Factors affecting dental restorative treatment decisions

Elizabeth Jane Kay, BDS MPH FDSRCPS

Submitted for the degree of Doctor of Philosophy, University of Glasgow
Department of Oral Medicine and Pathology and Department of Public Health

June 1991

For my Dad

This thesis is submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy at the University of Glasgow, Faculty of Medicine. Unless stated otherwise, the work is that of the author.

LIST OF CONTENTS

PAGE

LIST OF TABLES

List of Figures

ACKNOWLEDGEMENTS

The list of people who have contributed in some way to a thesis is usually lengthy. These acknowledgements are no different from any other in that respect, except, that I truly wish all those who have helped and encouraged me to know that I am deeply, deeply grateful to them. Some of them will be as amazed as I that I am submitting an PhD; others have had unwavering faith in my ability to complete the task - often more faith than myself.

Firstly I must thank my dear Annelise (Nielsen), who, besides typing the manuscript so frequently and so beautifully, has offered me her friendship, support, nagging (when it was needed), comfort and a shoulder on which to cry. I will never forget her, and this acknowledgement is testimony to that fact. Likewise Dr Robin Knill-Jones, who first inspired me with interest in decision theory in 1983 and has continuously fertilised that 'initial' interest since then. He has been a great supervisor. Professor Stephen I must also thank, not only for his expert supervision, but his encouragement, and most of all, his belief in me. He was a great "boss" to me during the period when the majority of this work was carried out and without his support this thesis would not have been finally beaten into shape.

Two others have been central to this work, encouraging me to form my theories and ideas, and increasing my own understanding of this thesis by their patient listening. These are Dr Raymond Jones and Dr Michael Lewis. I offer them my fondest thoughts and great gratitude.

I received technical assistance and advice from Dr R Strang, and Mr D Weetman, and most of all from Mr Gordon Conell. Gordon is the one person for whom it is 100% true that the work could not have been undertaken without his help. Thanks, Gordon. I owe you.

I must also thank Dr AS Blinkhorn who first introduced me to research and, who has been a mentor of mine for many years, and my good friends in Dundee, - Hazel, Chris and Patsy, who have managed to suppress many a yawn whilst I was in the desperate final throes of completing this work. I must also thank Dr NB Pitts for all his encouragement, and for allowing me time to finish the thesis.

Lastly, I thank Dave for enduring all the unwashed shirts and undusted rooms which this thesis has brought about, and my mum and step-father for always supporting me in whatever I do.

SUMMARY

Dental caries is a dynamic process rather than a diagnosis. It is therefore inappropriate to cite an exact 'threshold' at which interventive operative dental care becomes the 'correct' treatment option. This thesis explored this problem by applying decision theory to one diagnostic test for caries - namely the bitewing radiograph.

Fifteen 'mock' dentitions were assembled, using materials of similar radiodensity to oral structures. A number of cariously involved teeth were included in these dentitions. Radiographs were taken using a standardised technique to ensure that the quality and density of the simulated radiographs were as similar as possible to those taken *'in vivo'*. Twenty dental practitioners each made 360 treatment decisions based on these bitewing radiographs.

Subsequently the teeth were serially sectioned, examined histologically, and the extent of any caries in each tooth was recorded and correlated with the visual appearance of the tooth surfaces. The dentists' decisions were then subjected to Receiver Operator Characteristic (ROC) and Kappa analysis.

The work described also explored the value placed on various tooth-states by the population. By using the available literature and the results of the two studies described above, a decision tree was constructed which permitted examination of the expected values of treatment decisions. This decision analysis allowed recommendations to be made concerning the setting of treatment thresholds.

The first part of the study showed that the microscopic, radiographic and visual appearances of a tooth were not always directly related to one another. The analysis of the dentists' decisions indicated that individual practitioners have differing views about the importance of sensitivity and specificity, and therefore have different thresholds at which they institute treatment. These values and attitudes influenced the treatment decisions made to a greater extent than the dentists' views about the depth of lesion needing restoration.

The decision analysis indicated that, for the sample populations of dentists and patients investigated, the detectability of caries on bitewing radiographs was not great enough to warrant their routine use as diagnostic tools. The treatment thresholds held by the group of dentists in the study were inappropriate in relation to the utilities for dental health held by the population examined.

This study has highlighted the importance of three methods of evaluating diagnostic tests. It has shown that the values dentists place on the outcomes of their decisions affect their treatment thresholds more profoundly than lesion depth. Therefore restorative treatment thresholds should be reviewed and appropriate training given so that dental diagnostic skills are improved and clinical decisions made according to the values that patients hold.

CHAPTER 1. INTRODUCTION

1.1 THE PROBLEMS OF DIAGNOSIS

The diagnosis of caries is fundamental to the practice of dentistry, yet it is only comparatively recently that the difficulties involved in correctly identifying a carious lesion which is in need of restorative treatment have been recognised. [1-4]

One of the major hurdles to be overcome in this field is the apparent inclination of the dental profession to envisage caries as either "present" or "absent", rather than to recognise that caries presents as a continuum of disease and that the disease is a dynamic process, rather than a discrete entity. In effect, each dentist sets for himself a level, or threshold, at which he considers a lesion to require operative treatment. A commonly cited threshold is the point when caries extends to the amelo-dentinal junction, since after this point, the lesion is thought to be incapable of regressing[5]. However, the exact point at which a dentist intervenes surgically is likely to depend on his own treatment philosophy, on his knowledge of the patient and the patient's past caries activity, and on the personal traits of the patient concerned.

It seems that the dentist subconsciously assesses the likelihood that a lesion is present or, if a lesion is already established, that it will progress. The dentist also implicitly weighs the long term outcomes of treating or leaving the tooth unrestored. For example, if the patient is not a regular attender the dentist may be more likely to intervene at an earlier stage, since the probability that he will see the tooth again before pain becomes an intervening factor is low. Therefore from consideration of the large number of factors which influence a dentist's assessment of the outcomes of action or inaction, it is clear that the decision path is complex. This thesis will apply decision theory to clinical dental decision making and makes explicit, some of the implicit calculations which are subconsciously made by dentists.

1.1.1 The approach to the problem

Decision analysis provides a means by which a dentist's decision to treat a tooth can be analysed to investigate the role of personal judgment, to relate this to costs (both to the dentist and the patient) and to describe the "system" by which dental surgeons take decisions to restore teeth. The role of intuitive judgement need not be ignored. It may be introduced into the decision analysis pathway by using numerical scales to express uncertainty about the final decision.

Signal detection theory[6] incorporates both probability theory and statistical decision theory which are used in order to separate the sensory process of detection of a "signal" (e.g. detecting a radiolucency on a radiograph) from the "noise" arising from all the other factors involved in the process of making a decision to restore a tooth[7]. The essential feature of signal detection theory is the construction of *receiver operating characteristic (ROC) curves*[8], which allow compression of all factors influencing the dentist's attitudes towards restoring teeth. An ROC curve plots the relationship of *sensitivity*, or rate of true positive diagnoses, against the rate of false positive diagnoses, at various diagnostic thresholds. A second term *specificity* refers to the extent to which a diagnostic test successfully identifies cases who do not have the disease in question. The false positive rate is therefore equivalent to 1 minus the specificity. Plotting an ROC curve thus allows a description of the "accuracy"[9] of a test at any given diagnostic threshold.

The factors affecting each individual dentist's decisions, if treated cumulatively, will have considerable public health implications. By modelling this situation, it is possible to evaluate the outcomes of dentists' current decision criteria from the point of view of a population. The costs to the community of adopting a particular diagnostic policy, or level at which surgical intervention is undertaken, can be assessed in terms of *health state utilities*[10]. The analysis presented in this thesis allows recommendations to be made about the most appropriate diagnostic criteria, in terms of both 'accuracy' and population utilities.

Decision analysis is a method whereby the probability of the outcomes of a decision are weighted by their "usefulness" or utility. The construction of *decision trees* allows exploration of the likelihood of valuable outcomes and thus enables recommendations to be made concerning the most beneficial decision path under a particular set of circumstances. The value of such analysis lies in the flexibility of the decision tree. For example, the probabilities of a clinical decision resulting in benefit for the patient will depend upon the treatment threshold chosen, and the resultant probability of 'mistakes'. Furthermore, the outcomes of a clinical decision, whether it be negative or positive will depend on the natural history of the disease, once the decision has been made. Decision trees allow the influence of these factors on the utility of the decision to be examined in a logical fashion.

1.2 THE SPECIFIC PROBLEM TO BE EXAMINED

The theoretical background of this thesis may be applied to any dental clinical decision. However, since longitudinal data is available concerning the approximal lesion, and because radiographs are frequently employed for detection of this type of lesion, this thesis uses data concerning approximal lesions as the basis for developing a decision tree for analysis.

It has been highlighted that bitewing radiography should be employed in order to detect carious lesions within the approximal enamel of posterior teeth so that they may be treated preventively rather than restoratively[11]. Bitewing films may also be used to compare a series of radiographs, taken over time, allowing monitoring of lesion behaviour[12]. This allows probabilities of lesion progression, arrest and regression to be calculated. Recent work by Pitts and Renson[13] has led to a better recognition of the nature of the approximal lesion, and this understanding is essential for the development of the models to be used in this thesis. Also, when an approximal lesion is designated for restorative treatment, a Class 2 restoration is placed. This is a type of filling known to be prone to failure and thus frequently replaced[14]. Decisions to restore an approximal lesion therefore have greater costs associated with them, than do decisions to restore occlusal surfaces. This is for two reasons, firstly because radiography is the basis on which the restorative treatment decisions are made, and secondly because the long-term outcomes of restorative treatment are less certain.

1.3 EVALUATION OF DENTAL DECISION MAKING

When a dentist decides to radiograph a patient's dentition he has an impression, through his training and his experience, as to how informative the radiographs will be. He also recognises that radiographic findings may, in some cases, influence his treatment. At one extreme the radiograph will merely complement the information the dentist has already gleaned from his assessment of the patient, and the clinical examination of the mouth. At the other extreme the radiographic findings will comprise most of the information which determines the dentist's decision to fill a tooth. On many occasions the radiograph will have an intermediate influence on the decision process.

There have been few critical evaluations of the diagnostic procedures for detecting approximal caries. The use of bitewing radiography should be justified by clear scientific evaluation, especially in the light of concern about radiation doses[15], limited resources for public health dentistry, increasing costs of restorative care[16], and increasing consumer interest in dental health care procedures[17].

This chapter will describe four methods of critically evaluating diagnostic procedures, with specific reference to the diagnosis of approximal carious lesions. These methods of evaluation are unfamiliar to most dental practitioners, and yet the basic principles underlying them are simple. In essence, they measure the ability of the diagnostic system (in this case bitewing radiography) to detect surfaces in need of treatment, while simultaneously excluding teeth which do not require surgical intervention. Evaluating a diagnostic system must also take into account the priorities of the dentist taking the clinical decision. For example, some may be prevention-orientated and wish to retain unfilled dentitions; some may consider their patients ability to pay for treatment, and will institute treatment accordingly; finally some may be concerned with the collection of accurate information about their patient. Therefore their eventual treatment decision will rest primarily upon the results of diagnostic tests.

The four methods of decision evaluation which will be described in this chapter (1.3.1 - 1.3.4) are; the decision matrix, the receiver operating characteristic curve, information theory, and decision analysis.

If treatment decisions based on bitewing radiographs are examined by just one or even all of these techniques, algebraic equations can evaluate the results of dentists' treatment decisions in a particular patient, or in a particular population.

1.3.1 The decision matrix

A decision matrix allows the relationship between the dentist's assessment of the appearance of the radiograph, and the actual pathological state of the tooth to be examined. This type of analysis is an oversimplification of the clinical problem, since it rests on the assumption that caries is either present or absent from a tooth, and furthermore that the radiograph will have either a "normal" or "abnormal" appearance. This is not the case. As stated earlier, caries exhibits a spectrum of disease severity and this spectrum is represented in the radiographic appearance. However, decision matrices are described as they are the simplest method of evaluating treatment decisions.

The two states - caries present/caries absent and normal/ abnormal radiographic appearance, can be represented as a 2 x 2 table to show the four possible combinations of radiographic appearance and treatment decisions.

Radiographic Test Result	**Caries**		
	+	**-**	
Abnormal	**a**	**c**	**a+c**
Normal	**b**	**d**	**b+d**
	a+b	**c+d**	**a+b+c+d**

From this matrix the appearance of the radiograph can be related to the treatment decision made, i.e. if caries into dentine is taken as the 'correct point of intervention the proportion of positive treatment decisions in all patients who actually have dentinal caries, (a/a+b), represents the true-positive (TP) ratio, or **SENSITIVITY** of the test. This measures the fraction of tooth surfaces with dentinal caries which will be restored after a dentist's assessment of a bitewing radiograph.

