z=0.42 P=0.3372</td>
<td>Z=2.08 P=0.01876</td>
<td>Z=0.10 P=0.4602</td>
<td>Z=2.50 P=0.00621</td>
<td>Z=1.54 P=0.0618</td>
</tr>
<tr>
<td>Junior forensic expert</td>
<td>Z=1.85 P=0.0322</td>
<td>Z=0.40 P=0.3446</td>
<td>Z=2.31 P=0.01044</td>
<td>Z=1.28 P=0.1003</td>
<td></td>
</tr>
<tr>
<td>Dentist</td>
<td>Z=2.36 P=0.00014</td>
<td>Z=0.00 P=0.5000</td>
<td>Z=1.59 P=0.0455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social worker</td>
<td></td>
<td>Z=0.47 P=0.3192</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
had not seen any human bite marks professionally. The dental students were in their final year in the Cardiff Dental School and all had attended an undergraduate course in forensic dentistry including theoretical aspects of bite mark analysis.

The police officers were chosen from those working in family support units in South Wales and all had an interest and in many cases personal experience, in bite mark cases. The social workers were chosen from those who had experience of non-accidental injury to children and bite mark injuries.

The senior forensic dentists showed no significant differences from the juniors in their decision making. It is commonly supposed that extensive experience is required to assess bite marks [19] but in so far as the decision of a child or adult bite is concerned, the seniors and juniors related to the "gold standard" in the same way. In the absence of more detailed information about the case, experience seems to be unimportant when photographs alone are available for study.

The seniors were better than general dental practitioners and police officers and there was a non-significant trend for them to be better than social workers. Juniors were also better than general dental practitioners and police officers. It seems likely that formal training (which had not occurred for the general dental practitioners) is more important than extensive experience in this decision making process. It is interesting that only two senior forensic dentists pointed out the limitations of being able to make a satisfactory decision from photographs alone. Those who responded understood the purpose of the exercise which was not to be able to make a clear decision as to child or adult - but to respond if necessary that the information provided was insufficient to allow a firm decision to be made.

It is interesting that the final year students were better at approaching the "gold standard" than were the more generally experienced qualified dentists. It appears that conventional undergraduate training and subsequent clinical experience does not assist a practitioner in making good judgements about this aspect of bite mark analysis. The brief exposure of the students to a formal course, perhaps coupled with fewer pressures on their time may be important. The students, but not the dentists were better than both police officers and social workers, implying that a knowledge of teeth and their arrangement per se may not be as important as a theoretical knowledge, however simplistic, of the principles of forensic dental bite mark analysis.

Lay persons with an interest in and a knowledge of bite marks were not as good as any of the dental professionals at reaching satisfactory decisions in relation to whether bites were by children or adults and this may be due to a different appreciation of the definition of child and adult in this context or to difficulty in relating information in the photographs to their decisions.

This first attempt to study one single aspect of decision making in bite mark cases confirms the widely held view by the courts that an "expert" in this field should be dentally qualified. However, it suggests that conventional undergraduate education does not equip the general dental practitioners to deal with at least this aspect of decision making any more than does an interest in the subject by lay personnel and provides support for the concept of formal postgraduate training in this area. The effect of training is paramount and suggests that benefits would be gained by ensuring that all undergraduates receive some forensic training [20] and that police and social workers
involved in these cases would also be better equipped to make decisions concerning the need for further advice in cases of this nature. The precise effects of training schemes needs to be assessed in more detail in future studies, as do the effects of providing comprehensive information about the bite marks and the circumstances surrounding them.

The use of this ROC technique, although widely used in other studies, has not been applied to the field of forensic odontology (and rarely to forensic science as a whole), this technique excels at assessing objective differences in performance of different observers and as such, would be applicable to other areas of the discipline.

Appendix A

You are presented with photographs of 50 cases, each showing a human bite mark. Examine each bite mark and rate using the following scale, the certainty with which you would assign the bite mark being made by an adult or a child.

1. I am certain that this bite mark was made by an adult.
2. I am fairly certain this bite mark was made by an adult.
3. It is slightly more likely that this bite mark was made by an adult than by a child.
4. I am unsure whether this bite mark was made by an adult or by a child.
5. It is slightly more likely that this bite mark was made by a child than by an adult.
6. I am fairly certain this bite mark was made by a child.
7. I am certain that this bite mark was made by a child.

References