The false-positive (FP) ratio is the proportion of tooth surfaces for whom treatment is planned on the basis of the radiographic test but in which caries is not present. i.e. it is the probability that teeth without dentinal caries will be deemed to have disease in need of treatment (c/c+d).

The true negative (TN) ratio is the proportion of negative tests (decisions to leave the tooth unrestored) of all teeth which do not have caries (d/c+d). It is the probability that patients without caries will be deemed not to be in need of treatment. This ratio is otherwise known as the **SPECIFICITY** of the diagnostic test. It measures the fraction of teeth which will be correctly identified as requiring no treatment.

The false-negative ratio (FN) is the proportion of teeth which are deemed not to require treatment amongst all the teeth which actually have caries (b/a+b).

A good diagnostic test has a high TP ratio (sensitivity) and a low FP ratio. That is, the radiographic appearance correctly identifies a large proportion of teeth which require operative treatment without incorrectly indicating surfaces for treatment when they are, in fact, caries-free.

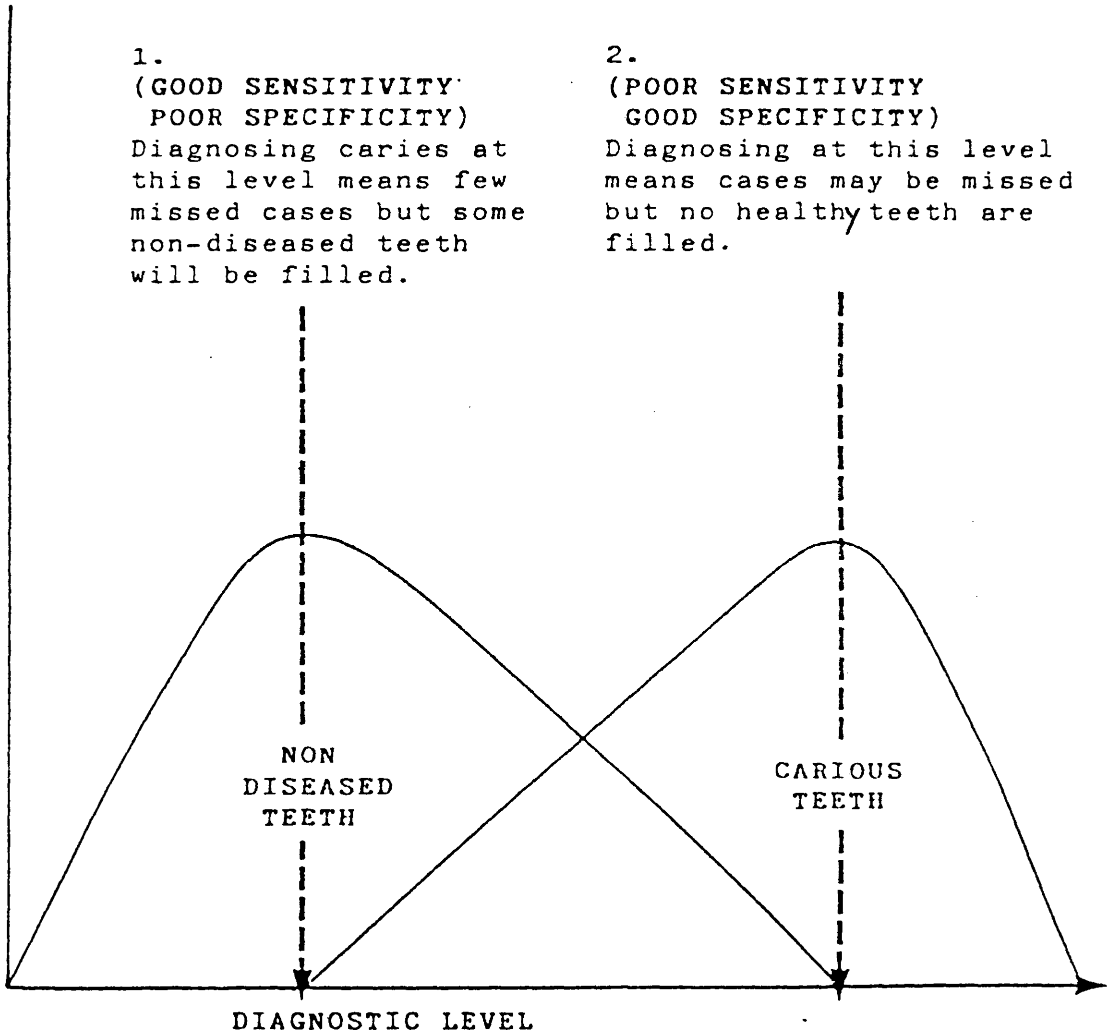
A test which gives high sensitivity and low specificity levels will lead to few untreated carious teeth but a high rate of treatment of teeth which do not need restorative treatment. Unfortunately very few diagnostic systems achieve high levels of both sensitivity and specificity. Therefore, if bitewing radiography for approximal caries diagnosis is to be regarded as a worthwhile exercise, one must first consider whether it is more important to avoid unnecessarily filling sound teeth or, to avoid failing to restore some teeth which are carious. The alternative to risking overtreatment is to accept that some carious lesions will be left unfilled. This entails setting the diagnostic threshold at a level where few healthy teeth will be restored. (See Figure 1.1 overleaf.)

A dentist who restores a tooth, for whatever reason, commits his patient to a lifetime of repair and cavity extension. It would, therefore, seem that the avoidance of restoring wherever possible, especially in disease free teeth may be the preferred option. It is the ultimate aim of this thesis to examine whether dentists' current treatment decisions achieve the aim of increasing the probability of favourable long-term outcomes.

1.3.2 Receiver operating characteristic curves

As stated earlier, the division of a radiograph's appearance into simple binary outcomes is an over-simplification. Like most diagnostic tests, the radiograph yields a range of appearances. Only one appearance can be chosen as a cut-off point to differentiate teeth which need, or do not need, treatment. Where this cut-off point lies will depend on the

Figure 1.1 The effect of treatment threshold on sensitivity and specificity.



relative costs associated with failing to treat diseased teeth versus the costs of treating disease-free teeth. To determine the most advantageous cut-off point, a graph may be constructed of true-positive (TP) ratios (sensitivity) against false-positive (FP) ratios, when various levels of cut-off criteria are employed. The resulting concave curve is known as a receiver operating characteristic (ROC) curve[18] (see Figure 1.2 overleaf).

An ROC curve is generated by holding constant the physical characteristics of the examined radiograph while the examining dentist changes his decision attitude. For example, with a "restorative" attitude a dentist would adopt a policy of "when in doubt, restore the tooth". Such a policy increases the sensitivity of the diagnostic system, i.e. all true lesions are likely to be restored, but at the same time the number of false positive diagnoses must increase.

In Figure 1.2 (overleaf) the diagnostic threshold brought about by a "restorative" attitude will generate points on the upper portion of the curve near A. Adopting a "preventive" attitude of "when in doubt consider the tooth to be sound" generates points on the lower part of the curve, near B. This part of the curve represents the outcomes from dentists who decide not to treat some "suspicious" teeth in order to avoid overtreatment. If it is accepted that decision attitudes vary over time in one dentist, and also vary between dentists, this theory can explain the inter- and intra-observer errors which have been reported in several studies involving treatment decision making[19-22].

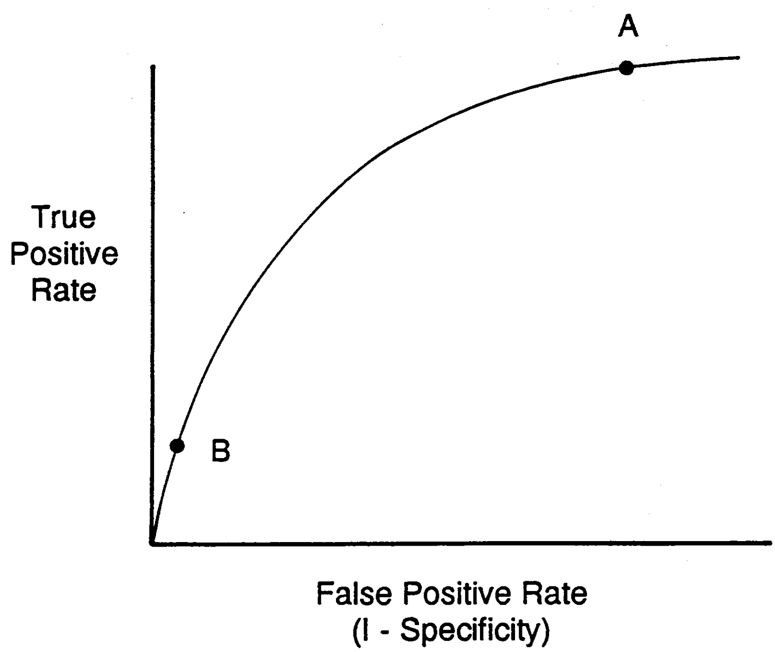
It is clear that individual clinicians will accept different FP and FN rates in their treatment decision making. This concept is examined in this thesis in the context of decision rules and likelihood ratios (Chapters 5 and 6).

However dentists' decisions cannot be evaluated unless their clinical objectives are specified. Put simply, one objective might be to maximise numbers of true positive diagnoses; another objective might be to restrict false positive diagnoses to, say, ten percent. However, in the clinical situation, the objectives of the decision-maker will vary according to numerous factors. These might be: the attitude of the patient; the time available; the patient's previous caries experiences. ROC curves allow simple, optimal decision rules to be described, which can then be applied to several clinical objectives. The choice of clinical objectives depends upon the costs associated with errors and the benefits associated with correct positive and negative decisions.

For example, if the cost of failing to treat a carious tooth is high (i.e. the tooth is likely to require an extraction before the next check-up) this might cause a dentist to adopt an interventionist attitude. Likewise, if the cost of filling sound teeth is low (i.e. restorations are cheap and ensure a long-lived and pain-free tooth) then clearly, the cut-off point, or

Figure 1.2

The Receiver Operating Characteristic Curve



diagnostic threshold must lie near A in Figure 1.2. Such circumstances would suggest that all teeth with any radiolucency should be filled. In contrast, if restorations are expensive, and likely to fail rapidly; if they are uncomfortable for the patient, and if the value of a filled tooth is not greater than that of a decayed tooth, intuition indicates that surgical intervention should take place at B in Figure 1.2, where the slope of the ROC curve is steep. In these circumstances dentists would avoid placing a restoration wherever possible.

This thesis will concern itself with clarifying the probabilities, costs and values of the four components of the decision matrix which exist when dentists decide to restore approximal lesions on the basis of radiographic evidence.

1.3.3 Information theory

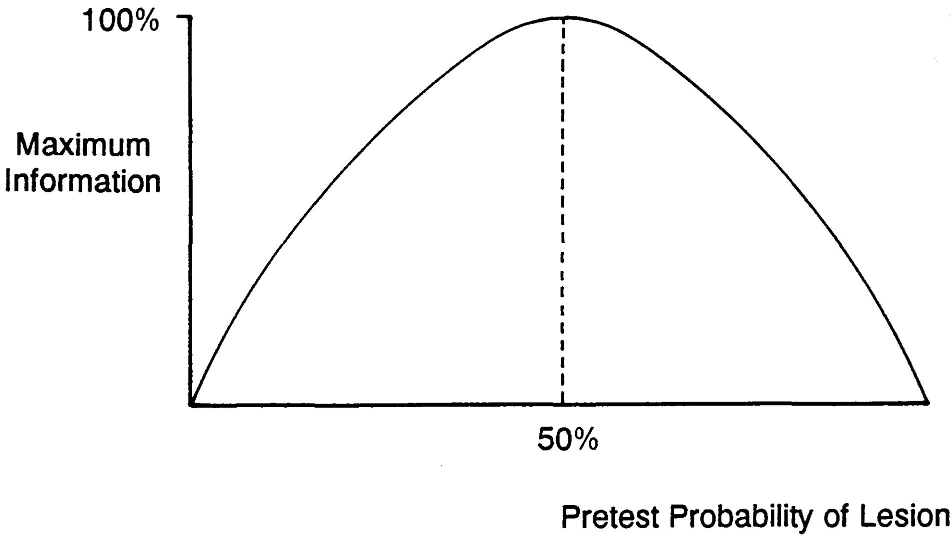
Another method of examining the role of bitewing radiography is information theory. Information from a test or action can be defined as a reduction in uncertainty. It can be measured by the increase in certainty about the diagnosis which the test brings about. For example, the increase in a dentist's certainty that a lesion in need of treatment exists, after a radiograph has been taken, compared to his certainty about the presence of a lesion in need of treatment **before** he examined the radiograph, gives the information content of the radiograph.

Thus, if a lesion is definitely thought to be present i.e. it is visible to the naked eye, can be felt with a probe, or the patient has experienced symptoms of pulpitis, then relatively little is gained from the radiograph. On the other hand if the dentist feels that the probability of a lesion being present is low, the information content of the radiograph will be expected to be low, as it is unlikely to reveal disease in need of treatment. In the former case the pretest estimate of the probability of a lesion being present is already close to 1 before the radiograph is taken. In the latter instance the pretest probability estimate of disease is close to zero. In both instances the gain in information from the radiograph is small. However, if there is a 50% probability that a lesion is present, or not present, then the information content of the radiograph is high. (See Figure 1.3, overleaf.)

The pretest probability of lesion existence is related to the frequency of the disease within a population. For example, for a person who is known to have demographic characteristics which suggest high disease incidence, the information content of a bitewing radiograph is higher than one taken for a middle-class health conscious individual, who is unlikely to develop caries. Similarly the information content of radiographs taken in 12 year old children, when caries incidence is raised, is higher than those taken in adults. There is therefore a relationship between the population prevalence of approximal caries and the information content of a bitewing radiograph. If bitewing

Figure 1.3

Theoretical representation of information content of a radiograph



radiographs were a perfect means of diagnosing approximal caries (i.e. TP ratio = 1, FP ratio = 0), this relationship would theoretically be represented by Figure 1.3. However, since radiographic tests are never perfect, the theoretical maximal information content is seldom reached. The actual value for the information content is therefore dependent upon TP and FP ratios.

1.3.4 Decision analysis and decision trees

A further method by which treatment decision performance can be evaluated, is by developing decision trees. This requires the assignation of probabilities to the outcomes of a dentist's decisions. A simple example can be developed by imagining that the outcomes of a decision not to restore a tooth with a radiographic shadow might be pain, or extraction, or perhaps full health as the lesion regresses. The probabilities of all of these outcomes may be small, but clearly the utilities or "usefulness" of these different outcomes vary considerably.

The formation of a decision tree requires that probabilities are assigned to all possible decision outcomes, and utility or "usefulness" scores are assigned to each of the possible outcomes of treatment, or non-treatment.

This thesis will concern itself with the construction of decision trees relating to dentists' decisions to restore approximal lesions. In the relevant study, described in chapter 9, the tree will only be applicable to the decision whether or not to treat an approximal lesion when the basis underlying the decision is a bitewing radiograph. However the decision analysis approach can be applied to other types of lesions and other tests on which treatment decisions are based.

1.4 MEASURING OBSERVER VARIATION

Variation amongst dentists when they are planning treatment may arise from two distinct, but overlapping, sources. Firstly, if dentists hold different views as to the depth of lesion which warrants intervention, the differences in their treatment plans will be systematic i.e. one dentist may fill many more teeth than the other, but the surfaces deemed to require restoration by the first dentist will include surfaces filled by the second. Such variation is systematic, and is otherwise described as bias. The first dentist requires less evidence of pathology than the second before he decides to restore a tooth. The first dentist's criteria for restoration cause him to operate with high sensitivity. The second dentist requires greater evidence before he restores teeth, and therefore he fills fewer surfaces. However, unless random error intervenes, all the teeth which the second dentist fills would be filled by the first dentist.

The second source of variation between dentists is random variation. It is not caused by consistent differences in dentists' diagnostic criteria. This kind of variation is due to random inconsistencies between dentists of which they may not be aware. This thesis will examine the importance of dentists attitudes to treatment decision-making, and therefore investigates the role of both systematic and random variation.

At the present time there is no universally accepted method of analysing observer variation. In the past, correlation coefficients have been used. This is inappropriate since correlation coefficients merely measure the strength of a relationship between two variables, not the agreement between them, i.e. for perfect **agreement** the points have to lie along the line of equality (y=x). For perfect **correlation** the points can lie along any line. i.e. two examiners may report that 50% of a series of teeth are carious - apparently they are in perfect agreement yet each could be diagnosing a different 50% as carious. The comparability would be 100% yet true agreement would be 0%.

Differences between observers can be examined in several ways. Initially the data must be recorded as a 2x2 table, the four cells of which contain the following information:

(a) the proportion of surfaces agreed on as sound
(b) the proportion of teeth Examiner I considers sound but Examiner II considers carious
(c) the proportion of teeth Examiner I considers carious, but Examiner II sound
(d) the proportion of teeth which both examiners agree are carious.

		Examiner I		
		Sound	**Carious**	
Examiner II	**Sound**	**p**	**q**	**p+q**
	Carious	**r**	**s**	**r+s**
		p+r	**q+s**	**p+q+r+s.**

Examiner agreement can be measured in three ways.

1) The proportion of agreement[23] $\frac{p+s}{p+q+r+s}$

2) <u>Dice's Coincidence Index</u> [24]
This index provides a measure of the probability that the diagnosis made about a tooth surface by one examiner will be agreed upon by another examiner.

e.g. probability that surface diagnosed as sound by Examiner I will be diagnosed as sound by Examiner II $= \frac{p}{(p+r)+(p+q)/2}$

probability that surface diagnosed as carious by Examiner I will be diagnosed as sound by Examiner II $= \frac{s}{(q+s)+(r+s)/2}$.

3) <u>Cohen's Kappa</u>[25]

The Kappa statistic is a chance-corrected measure of agreement. For example, if one observer records the presence or absence of caries, and a second observer records this in the same patient, then a number of the agreements between them will arise by chance. This is perhaps more easily explained by two observers who are asked to record the dental health of an individual. If one observer is recording periodontal disease, and another recording caries, they will still, by chance, both agree that some teeth are diseased despite the fact that they are seeking to record different entities.

The formula for calculation of the Kappa statistic is $\frac{p_o - p_e}{1 - p_e}$. Where p_o is the proportion of observed agreement ($p+s$), and p_e is the proportion of agreement which could be expected by chance

$$p_e \text{ (sound)} = (q+p) \times (p+r)$$
$$p_e \text{ (carious)} = (r+s) \times (q+s)$$

It is important to evaluate the differences in treatment planning between dentists. Systematic errors can be detected by comparing ROC curves to the dentists' reported intentions, while non-systematic errors can be evaluated by use of the Kappa statistic.

The Kappa statistic is, in essence similar to the chi-square statistic. If there is perfect agreement between dentists the resultant Kappa score will be equal to 1. A negative result implies that the decision criteria used are so dissimilar that the operators are tending to do the opposite to each other.

In general	K =	0.8 =	excellent agreement
	K =	0.6-0.79 =	substantial agreement
	K =	0.4-0.59 =	good agreement
	K =	0.25-0.39 =	moderate agreement
	K =	0 - 0.25 =	poor agreement.

1.5 CONCLUDING COMMENTS

Dental treatment has traditionally been evaluated on the basis of conventional wisdom, and treatment thresholds determined by 'expert opinion', rather than by logical analysis of the outcomes which occur as a result of treatment.

This chapter has outlined the four methods of analysing dental treatment decision-making. By applying such analyses to the restorative treatment decisions made by a group of dentists, it is possible to evaluate their treatment decisions in terms of their impact on the dental health of a population.

CHAPTER 2. LITERATURE REVIEW

2.1 INTRODUCTION

A large part of this thesis will concern itself with the incorporation of data from previous studies into a decision analysis. It is therefore necessary that a comprehensive review of the relevant literature is undertaken in order to describe the derivation of the data to be used in the decision analysis (Chapter 9). The review will be divided into nine sections as detailed below.

2.1.1 The approximal carious lesion

The first section of the literature review concerns itself with the prevalence of dental caries in general and, in particular, the prevalence of approximal caries. Changes in caries incidence and the consequent proportional changes in the prevalence of approximal or other lesions, will affect the utility of the outcomes of treatment decisions. The purpose of reviewing the literature concerning general caries prevalence is to underline the relative importance of the approximal lesion.

2.1.2 Progression/regression of lesions

The second part of the literature review concerns itself with the behaviour of approximal carious lesions. Clearly, an understanding of the natural history of the smooth surface lesion is a prerequisite to this research work. The rates and probabilities of lesion progression are required in order to examine the outcomes of treatment decisions. Decisions not to restore teeth at an early stage in the caries process may lead to several outcomes which depend on the probability and rate of lesion progression.

2.1.3 Signal detection theory

Section three of this review introduces the concept of signal detection theory with respect to the interpretation of radiographs. The purpose of this technique, is to enable a distinction to be made between the detectability of a lesion and the decision criteria of the observer. Since little work has hitherto been carried out concerning signal detectability and bitewing radiographs, much of this part of the review will involve interpretation of work reported in more medically related fields.

2.1.4 Restoration failure

The review will then examine the information available concerning the fate of Class 2 restorations i.e. the restorations traditionally placed when an approximal lesion is deemed to need surgical intervention. The decision to restore a tooth cannot be made rationally without prior knowledge as to the likelihood of success or failure, of treatment. Therefore it is necessary to review the possible outcomes of the decision to restore a tooth. Success or otherwise, of new restorative techniques and materials can only be truly evaluated if both restoration longevity and the point in the caries process at which surgical intervention takes place, are known.

2.1.5 Dental radiography

This thesis will attempt to describe the optimum timing of surgical intervention following radiographic examinations of a dentition. Part Five of the review will examine the uses, advantages, disadvantages and limitations of dental radiography. The health costs of using bitewing radiographs will be studied. This section will also discuss what is needed from bitewing radiographs if they are to provide maximum information to the operator. The ultimate aim of the thesis is to examine the role of the bitewing radiograph as an adjunct to clinical decision-making. It is therefore necessary to clarify the current usage of radiographs.

2.1.6 Relationship of radiographic appearance to actual lesion depth

Past research shows radiography to be an important, but occasionally unreliable, method for examining the extent of a carious lesion. Hence, the sixth part of this review will focus on the relationship between the radiographic image of a carious lesion and the actual extent of the same lesion. The closeness of this relationship will influence the information content, the usefulness, and the effectiveness of bitewing radiography.

2.1.7 Decision analysis

The foundation of this work lies in the rational analysis of a dentist's clinical decision-making process. Part seven of the review will discuss some uses to which decision theory has been put in both the medical and dental fields, and will attempt to demonstrate the under-utilisation of this important tool in dental clinical decision-making.

2.1.8 Variation among clinicians in treatment planning

In section eight, existing evidence regarding the variation among dentists when diagnosing carious lesions will be examined, and the review will investigate the effect that such variation has on both measurements of treatment need, and epidemiological studies of caries prevalence. Variation between dentists has recently caused controversy in the British media[26] and is thus worthy of careful scientific scrutiny. This thesis aims to explain some of the causes of these variations.

2.1.9 Evaluation of health outcomes

The value placed by the public on dental health, and therefore on the outcomes of dental decisions, must be delineated before a decision analysis can become meaningful. Therefore the final section of the review will examine previous attempts to develop health status profiles. In the past, such profiles relating to oral health have related only to population evaluations of dental ill health. The research described in Chapter 8 attempts to determine the value which populations place upon individual teeth within their mouths, rather than to evaluate the health of the mouth overall. Information about the public's and profession's views on "health", as it relates to the mouth, can be gathered by examining the literature on this subject. Such information allows decisions to be made which enhance "health", rather than them being based merely on the assumption that pathology must be treated.

2.2 REVIEW OF LITERATURE

2.2.1 Prevalence of dental disease

The patterns of dental disease are changing in both industrialised and developing nations. For example, evidence is now readily available from many of the industrial countries that dental caries, particularly amongst young adults and children, has declined dramatically since 1970[27-33]. Such a decrease in caries prevalence has had important implications for those planning, implementing and evaluating dental health services. The increasing rarity of dental caries has also affected professional training and the manner in which diagnosis of this previously common dental disease is executed.

In Britain, for example, Anderson *et al*[27] reported on the dental health of 12 year-old children in two schools in the South-West of England, first in 1963 and again in 1978. There was a 35% reduction in the mean numbers of decayed, missing and filled teeth (DMFT) values between the two studies, with more caries-free children, and fewer with untreated caries being seen in 1978 when compared to 1963.

Furthermore, important sources of caries prevalence data are available from a large number of clinical trials carried out in Britain since the late 1960's, with six having been conducted in the North West of England from 1968. Here, the subjects were school children aged between 11 and 12 years at the beginning of each trial[34-38]. From the extensive data generated in one area over this time-period, it can be shown that there has been a steady reduction in caries in the population described. However, the results are not directly comparable because each set of examinations was conducted by different examiners in slightly different geographical locations and therefore may be questionable. The most important finding from these studies, from the point of view of this thesis, was that the percentage caries reduction across the years was greater for approximal surfaces than for occlusal surfaces. Since, on a population basis, disease prevalence affects the information content and the predictive power of a diagnostic test, such a finding has important implications regarding the use of diagnostic tests such as bitewing radiography.

It is clear that one of the effects of the reduction in caries prevalence is an alteration in the distribution of carious tooth surfaces in the mouth and, as stated above, this has important implications for the every-day practice of dentistry. Unfortunately, many epidemiological surveys describe caries on a tooth-alone basis[39,40] with the result that the distribution of caries on various tooth surfaces is unknown. Thus it is impossible to determine the surfaces which are benefitting most from the overall caries prevalence decline, and it is difficult to define treatment needs with any certainty. Likewise, little work has examined caries activity in terms of the progression of established lesions. Despite the fact that many of the reported trials were undertaken before the widespread use of fluoride dentifrice, and therefore may not necessarily reflect the situation as it stands today, this type of study is crucial to the models to be constructed in this thesis, and will therefore be dealt with in detail in the next section of the review.

2.2.2 Progression and regression of lesions

In recent years, there has been a marked increase in understanding of the natural history of the smooth surface carious lesion. This has been due primarily to the standardisation of radiographic techniques[41], the increased use of bitewing radiography in epidemiological studies[42], and the introduction of computer-aided imaging[13].

The first studies concerning the natural history of carious lesions tended to concentrate on the occlusal surfaces, mainly because of the ready availability of these lesions to clinical examination. Boyd *et al*[43-46], reported on the progression of caries in 212 teenage children in a series of papers published in the early 1950's. These workers suggested that the median time for progression of an early occlusal lesion to dentine involvement was three years, and that many incipient lesions did not progress at all.

In 1956, Parfitt[47] again addressed the issue of carious lesion development from its early stages to cavitation. He pointed out that it is important to know the speed of caries development in order to assess the most beneficial recall time for patients. His study, like that of Boyd and colleagues[43-46], dealt only with occlusal caries, but the data are nevertheless of interest, since they led Parfitt to be the first person to document the considerable variation seen between individuals in their rates of caries progression. He showed that between 20 and 53% of lesions remained at the "incipient" level for two years or more. Most importantly, he also pointed out that the transition from incipient caries to clinical caries could not be observed precisely, since it represented merely the imposition of a clinical grading on a pathological process.

The concept of caries progression and arrest, as first introduced by Boyd, appears to have been largely ignored as a research subject for the next decade. However, in 1973, Berman and Slack[48] investigated caries progression in the approximal surfaces of 11 year-old children's posterior teeth when they confirmed the hypothesis that caries progression appeared to be a slow process. It was shown that 50% of observed enamel lesions did not progress within the three years of their study, although they added the proviso that this was within the limits of the diagnostic system used, i.e. radiographic examination under standardised conditions.

In 1975, Haugejorden and Slack[49] clarified these ideas by comparing two systems of caries scoring, despite the fact that Rugg-Gunn[50] had previously questioned the validity of assessing the precise depth of a lesion from a radiograph. In Haugejorden and Slack's study, they identified and clarified the problems of converting a pathological condition to a numerical score, especially when employing an imperfect diagnostic system such as radiography, and particularly when lesions can both regress and progress. They showed that the progression rate of lesions could apparently vary by 50%, merely by altering the scoring system. Nevertheless, they concluded that the findings of earlier workers (i.e. that few lesions appeared to progress within a twelve month period), were correct, although they reported that 23-25% of lesions progressed. They also noted that reversals, or regression of lesions, were relatively rare (6-8%) and that the majority (53-58%) remained apparently unchanged over three years. These authors described their own and others' diagnostic criteria in detail, a trait unfortunately uncommon in earlier work. They used a system similar, although not identical, to that employed by Dirks *et al*[41], and if their results are interpreted in terms of initial lesion depth, 30% of lesions initially limited to the outer half of enamel, progressed within a year. Of those within the inner half of enamel, 19-26% were seen to advance. There were too few observations to assess progress of lesions which were observed at baseline to extend well into dentine.

In 1976, Zamir and colleagues in Israel[51] added to the growing collection of longitudinal studies, and measured the speed of propagation of initial approximal caries using bitewing radiography. One of the purposes of their investigation was clarification of the most suitable frequency for this type of test. They showed that of 96 initial lesions observed in 51 patients, only 16% reached the amelo-dentinal junction within 24 months of the lesion being noticed. However, of 27 lesions first recorded as extending into enamel and not extending to the amelo-dentinal junction (ADJ), 48% reached the ADJ or beyond within 12 months. It is also worthy of note that after 36 months, 12% of the initial lesions, and 15% of the more advanced lesions, had remained unchanged. The only factor amongst oral hygiene, DMFT and number of untreated lesions which appeared to correlate with the rate of lesion progress, was the latter. The authors concluded, in accordance with Berman and Slack[48] that only a small proportion of carious lesions detected in the outer half of enamel were likely to progress far into dentine in two years, and none would reach the pulp within that period. The authors therefore recommended that the minimum time-interval between bitewing radiographic tests should be two years.

In a similar longitudinal study, Kolemainen and Rytomaa[52] examined the progress of approximal lesions in Finnish dental students. This 1977 study allowed insight into the effect of good oral hygiene and preventive dentistry practices on caries progression rate. They showed that only ten incipient approximal lesions advanced to become dentinal caries over a twenty month period.

In contrast to the above papers, all of which appear to confirm the hypothesis that caries is a slow process, a study of approximal caries progression in deciduous teeth, reported in 1978 by Murray and Majid[53] in Great Britain, showed that 69 of 71 newly initiated lesions progressed into dentine in one year. Although there is doubt about the comparability of the diagnostic criteria between this and the earlier studies of permanent teeth, it appears there is strong evidence that caries advances much more rapidly in the deciduous dentition, than in the permanent.

In a series of investigations by Gröndahl and Hollender, and Gröndahl *et al*[54,55], 158 teenagers were followed-up for more than six years, making these studies the most comprehensive research investigation of the development of the approximal lesion. Initially, the authors showed that only ten per cent of intact surfaces developed caries during a three-year interval in 16- to 19-year old participants. They commented on the phenomenon, raised earlier by Boyd *et al*[43-46], that there was a marked variation in caries progression between individuals. The authors qualified this observation by showing that 20% of individuals in the study accounted for about half the newly observed carious lesions. The authors attempted to examine the group of 158 individuals once more, after a further 3 years, when the subjects were aged 22 years, and succeeded in radiographing 100 individuals. During the second three-year interval, 4.7 per cent of the remaining intact surfaces developed caries, or were restored, whereas the corresponding value for the preceding three year interval was 10.1 per cent. When the differences in decayed and filled tooth surfaces for each individual in the two time periods were compared, it was shown that 22% had an increase, 10% showed the same values for each period, whilst 68% had experienced a decrease in their decayed and filled tooth surface score. These findings thus confirmed the reports of previous authors, that caries is both a slow process, and is apparently reversible.

In 1980, Granath *et al*[56] demonstrated that 8% of 530 incipient enamel lesions in 12-13 year olds progressed into dentine within one year, whilst 51% progressed further into the enamel. However, of lesions which were found to be deep into enamel at baseline, 35% entered dentine or were restored within the study period. Two years subsequent to baseline, 67% of the 236 shallow lesions had progressed, and 24% of these had reached dentine, while the corresponding figure for deep lesions was 58%. The findings of these workers were therefore in contradiction to those of previous authors and suggest a high rate of progression of radiographically detectable carious lesions.

In general, studies in low fluoride areas have shown that the initial lesion progression rate in approximal surfaces is low, and many such defects show little change over long time-periods. It has also been shown that application of fluoride in dentifrices and mouth rinses could reduce the caries progression rate[57,58]. Indeed Muhler *et al*[59], claimed that topical fluoride treatments could reduce expansion of incipient carious lesions by 80%. Given the wide availability of fluoride dentifrices, it is important to understand the effect which these substances have on lesion progression.

In 1981, Powell, Barnard and Craig[60] reported on the effect of application of stannous fluoride (SnF) on approximal lesion advancement in permanent teeth of 12-14 year olds. They showed that over a 48 month period, 89% of initial lesions had progressed into dentine. In the group using stannous fluoride and dentifrice, 49% of initial lesions reached dentine, whilst in the group receiving both professionally applied topical fluoride and stannous fluoride dentifrice, only 29% extended into dentine. The dentifrice exhibited a marked inhibitory effect on lesion progress since, at 48 months, 51% lesions remained within enamel, and 23% showed no change at all. These authors suggested that a suitable time-interval between radiographic examinations would therefore be 12-18 months, provided that a topical fluoride regime was being employed. They also stated that treatment should be delayed until there was radiographic evidence of lesion extension into dentine.

In 1983, considerable progress was made in the interpretation of caries progression. S hwartz *et al*[61] attempted to minimise the biases in estimating progression rates. They claimed that previous authors had overestimated the speed of progression by ignoring filled and non-progressing lesions. Their results showed that in newly-erupted first permanent molars, it took 21-23 months for a lesion to progress through the outer half of enamel, and between 19-28 months for a lesion to progress through the inner half of enamel to reach the amelodentinal junction. In older individuals, caries advancement was slower, with lesions taking between 38-41 months to progress through the outer half of enamel, and 47-56 months through the inner half. The authors then made the assumption that lesion duration in each half of the enamel followed an exponential distribution with a mean of two years. This, they suggested, implied that about 10% of new lesions would progress through enamel in one year, and 25% in two years. However, it is also worthy of note, that the same assumption implied that over 40% of lesions would **not** have progressed in four years. Furthermore, the authors also concluded there were no significant differences in progression rates between high and low risk individuals.

Additional support for the hypothesis that caries, once initiated, need not tread an inevitable path ending in restoration, came from a lesion progression study in dental students[62]. Cook showed that 68% of initial enamel lesions either remained constant or exhibited some improvement over a 32 month period. For larger lesions, 58.3% did not progress and the author concluded, like Zamir and colleagues[51] that radiographically detectable lesions should not be restored until there was clear clinical evidence of dentinal disease. However, when interpreting results of studies such as these, it is important to remember that dental students have relatively fewer carious tooth surfaces compared to other subjects of similar age[54].

In 1984, Darvell and Pitts[63] reviewed earlier studies and pooled the available data in order to construct a mathematical model to represent caries progression through approximal enamel. These authors highlighted a concept which is to be a major theme of the current research, namely, that the presence of an enamel lesion merely indicates a "risk", of lesion progression rather than an inevitable progression. Therefore the likelihood of a shallow lesion "surviving" as such within a given period, must be considered when planning treatment for lesions of this type. The final model developed by the authors permits an assessment of the risk that an enamel lesion will progress into dentine.

By using the formula they developed, one can calculate that even at the peak rate of lesion progression, the carious process is a relatively slow one and that 50% of early lesions do not progress within 2.5 years. This represented a slightly lower value than an earlier estimate made by one of the authors[64]. Pitts had previously reviewed 19 studies and concluded that the progression of interproximal lesions takes, on average, 3-4 years. However, if the lesions considered were confined to those in the outer half of the enamel, 5-6 years was the time taken for a lesion to progress to dentine.

Shaw and Murray[65] reported one of the most recent, and therefore most relevant, longitudinal studies investigating the progress of approximal lesions in Britain. These authors examined standardized bitewing radiographs of 1162 11-12 year olds, taken at annual intervals over a three-year period. Their study differed from those reported previously, in that an initial lesion was not considered to be present unless there was a radiolucent area reaching the inner half of the enamel. Of these, 19.8 per cent did not progress from baseline over a three-year period. However of lesions with dentinal involvement at outset, only 10.9% had not progressed, or been restored. Unfortunately, the above statement immediately highlights a difficulty with interpretation of such longitudinal studies. The number of lesions restored in a given period of time does not depend solely on the progression of the carious lesion, but also depends on the treatment philosophies prevalent among dentists treating the study population. In Shaw and Murray's study[65] over 40% of surfaces restored during the period of study had never been recorded by the investigators as having any type of radiographic lesion. Furthermore, almost 80% of approximal lesions diagnosed in the inner half of enamel had either progressed or been restored within the study period. The rate of caries progress to dentine detected by these authors was thus higher than any reported previously, but this difference could be accounted for by the high restoration rate of apparently disease-free teeth.

In 1987, a study by Ekanayake and Sheiham[66] showed that rates of lesion progression had slowed between the 1970's and 1980's. They suggested that in the 1980's only 11% of outer enamel lesions progressed to the inner enamel within a three-year period. Finally, the most recent report of approximal caries progression[67] showed that 2% of surfaces, initially diagnosed as sound in 11 and 12 year-old children, had been filled by the time they were 15-16 years of age; 4% had developed caries in the outer half of enamel; 7% had developed caries up to the amelodentinal junction, and approximately 3% had developed caries into dentine. Of the surfaces which had shown minimal lesions restricted to the outer half of the enamel, 28% were considered sound at 15-16 years of age, 16% had been filled, 14% had remained static, 21% had progressed to the amelodentinal junction, and 18% had advanced into dentine. Similarly, for lesions which had been noted in the inner half of enamel at age 11-12 years, 13% appeared to have regressed completely, 29% were restored, 24% remained static, and 25% had reached into dentine by the age of 15-16 years. With respect to those lesions which had already progressed as far as the outer half of the dentine at baseline, 59% had been filled, 23% had remained static and 9% had entered the inner half of dentine. Where caries had entered the inner half of dentine when the individual had been 12-13 years of age, the majority of the teeth had been extracted by the time of the follow-up study at 15-16 years, However, although 19% of the deep lesions in teeth which had not been extracted had been filled, 19% appeared not to have increased in depth.

The conclusion which can be drawn from these studies is that lesion progression is a slow, and with respect to enamel lesions, sometimes a reversible process. This review also demonstrates that caries is certainly not inevitably a progressive process. However, knowledge of the exact manner of caries progression is still scant, due to the fact that longitudinal studies are rare, and are seriously affected by the biases in the diagnostic criteria applied by each examining dentist. In Table 2.1, a summary of the key papers reviewed in this section is given. This information will be used in the construction of decision trees in Chapter 9.

Table 2.1 Rates of approximal caries progression

Author	Year	Population	Time	Enamel only into dentine	Outer into dentine	Inner into dentine
Berman (48) + Slack	1973	353 11 yrs olds	3 yrs	50%	-	-
Haugejorden (49) + Slack	1975	40 13-15 yrs olds	3 yrs	30%	30%	23%
Zamir et al (51)	1976	51 14-24 yrs olds	2 yrs	20%	-	50%
		51 14-24 yrs olds	3 yrs	28%	-	85%
Kolemainen (52) + Rytomaa	1977	59 Dental Students 23-24 yrs olds	2 yrs	4%	-	-
Granath et al (56)	1980	126 12-13 yrs olds	2 yrs	24%	51%	58%
Cook (62)	1984	95 (young adults) Dental Students	<3 yrs	15%	21%	8%
Shaw (65) + Murray	1986	1431 11-12 yrs olds	3 yrs	-	80% (64%)*	67%
Dummer (67) Oliver and Shaw	1988	4810 11-12 yrs olds	4 yrs	48%	21%	58%** 25%**

* Excluding surfaces restored during study period

** Excluding surfaces extracted or restored during study period

2.2.3 Signal detection theory

Accurate interpretation of a radiograph depends on the visual perception by a dentist of the images recorded on a film. The application of signal detection theory implies that one considers the image of a pathological lesion as a signal, or group of signals. The dentist observes these signals against a background of shadows and images of structures other than lesions, which introduce difficulty and confusion, known in this context as "noise". The dentist, in the role as clinical decision-maker, tries to identify the signal from the noise in order to choose among the available treatment options. The dentist therefore usually makes decisions about the optimal treatment pathway under conditions of uncertainty.

Thus, each time a dentist makes a decision as to whether a patient requires treatment, an estimate is actually being made of the probability that a carious lesion exists. Many dentists would perhaps feel uncomfortable that this implicit behaviour can be formalized by using probabilities based on relative frequencies. Most dentists would feel that such a procedure would not make sense when applied to an individual patient. And yet, if this same decision process is placed in an informal context, dentists are happy to be considered to be individualising diagnosis and treatment for each particular patient, on the basis of their past experience and personal judgment. All dentists, when making decisions under uncertainty are, in fact estimating probabilities from a small sample[68].

The general theory of signal detectability was developed in the early 1950's by mathematicians and engineers. The theoretical material arose from an analysis of radar and other information systems. For example, a human observer may be asked to watch a radar screen and detect a signal on the screen indicating the presence of an aircraft. This signal might be a small "blip" or a bright spot of light. This sounds simple, but it must be remembered that the blip is always observed against a background of "noise" caused by electronic interference in the background receiver, and, by extraneous radar echoes from hills, clouds, ground and buildings. The evaluation of such equipment requires a "gold" standard to which the performance can be compared. The concept of an ideal, or optimal detector, is central to the theory of signal detectability.

Investigators have applied detection theory to experiments in vision and learning[69], and have shown that the principal advantage of detection theory is that it allows the compression of all factors affecting an observer's attitude, into a single variable called the 'decision criterion'.

The main purpose of using signal detection theory in this thesis is that it allows one to separate the sensory process of the visual detection of an approximal lesion, from the treatment decision. It is important to realise that, in using signal detection theory a sensory "threshold" below which a lesion cannot be perceived is no longer considered. In detection experiments, one accepts the premise that the usefulness of a test is continuously variable. An examination of the detection behaviour of the dentist examining a radiograph aims to answer several questions. Firstly it may lead to some explanation of the long-standing problem of inter- and intra-observer variation, and it might also be possible to improve the consistency of clinicians' diagnoses.

An ROC curve allows a distinction to be made between the inherent detectability of a signal (the visual image of caries on a radiograph), and the judgment of the observing dentist or dentists. The very important, and almost unique feature of this method of presenting data about treatment decisions, is that the results it gives are completely independent of any assumption one might make about the statistical distribution of the signals (the prevalence of carious lesions), the results are therefore "distribution free"[70]. Another more pragmatic, but less important reason for the use of ROC curves is that they can be generated without changing any of the physical parameters or conditions under which the test takes place. Thus one can employ the same series of radiographs throughout a study. This is very important when examining for caries, the true extent of which can only be verified by actual sectioning of the relevant teeth. Using the same radiographs throughout the study means that any variation between dentists in their treatment decisions is due solely to the method each individual uses to decide among the various treatment options.

Signal detection theory has been widely applied to medical problems[71-75]. Garland[74] found that a radiologist with a policy of "when in doubt, call a shadow 'pathology'", had a high rate of true-positive diagnoses, but a concomitantly high rate of false-positive diagnoses. He also showed that a radiologist could change the way in which he operated by consciously changing his decision attitude. As a result of this work, Lusted[71] surmised that it might be useful for a radiologist to know what his ROC looked like for each type of x-ray examination, and that an individual might be able, by being aware of his decision attitude, to hold his false-positive error-rate within certain limits. From this premise Lusted[71] introduced the idea that it is possible to determine the weighting that an individual radiologist assigns to the values of a correct decisions, as compared to the negative values placed on errors.

A 1968 study by Alcorn and O'Donnell[76,77] compared the ROC curves produced by trained and untrained radiologists when assessing mammography results. The "gold standard" or reference test used to construct the curves was a pathologist's diagnosis. Now, while this reference diagnosis is still, in part, "opinion", it is important to recognise that in ROC analysis a reference test is considered to be the "gold standard", or "recording angel" version of the truth.

In 1971 the concepts underlying ROC analysis were introduced to the dental profession by Goldstein *et al*[78]. Although they analysed the error rates of students examining dental radiographs, they did not construct ROC curves of the data. It is interesting to note the long interval between this study, using signal detection theory, and the interest which has recently been shown in the technique, particularly by Mileman[79].

Following Alcorn and O'Donnell's[76,77] indication that training affects observer response, in 1975, Herman and Hessel[80] attempted to examine more thoroughly the relationship between formal training and the accuracy of x-ray diagnosis. Such a study is of interest as cognitive processes such as interpretation of radiographs are affected by individual abilities, informal education, and life experiences. These workers[80] randomly selected 100 chest radiographs and asked eight radiographers at various stages of training to examine and report their findings from the films. These authors found there was no association between length of training and accuracy, and that, interestingly, a first year resident had the lowest error rate! These findings were in agreement with Sheft, Jones and Brown[81], who also found little difference in the reporting of radiographs between resident and staff radiologists. Herman and Hessel[80] concluded that innate abilities have a stronger influence on accuracy, than the length of formal training, since much sophistication in visual perception has been acquired by individuals long before they are trained in radiology.

The view that an examiner's report of an abnormal finding on a radiograph is a result of a statistical decision made under conditions of uncertainty, suggests that errors are inevitable. The examiner guards against excessive false-positive reporting by applying a criterion or threshold derived from prior experience i.e. true-positive findings are omitted if they do not look "abnormal" enough. The counter-arguement is that faulty visual search may also be a major, and avoidable source of error. Several studies of radiologists, in which eye movements during film interpretation have been recorded, show that large areas of a film are never examined by foveal vision[82-84]. Such omissions are thought to arise because the examiner terminates a visual search prematurely, usually when the eye alights on what is considered to be a key finding. This view would suggest that other features would be reported and acted upon if they had been noticed, or pointed out to the examiner.

Swensson, Hessel and Herman[85] asked 10 radiologists to interpret selected, difficult chest radiographs, then to reread them in conditions which directed the examiner's attention to film regions containing frequently ommitted findings. The authors reported that, although the percentage of true abnormalities reported increased, the corresponding percentage of false-positive findings also increased. They concluded that the reduced omissions in the focussed search experiment were due to less stringent criteria for abnormal findings, rather than an enhanced ability to detect abnormalities i.e. the examiners were not seeing more of the pathologies, they merely labelled more types of appearance as 'pathology' than they had previously. Thus faulty search was not considered to be the reason for abnormalities being overlooked.

2.2.4 Restoration failure

When an approximal cavity is diagnosed, whether or not the cavity should be restored must depend to some extent on the long-term prognosis for the tooth if it is left unrestored or if it is filled. The natural history of the unfilled lesion is important as highlighted in Section 2.2.2. However, the rate of failure of approximal restorations is also germane to the decision process.

Until recent years, the only material used to restore the approximal lesion was amalgam and, for this reason this discussion of the rate of restoration failure will be centred around the survival and replacement rate for Class 2 restorations. There are several causes of amalgam failure, and the relative importance of each is still under debate.

An early investigation which shed some light on the extent of failures among restorations was reported in 1925 by Ottolengui[86], who considered that 51% of the 1067 amalgam fillings which he examined, had failed. However, since many of the teeth examined must have been extracted due to caries, the sample was biased towards a high failure rate. Tingley[87], claimed that more amalgams failed due to poor workmanship than to any other cause, and his view was supported by other authors[88,89]. Healey and Philips[90] demonstrated that incorrect cavity design accounted for 56% of all amalgam failures, and incorrect manipulation of materials, for 40%.

Clearly, with the introduction of new materials, modern methods of cavity design and preparation, and longer clinical training times for dentists, the figures for restoration failure reported in early studies may no longer apply. Harvey[91] examined 1197 restorations involving either the mesio-occlusal surfaces of first molars, or the disto-occlusal surfaces of second premolars, since he considered that these surfaces were most prone to amalgam failure. He showed that 6% of the amalgam restorations had failed. However, these unusually low failure rates are likely to be due to the very low response rate of persons involved in the study, only 25% of whom actually attended for examination. The sample was thus biased in favour of motivated patients, many of whom may have recently completed treatment.

In a large 1967 study Moore and Stewart[92] examined the records of 21,728 teeth in 907 patients. Of these, 39% of teeth contained restorations, over three-quarters of them being amalgam and 4-5% were considered to be defective and in need of replacement. Unfortunately, no criteria concerning the assessment of restorations were given in this investigation.

In a study by Attalla and Gibb[93] in 1968, 560 teeth with approximal amalgam restorations were examined to assess their principal reasons for failure. They stated the majority were due to recurrent caries, but gave no data to support this conclusion. This work is, however, of importance, as it was the forerunner of all subsequent studies which provide vital data concerning the reasons for failure of approximal restorations. The authors also concluded that Black's[94] principle of "extension for prevention" greatly minimised the advent of recurrent caries.

The first statistical study of restoration failure was carried out by Allan in 1969[95], who examined unselected patients attending for conservative treatment, under standardised conditions. Clinical criteria were used solely to determine whether a filling had come to the end of its useful life. The length of time the restoration had been present in the mouth was recorded, but, unfortunately, this figure could only be authenticated in approximately one third of the patients examined. The failed restorations were divided into "false" failures - i.e. teeth destroyed because of the advent of caries in other parts of the tooth - and "true" failures - where the failure could be blamed on restoration quality. Overall about 33% of the 887 restorations examined were considered to require replacement. Amongst the Class 2 restorations, whether restored with gold or amalgam, none had lasted more than 15 years, and many had lasted for much shorter periods.

In 1971, Robinson[96] performed a retrospective study on patients attending a general dental practice, in order to assess the useful life of amalgam and silicate fillings. The study involved examination of 43 patients' records of who had continuously attended a general practice for a period of 21 years. Allan[95] had shown that one quarter of the amalgam fillings failed within 5 years and one half within 10 years. About a quarter of the fillings in Robinson's study lasted for more than twenty years. Although Robinson's study had relatively fewer patients, and causes other than amalgam failure may have produced the loss of some restorations, it was the first to introduce the concept of an amalgam "half-life". Robinson[96] noted that Class 2 lesions had a remarkably high survival rate in his investigation, but commented that this was most probably due to the 'survival' criteria used.

In 1973, a study carried out in Canada[97] showed that the average dentist was replacing 6.6 amalgam surfaces per day. The most important conclusion which these authors drew from their findings was that perhaps Black's principle of "extension for prevention", was not effective, as recurrent caries was the main cause of amalgam restoration removal.

In 1975, Elderton[98] showed that clinicians judged their own work to be excellent - i.e. they tended to assume that the work they performed would be successful in the long-term. Such an attitude among dentists implies that, when the initial decision to treat a tooth suspected of being carious is taken, the dentist overestimates the probability of a favourable outcome if the decision is to restore the tooth. Such an attitude will encourage a dentist to err on the side of over- rather than under-treatment. This attitude is not congruent with the findings that approximately one in three of all restorations at any one time are in need of replacement. Such data give cause for concern, especially as Elderton has also shown that cavities increase in size when restorations are replaced[98]. Hence the placement of a restoration, which is bound to be replaced during a patient's life, should be avoided for as long as possible, as long as delay does not compromise the overall longevity of the tooth.

In 1977, Allan[99] reported on findings from an analysis of practice records in the North East of England, for two periods of fifteen and twenty years, respectively. The data showed that half the amalgams were lost within 5 years, and 90% within 15 years. The author admitted that an improvement in these figures would be likely in the future, since the restorations examined were placed prior to substantial material modifications affecting amalgam durability. Nevertheless, these figures are an important indication of amalgam replacement rates.

More recently Crabb[100] undertook a similar study to those of Allan[95] and Robinson[96], but used dental hospital records rather than those from a general practice. His study analysed 1641 restorations placed by dental students in 155 patients who had been attending for ten years or more. Similar criteria to those of Robinson[96] were employed, and a similar "half-life" approach used to analyse the data. Crabb[100] showed that the half-life for all amalgam restorations was approximately 8.5 years. The five year failure rate was higher than in Robinson's study[96], being about 35%, and the ten year failure rate was 56%.

The differences between the three studies described above[96,99,100] can be accounted for by the fact that the procedures used by a single operator in his own practice should be more consistent than those of numerous operators of varying degrees of experience and competence, as in Crabb's[100] study. Also, the population continuously attending a dental hospital over many years is self-selected and may contain a higher proportion of patients in whom restorations require frequent replacement.

In 1981, Patterson[101] reported the findings of a study carried out in a sixteen-operator practice, again in the North East of England. This investigation involved analyses of the dental records of 200 patients who had attended the practice regularly. Here, 2344 amalgam restorations were included in the analysis and it was shown that occlusal amalgams had a similar half-life to that reported by Crabb[101] for all amalgam restorations. However, for mesio-occlusal (MO), disto-occlusal (DO) and mesio-occlusal-distal (MOD) restorations, a 50% survival time of just over 7 years was calculated. For children under the age of 13, Class 2 amalgam restorations yielded an estimated survival of slightly less than three years.

The results of these restoration failure studies are summarised in Table 2.2. These figures are utilised in the decision analysis described in Chapter 9.

2.2.5 Dental Radiography

Since the ultimate aim of this thesis is to describe optimum treatment decisions based on dental radiographic examination, this review will now deal with the effectiveness, advantages, disadvantages, and limitations of dental radiography.

Traditionally such evaluation has been carried out by comparing clinical and x-ray findings, to histological findings, with the earliest attempt to validate radiographic diagnosis being carried out by Arnett and Ennis[102] in 1933. They reported that clinical examination failed to detect many teeth which were shown, on radiographic examination, to be carious. However the reverse was also true, in that clinical examination determined some teeth to be carious, while the radiographic appearance did not support such a diagnosis. They found that 1425 teeth shown clinically to be carious did not show radiographic evidence of caries whilst for 1135 the reverse was true. The two methods of examination were in agreement for 237 teeth.

In 1941, Burket[103] reported a study in which 920 approximal tooth surfaces were examined by gross microscopic and x-ray examination. He found that 347 surfaces (38%) were diagnosed as carious by one or more examination methods. Using Burket's criteria, the clinical method of detecting caries was more frequently confirmed in the histological examination, than by the radiologic examination. This finding gives credibility to the method proposed for validating the radiographic findings in this thesis, i.e. by means of histological sectioning. In Burket's study, clinical and radiological examinations were only in agreement for 40% of teeth, but the highest agreement in positive diagnoses by all three methods of examination employed in Burket's study, was for approximal carious lesions. Both the clinical and radiographic examination method failed to detect many of the histologically apparent carious lesions. This suggests that Burket's histological criteria included as "carious", early microscopic and clinically unimportant lesions. This makes it apparent that the "gold standard" used to validate methods of diagnosing caries must be considered carefully.

Table 2.2 Rates of failure of approximal restorations

Author	Year	No. of restorations	Time	Half-life	% Failure
Harvey (91)	1926	1197	Unknown		6%
Moore (92) + Stewart	1967	21,728	Unknown		5%
Allan (95)	1969	887	Unknown		33%
Robinson (96)	1971	43 patients	5 yrs 10 yrs		25% 50%
Allan (99)	1977	Unknown	5 10		50% 90%
Crabb (100)	1981	1641 restorations 155 patients	5 10	8.5 yrs	35% 56%
Paterson (101)	1981	2344 restorations 200 patients		7 yrs	

Burket[103] was also one of the first authors to introduce the concept that it is impossible to diagnose all carious lesions i.e. he recognised that perfect sensitivity and perfect specificity in diagnostic screening tests, are goals which are not achievable. This being so, it is important to be able to justify the use of radiographs on clinical grounds. as a dentist exposes his patient to ionising radiation for the sole purpose of gaining clinically relevant information. Trithart and Donelly[104] stated that when approximal tooth surfaces have established contact, a radiographic examination is "a necessary prerequisite" in order to arrive at the best possible diagnosis, in terms of the number of carious lesions and their extent.

In 1968, Webber, Benton and Ryge[105] carried out a study in which seven dentists examined radiographs which had been produced by mounting teeth in a human skull. These simulated radiographs were apparently indistinguishable from those taken routinely in patients. One of the most important findings of this study was that practitioners appeared to agree that few dentists would consider failure to detect an incipient lesion, as being as serious as wrongly diagnosing a sound approximal surface as being diseased. Hence one of the aims of the current study is to ascertain whether this is still the case. Burket's[103] study suggested that the ratio of Type I (false positive) to Type II (false negative) errors is controlled more by discriminatory criteria than by any lack of interpretive aptitude on the part of examining clinicians. In other words, the examining dentists were hesitant to recognise the existence of a perceived, questionable radiolucency. Burket suggested that examiner training would have little influence on the incidence of errors, and secondly that the most efficient way to reduce such errors would be to change the method of radiography, or the viewing conditions. This assertion is not in agreement with studies of radiographic diagnostic accuracy evaluated by ROC curve analysis[76,77]

In 1973, Hefferen[106], for the Council on Dental Materials and Devices in the USA, confirmed that radiographic evidence was the best current means of detecting approximal carious lesions, although he claimed that considerable and important loss of tooth substance could occur before caries was evident on radiographs. This statement is in direct contrast to the findings of Gröndahl[107], who suggested there was a considerable risk of over-prescription if all surfaces with radiologically registered lesions were restoratively treated.

The basis for such a statement is derived from investigations by Rugg-Gunn[50] who found that in 13-year-olds, 65% of all radiographically registered lesions had not reached the amelodentinal junction. In addition, Gröndahl *et al*[55] had found that in 16-year-olds, almost 80% of all radiographically diagnosed lesions were found to be in enamel only, and more than 60% of these were confined to the outer half of enamel. The fact that subsurface lesions, without any break in the continuity of the enamel surface, can be seen on radiographs, has been confirmed by several authors[108-111].

The cumulative evidence highlights the central theme, that is, in the interpretation of radiographs one is dealing with a series of probabilities relating to lesion depth and 'chance' of cure. Since the presence of a lesion is a probability, rather than a certainty, it is impossible to state categorically that a cavity is, or, is not, present, in an individual patient. Nor is it possible to state that a lesion should definitely be treated, solely on the information on lesion depth, gained from a radiograph. The estimated probabilities of disease will vary with the prevalence of the disease under study[112]. If a patient belongs to a selected population in which the prevalence of carious cavities is expected to be higher than in the general population, the probability that a positive radiographic result is associated with a clinical cavity is greater than would be the case in the general population. Thus a decrease in the prevalence of dental caries gives rise to higher demands on the diagnostic methods employed.

The results of earlier studies relating treatment need, cavitation and histological appearance, seem to indicate that radiographs underestimate lesion depth, and that dentinal caries has a high probability of being associated with cavitation[50, 110-118].

In 1983, Mileman, Purdell-Lewis and van der Weele[119] showed that dental teachers, when making treatment decisions from bitewing radiographs, would overtreat teeth for 19% of their decisions (Type I errors), and would undertreat in 36% of their decisions (Type II errors). These teachers were, however, using "caries definitely in outer third of dentine" as their criterion for interventive treatment, rather than that promulgated by Howat and Holloway[120], who stated that the stage of caries requiring restorative treatment is when enamel is cavitated. Presumably Mileman *et al's* dental teachers were using their experience to predict how cavitation is depicted on a radiograph. Howat and Holloway and Brandt[121] did, however, suggest that caries into dentine was a suitable point at which interventive, rather than preventive, treatment should take place.

A recent and comprehensive review by Kidd and Pitts[11] has detailed the precise rôle of bitewing radiography for the diagnosis of early carious lesions. These authors state that many more lesions are detected by radiography than would be found by clinical examination alone. However, although the results of their review confirm this belief, further evidence as to the efficacy of preventive regimes in high risk individuals is required, before routine screening by radiography can be advised as a public health measure. This thesis will not directly address this issue as it is concerned only with the rôle of bitewing radiography as a aid to *restorative*, rather than preventive treatment decision-making.

It is clear from the preceding evidence that decision-making performed by a dentist must take into account, firstly, the general decline of caries throughout Europe, and secondly the demands being made for increasing the time between radiographic examinations for caries[122].

In 1977, the number of dental films used in the UK was around 12 million, almost six times greater than in 1957[123]. Of these, approximately 95% were intra-oral films, although the percentage of those which were bitewing radiographs used to identify approximal lesions is not known. It is also interesting to note that the frequency of dental examinations involving radiographs is much lower in Scotland and Wales, than it is in England[124].

In Scotland, 37 courses of dental treatment per thousand of the population involve radiography, compared to 106 courses of treatment per thousand persons in England. Therefore in 1977, approximately 20% of treatments entailed x-ray exposure. The number of dental treatments involving radiography appears to have increased linearly between 1957 and 1977, rather than exponentially[123]. The UK level of radiography is well below that found in other countries. e.g. Japan in 1974 - 855 per 1000 population; USA. 1970 - 1400 per 1000 population. Sweden 1974 - 1500 per 1000 population. Levels of dental radiography utilisation are of great importance as the patients who undergo radiography are predominantly children and young adults. with 60% of radiographs involving those below thirty years of age. Thus, although the gonadal doses from dental radiography are low. the high proportion of children in the population of patients involved will tend to increase any genetic effects of such examinations. This can be quantified using the genetically significant dose (GSD), a measure which, if given to every member of the population, would cause the same genetic effect as the doses actually incurred[125]. If the mean gonadal dose from a dental film is rated as 0.1 millirem[126], the per caput gonadal dose would be 0.23 millirem, and the GSD about 25% higher. However, this is very low when compared to the GSD calculated from all sources at approximately 110 millirem per year[127]. On the other hand, when somatic, rather than gonadal doses are considered, the per caput bone marrow dose from dental radiography (1 millirem) makes a larger, but proportionally, still small, contribution to the bone marrow dose from all sources (170 millirems)[128]. Although these values are low enough to suggest that the benefits of dental radiography outweigh any radiation risks to patients, the figures highlight the fact that it is important to avoid unnecessary examinations. Also, unless the treatment decisions made on the basis of x-ray examination give a high rate of avoidance of unfavourable outcomes, it should not be used as a test.

In 1983, Wall and Kendall[126] reported that the expansion of dental radiography had continued up to 1981, and that the rate of increase was much in excess of that seen in medical radiography. However it is apparent that much of the increase was in the use of orthopantographic (extraoral) radiography, rather than in the use of intra-oral films. In 1981, approximately 7.8 million dental x-ray examinations were undertaken in the UK, involving 14 million films, and the great majority of these were intra-oral. Similar data, although concerning a shorter time period, were analysed by Hirschmann[129]. The most important finding of Wall and Kendall[126] was that, assuming a linear dose-response with no threshold, the anticipated maximum number of malignancies attributable to dental radiography would be three per year. Since the total cancer death rate in the UK is 140,000 per year, the contribution to this figure from dental radiography is extremely small. It is difficult to compare the benefits and risks of radiography in a quantifiable fashion, but it is clear that as the growth of dental radiography continues unabated, the situation concerning its use and usefulness, should be reviewed. This is particularly true when caries appears to be declining.

Schwartz *et al*[61] have shown that the average duration of carious lesions in each half of the enamel might be expected to be at least 36 months in young teenagers, and over 48 months in older teenagers. In exceptional circumstances, where the individuals are at very high caries risk, with no exposure to fluoride, enamel caries' duration may be in the 18-24 month range. However, as the authors point out, in deciding upon the interval between radiographs, clinicians must weigh-up the probabilities of lesion development in the time between examinations.

It has therefore been recommended that radiographs should not be taken at purely arbitrary intervals. This thesis will question the need for bitewing radiographs for the routine detection of approximal caries, and examine the hypothesis that radiographs provide unmistakable evidence of treatment need.

2.2.6 Relationship of radiographic appearance to caries

This section of the literature review will be divided into two separate parts discussing different aspects of the use of bitewing radiographs for the diagnosis of caries. Dental caries activity has been defined as the "rate at which new caries develops and old caries progresses" i.e. the caries increment plus the caries progression.

Bitewing radiographs may also be used to ascertain the caries prevalence i.e. the number cavities diagnosed given particular criteria for assessing whether a lesion is present or absent, **OR** bitewing radiographs may be used to monitor the progress of a lesion once it is established. In the first section below, the ability of the bitewing radiograph to estimate the probability of lesion as present or absent is examined, and later the relationship between the extent of caries seen on a film, to the actual depth of the carious lesion, is discussed.

As early as 1937, Fixott[130] showed that 25 to 40 per cent more cavities existed in children's mouths than could be seen if radiographs were not employed as an aid to clinical diagnoses, while even earlier studies had shown that the number of lesions detected more than doubled if radiographic examinations were used[131]. From a more practical, treatment-oriented point of view, it has been reported[132] that only 13% of patients could have been adequately orally rehabilitated without the use of bitewing radiographs. Similarly, Smith[133] found that bitewing x-rays revealed additional cavities to those seen in clinical examinations in 68 per cent of patients. In 1945, Barr[134] showed that, of the total number of approximal decay areas, 56% were detected by clinical examination, 93% by radiological examination and 49% by both means. Of the cavities detected by radiographs, 47% were not found on clinical examination.

In Trithart and Donelly's study in 1950[104], the value of the bitewing radiograph for detecting approximal caries was again confirmed. These authors showed that x-ray examination revealed up to 57% more cavities than clinical examination, and concluded that a dental examination was incomplete if bitewing radiographs were not taken. In contrast to these early findings, Jung[108] and Leijon[109] showed that if an approximal area could be viewed by direct vision, more lesions would be detected than by radiographs. However, Trithart and Donelly's[104] findings gave weight to the statement that bitewing radiography provided "an opportunity for very minimal operative interference"[135]. When the article was written, the automatic restoration of every lesion detected became normal practice, and this may still be the case amongst today's practitioners.

In 1970, Marthaler and Germann[110] compared the radiographic appearance of smooth surface lesions to the extent of surface cavitation (a commonly accepted point at which restorative therapy should be instituted), and also to the lesion depth as seen in 3 mm thick, visually-inspected sections. These authors showed that when the lesion depth on the radiograph appeared to penetrate the full thickness of the enamel, 88% of the surfaces had small cavities, whilst when the radiolucencies extended into the inner enamel, cavities were present in 66% of cases. If the lesions were limited to the outer enamel, cavities were present in 34% of cases. This work showed that if all radiolucencies were considered to be an indication for treatment, considerable over-treatment would take place, whilst if only dentinal radiolucencies were considered to be indications for treatment, then under-treatment would occur. This is, of course, only true if cavitation is considered to be the appropriate stage in the decay process for restoration to be undertaken. Thus it is clear that the extent of visible radiolucency cannot be directly correlated to treatment need, nor with the actual depth of a lesion.

Gwinett[135] restated that the radiograph could not always be relied upon to give a true reflection of the carious process, and emphasised that caution should be exercised when interpreting clinical radiographs if used as an adjunct to caries diagnosis. This was a quite different mode of thinking from that of earlier workers who placed great faith in the ability of the bitewing to determine accurately the need for treatment of smooth surface carious lesions.

Rugg-Gunn[50] was the first author to apply Bayesian principles to the diagnosis of caries from bitewings, in that he understood that one can only make estimates of probabilities, rather than lay down hard rules for radiographic diagnosis which lead to correct treatment in all cases. Furthermore, Rugg-Gunn[50] emphasised that the association between approximal clinical lesions and radiological gradings was not close.

Considering all these studies together, the implication is that radiological diagnosis produces only an estimate of the lesion depth. The discovery that, when a radiolucent area has reached dentine, a cavity has commonly occurred, has led to this criterion forming a basis for evaluating restorative treatment - needs in approximal surfaces[136,137]. Although it is agreed that the radiograph is an insensitive measurement of histological changes, especially at the early stages, it is generally accepted that as a carious lesion's depth increases, so does the size of the corresponding radiolucency.

Other methods have been used in order to examine the extent of carious cavities, and to compare them to radiographic images. For example, densitometry has been employed in conjunction with micro-radiography in order to analyse the degree of mineralisation of histologic sections of carious teeth[138,139]. These methods have been refined further in later studies[140,141] which have shown that the demineralisation within a lesion is not always homogeneous, and that radio-opacities sometimes exist within the radiolucent defect. This finding may well account for the observation that the radiographic image does not correlate in a uniform way with the type of lesion seen either clinically, on dissection, or histologically.

The most recent work carried out in the field of lesion depth and radiologic appearance has been undertaken by Pitts[142] who has enhanced considerably the understanding of the architecture of the carious lesion by the use of image analysis. Using 250 μm histological sections as the standard, he has shown that an image analyser has a far greater ability to assess accurately lesion depth, than does visual inspection of a radiograph. Pitts and Renson[13] have also confirmed that lesions can regress as well as progress.

Pitts[143] confirmed that any treatment strategy based on restoring radiologically-diagnosed approximal lesions could not be supported. This thesis will test this hypothesis by examining the $\frac{TP}{TN}$ ratios and probabilities of outcomes contingent upon treating surfaces with various radiographic appearances.

2.2.7 Decision analysis

Analysis of clinical decisions has really only been applied in medical clinical fields in the last 20 years[144-150], and only very recently to dental clinical situations[151-159]. However, review of this work indicates that this quantitative approach to difficult clinical problems warrants careful consideration. It is a tool which may be used in clinical dental practice, both for individual patients, and for analysis of different approaches to particular clinical problems. What decision analysis does is to enable one to gain insight into the outcomes of clinical decisions, without always having to undertake a clinical study, or rely on so-called "expert opinion". It allows a "what if" approach to different decisions using the data available. Although decision analysis has many advantages, it, like all other tools in the clinician's armamentarium, also has limitations. It is the purpose of this review to highlight the possible applications of decision analysis.

Decision analysis was derived originally from operations research and game theory. A 'tree' is constructed which diagramatically represents all the possible options and outcomes from a decision. The decision-tree therefore identifies all the possible choices available to a clinician, and all the possible outcomes of these decisions. Analysis of the decision-tree allows one to calculate the likelihood that each outcome will occur if a particular strategy is employed.

It is clearly a complex and time-consuming task to build a decision-tree which has relevance to "real-world" problems. The full benefit of fulfilling such a task is not achieved if the model simply determines the optimal treatment strategy under the conditions observed. The principle benefit of a decision model is that it has the capacity to answer the question "what if"?

For example, with respect to the problem under analysis in this thesis, it is possible to review the changes in optimal outcome if, say, approximal caries becomes less common, or progress faster or slower than was originally presumed. It can also deal with questions of diagnostic accuracy, i.e. inform us as to likely outcomes if the bitewing radiograph becomes a more accurate source of information, and can also deal with the question as to the risks of early intervention. In addition it can examine the effect of an increase in the longevity of amalgams, or the outcomes which would occur if repeat fillings become less common.

Such questions are dealt with by performing *sensitivity analyses* on the decision model developed. Sensitivity in the phrase "sensitivity analyses" should not be confused with the word "sensitivity" which is used to describe the accuracy of a diagnostic test. The technique is performed by varying the values assigned to one or several variables in a systematic fashion, and then repeating the calculations to determine whether or not they have any effect on the optimal decision.

In its simplest form, a sensitivity analysis involves changing the value of a single variable and recalculating the expected utility of each treatment strategy. This is called a one-way sensitivity analysis. The results of such an analysis can be presented as a graph. e.g. Barza and Pauker[158] investigated the optimal treatment strategy for suspected herpes encephalitis. Their results showed an interesting, and in decision analysis, frequent phenomena, that lines representing the optimal strategy may intersect. In their case, if the probability of herpes encephalitis is less than 42%, then a brain biopsy is required even without further investigation, whilst if the probability of the disease is greater than 42%, the condition should be treated immediately by empirical anti-viral drug therapy.

The levels at which the optimal strategy changes are known as "decision thresholds". However, although threshold values can tell the decision analyst how much change in a given variable is required to change the optimal decision, they do not indicate how much would be gained or lost by pursuing a given strategy. This requires information about the utilities of the various outcomes. Clearly, one-way sensitivity analyses provide only limited insight into a given problem, because they examine changes in a single variable, whilst the other variables are held at the baseline level. The clinician may, however, wish to explore which strategy is optimal when a combination of factors are varied. Multi-way sensitivity analysis is available for this purpose.

The following review of medical decision-making is not intended to be comprehensive, as the field is rapidly expanding. It is included solely to highlight the many diverse fields in which the theory may be applied.

Knill-Jones[159] has utilised decision analysis widely, in relation to computer questioning of patients suffering from dyspepsia; others[160] have constructed decision trees which allow early differention between obstructive and non obstructive jaundice. Examples of evaluation of diagnostic aids using decision theory include the use of Apache-II scores for assessing acute pancreatitis[161]. As another example, Goldman *et al*[162] examined the extent to which a computer protocol predicted myocardial infarction in emergency department patients presenting with chest pain. Decision analysis has also been used in the field of organ transplantation in order to construct prognostic models to assess the value of liver transplantation[163] and, in dermatology, to assess the accuracy of pigmented lesions clinical evaluation[164].

The use of decision theory in audit and health economics[165] indicates the importance of the technique as part of cost-effectiveness studies which are designed to help to decide which treatment is the most cost-effective for a particular patient. Hilden and Mooney[166] have emphasised the importance of decision-making tools at times when strategies for health care undergo radical change. Such changes are notable in the NHS at the present time, particularly in dentistry.

It is clear that decision analysis is one of the few methods of evaluation which allow treatment decisions and their resultant costs to be studied in a scientific way, whereas in the past expert clinical opinion has held sway over health care policy decisions. The first dental authors who examined the role of treatment decision-making in dentistry using a decision analysis methodology, were Settle and White in 1983[167]. The authors highlighted the fact that the frequency, distribution and costs of mistaken diagnoses had hitherto received little attention, and indicated that dentists were probably prepared to accept the risk of false positives, rather than miss the opportunity to treat disease. They also stated that the development of value scales, necessary to weigh the implications of diagnosis, were a long-needed requirement for the development of decision theory in dental fields. Also, in 1983, Douglass and McNeil[151] reviewed the methods available for evaluating diagnostic decisions in dentistry, and in 1987, Mileman's[79] thesis represented the first major contribution of the theories presented by Douglass and McNeil, applied to dental decision making. Mileman has subsequently published a series of papers on the use of bitewing radiography[119,168,169].

Tulloch and Antzcak-Boukoums[152] have applied decision analysis to the controversial problem of impacted third molar treatment. They examined the decision of whether extraction of asymptomatic third molars should occur prior to root completion, or delayed until there is no further potential for eruption, or whether extraction should be carried out only when pathology arose. The authors had difficulty in assigning utilities to the outcomes, which, as they point out, is odd, as most dental interventions take place on an elective rather than an emergency basis. They showed that awaiting pathology was always the risk-minimising treatment option, and that the probability of pathology and the severity of outcome would need to become extremely (and unrealistically) high before this strategy should be altered. They conclude that it is unlikely that clinicians will consider probabilities when choosing treatment options, while they are required during their training to concentrate on acquiring technical skills and biomedical facts. However their analysis is important to the policy maker.

Tulloch *et al*[156] also applied decision analysis to the problem of selecting the optimal threshold for diagnosis of approximal caries. They used a gold-standard of "consensus opinion" to calculate the rates of true positive diagnoses. Such a methodology ignores the inherent interpretative biases of respondents and the fact that caries is a process rather than a diagnosis. However, the authors conclude that no single operative threshold can be considered to be "correct", and that clinicians, when deciding upon treatment options, must consider the prevalence of disease in the population, the likelihood of disease progression, and the consequences of treatment. Although such conclusions may be seen as statements of the obvious, and include all the factors which any clinician might consider and implicitly utilise, this paper[156] is of great importance as it is the first to highlight the importance of making explicit the probability and value of treatment outcomes.

The increasing emphasis on clinical audit, and the heightened awareness of the consumers of health care, make it essential that dental clinical decisions are subjected to scrutiny.

2.2.8 Variation among clinicians in treatment planning

Many investigators[170-175] have reported that there are clinically important differences between dentists' interpretations and evaluations of many tests used in everyday practice. The existence of, and the extent of, the ensuing diagnostic errors are often under-appreciated. It will be seen that different clinicians take different test and patient factors into account when selecting a diagnosis and/or treatment. However, the somewhat idiosyncratic and subjective means by which physicians, surgeons and dentists reach clinical decisions, appear to be independent of the subject matter, or the diagnostic techniques or tools used.

The cause of the inconsistencies between practitioners may stem from two routes. The first is the systematic variation. Dental clinicians may actually hold different opinions of the severity of disease which must be reached before a diagnosis can be made[175]: For example, a dentist who believes that he should restore all carious lesions which show radiographic evidence of dentine penetration, will always fill fewer teeth than a dentist who believes that radiographic evidence underestimates lesion depth and therefore restores all teeth with evidence of any radiological changes in the enamel. The variation between such dentists is consistent, in that the second dentist always intervenes at an earlier stage than dentist 1, i.e. the differences are due to a systematic difference in the diagnosis of caries.

The other type of variation which may occur when clinicians diagnose disease, is non-systematic or random variation. This implies that although dentists may agree on which lesions should be treated, their diagnoses are often affected by factors other than the radiographic criteria by which they judge disease to be present or absent.

It is important to note that observer variation is not equivalent to observer error. Variation is used to describe the differences in diagnoses made by two or more decision makers, i.e. no judgment is made as to who is "right". Investigation of observer error refers to studies which compare diagnostic decisions to a reference test or "gold standard". Most studies which ask a number of clinicians to diagnose pathology, will yield data on both observer variation, and on observer error, if a "gold standard" test is available.

Here, studies of observer variability of diagnoses and treatment planning performed on real patients, (*in-vivo* studies) will be reviewed first. The limitation of such studies is the impossibility of validating the decisions, since it is rare that the teeth *in vivo* will be available subsequently for histological or other "gold standard" diagnoses. In such studies "expert opinion" is sometimes used as a reference test against which the other clinicians decisions are validated. However, since "experts" are probably as prone to errors and inconsistencies to the same degree as other practitioners, this type of validation exercise can be questioned.

To examine variability in diagnosis several studies have examined differences in dentists' treatment plans. In 1979 Rytomaa, Jarvinen and Jarvinen[176] reported a study in which 12 dental school teachers examined the same set of 10 patients (dental students), and made restorative treatment plans for each. Although these authors did not quantify the agreements seen between the dentists, the treatment plans varied dramatically, with the most "non-interventionist" dentist planning treatment for only 31 teeth whilst another planned to fill 72 teeth.

A similar study was carried out by Elderton and Nuttall[1] in 1983, in which 15 dental surgeons examined 18 patients. The range of treatment plans varied from one costing £147, to another involving a sum of £565 (1981 prices). For each filling placed by one of the dentists, on only 40% of occasions would a second dentist elect to fill the same tooth. The Kappa score for surfaces planned for filling, by each possible pair of dentists, ranged from 0.05-0.58.

In his thesis, Nuttall[177] proceeded to analyse the data from the above study and concluded (using majority opinion as a "gold standard" against which to evaluate each dentist's decision) that the sensitivity of a decision to fill was 0.69, and the predictive value of a positive treatment decision was 0.41. Whether such a gold standard is a useful measure with which to evaluate decisions is dependent upon the accuracy of the decisions themselves. It is important to remember that the majority opinion may not be the one which represents the patient's best interests.

In 1985, Hazelkorn[178] in the USA, attempted to examine the extent to which different reimbursement systems would affect dental treatment-planning. Two "actors" were sent into 22 dental practices (11 each) and gave the impression that they were seeking dental treatment. The two types of practices visited were either those providing a fee for item-of-service care, or those providing a "pre-paid" capitation service. This study was particularly interesting as the dentists were unaware that their decisions were being studied. However, a considerable disadvantage of such a study design is that the 'patients' had to record the treatment planned for them. The results of the study showed that dentists practising under the fee for item-of-service system, planned treatment, costing, on average, 50% more than the treatment planned by dentists who operated on a capitation system. Unfortunately these authors did not quantify the amount of agreement between the dentists, but the study shows that dentists vary widely in their opinions as to the treatment a patient requires.

In 1986, Nuttall and Davies[179] compared the results of the 1978 Adult Dental Health and the 1983 Children's Dental Health Survey, with the treatment subsequently provided by dental practitioners. Although epidemiological diagnosis of caries may differ from dentist's views of treatment need, these authors, by taking subsequent treatment as the validating criteria, showed the sensitivity of the epidemiological examinations to be 0.51, both for adults and children. The predictive value of a positive test was found to be 0.31 for adults and 0.24 in children. Although this study examined variations between epidemiologists and practitioners it did not investigate inter-examiner agreement as such.

A few authors have attempted to validate dentists' decision-making by using extracted teeth which can be sectioned and examined, subsequent to dentists diagnosing or planning treatment for them. Merrett and Elderton[2] in 1984 reported a study in which nine dentists examined 228 extracted teeth. The practitioners were asked to record caries and plan restorative care. They showed considerable variability in their treatment planning. The number of sites planned for restoration ranged from 28 to 119. The Kappa scores between pairs of dentists ranged from 0.22 to 0.67. Having examined the variability between dentists, the authors then validated the decisions by sectioning the teeth. The sensitivity of the dentists' decisions ranged from 0.14 to 0.84, while the positive predictive value of decisions to fill ranged from 0.27 to 0.46.

In 1988, Kay *et al*[180] reported on the validity of dentists' decision to restore the occlusal surfaces of permanent teeth. Ten dentists examined "suspicious" fissures in 30 extracted molar teeth. They were asked to plan treatment based on visual inspection of the teeth.

The number of teeth filled by an individual dentist ranged from 6 to 26. Validation of the dentist's' decisions was established by serially sectioning the teeth and examining for dentinal caries. The sensitivity values of the 10 dentists' diagnoses ranged from 0.24 to 1.

A number of authors have examined dentists' reliabilities and variability when examining radiographs. A major study of 845 dentists diagnosing from radiographs was reported by Sewerin and Stoltze[20]. Each dentist was asked to examine three sets of radiographs (46 approximal surfaces). The number of lesions assigned for treatment varied from 0 to 24, and the variation between dentists was not systematic across the three radiographs. Also, the authors found that a dentist who planned excessive numbers of fillings on the basis of one radiograph would not necessarily act in a similar manner with another radiograph.

A similar study, but this time using extracted teeth (thus allowing validation of the examiners' decisions) was reported by Espelid[181] in 1986. In this study 243 dentists examined nine radiographs. On average, 15.6% of the approximal surfaces were reported to be in need of restoration, and 80% of the total treatment planned was proposed by 25% of the dentists. Fortunately, agreement to "fill" improved as lesion severity increased. Decisions to fill sound or enamel-only lesions were made by less than half of these involved. Where cavitation was present, over 50% of the dentists planned a filling. Finally, Mileman *et al*[182] reported a study in which two groups of dental teachers were presented with 12 pairs of duplicated bitewing radiographs and the amount of error was established by validating the decisions against a standard decided by two "experts". These authors demonstrated that there were large variations between dentists in the amount of caries scored. Mileman *et al* ascribe these differences as being due to either differing treatment criteria or varying interpretations of the descriptions given of patients.

It is clear from this section of the review that observer variation has important implications in respect of the value of a diagnostic test, if the public health implications are considered. There is also an important interface between observer variation and signal detection theory.

2.2.9 Evaluation of health outcomes

In the late 1950's and early 1960's, there was a burgeoning of interest in obtaining quantitative measures of ill-health. In 1960 Hinkle *et al*[183] defined the 'seriousness' of an illness as "the likelihood that this episode of illness, or its sequelae, if untreated, will lead to the death of the subject." This scale of 'seriousness' of states of health implies that health can be measured merely by calculating (or estimating) the epidemiological probability of death. Hinkle *et al* then introduced a second concept, - the 'severity' of an illness, or health state, which he defined as "the extent to which a person is unable to carry out his social life." These workers measured the level of severity by using parameters such as the number of days missed from work due to the illness. Such definitions are clearly of limited usefulness in the dental context. For example, a dental clearance may mean only one or two missed work days, and only a minute expectation of death if a general anaesthetic is used. However, a clearance may be regarded by some individuals as having a serious adverse effect on their social and psychological well-being.

Wyler, Masuda and Holmes[184] later undertook a study, in which they attempted to take account of people's emotional responses to a disease. They developed a system of rank order scores for 126 conditions. The authors differentiated between medical and lay respondents, in order that they might determine the differential between 'weights' placed on the same disease states by physicians, and by the general public. Their study analysed the results of 291 questionnaires, in which the overall interpretation of "seriousness" was left to the respondents. Results showed that leukaemia, cancer and uraemia were regarded as highly serious diseases (ranks 126, 125, 124 respectively) while dandruff, cold sores, and bad breath were considered to be of little importance (ranks 1, 3, 6 respectively). An abscessed tooth was rated as the 26th most serious condition, being considered as more severe than laryngitis, ringworm and acne (ranks 16, 17, 22). The authors found a high degree of agreement between medical and lay assessments of the seriousness of the diseases.

In 1973, Patrick, Bush and Chen[185] compared several methods for eliciting measures of social preferences for different health states. This was also one of the first papers to point out that social indicators of the health status of a community would provide criteria for evaluating preventive and health care programmes. These authors showed that health cannot be expressed purely as a state at one point in time. This was also clear from Hinkle *et al's* results[183] in which a condition from which one is likely to recover rapidly, was found to be much less important than a chronic, although milder affliction. Therefore the level of well-being at one point in time, and the probability of transfer to another level of well being (i.e. the prognosis) are both of importance. Patrick *et al*[185] express this concept mathematically, in the form of an equation, which had been developed by others[186,187].

Patrick, Bush and Chen[185] measured social preferences for 29 levels of function, with a view to finding a method of eliciting health preferences which would be both valid and reliable, while being simple enough to use in a household interview survey - an important requirement for any useful method of assessing communities' preferences for health states. The authors' results were impressive and gave details of methods of measuring preferences for health states, yet they concluded that further methodological research should be undertaken before a valid and reliable method of measuring social preferences for health states could be achieved.

In 1972 Goldsmith[188] reviewed the state-of-the-art of the quest for general health status indicators. Like Patrick *et al*[185] he concluded that the methods available included value judgements, and that these were poorly understood, and required considerable further study. He was pessimistic that the future would herald major breakthroughs in the field.

In contrast to Goldsmith, Torrance[189] highlighted the importance of the new quantitative approach to 'health' for health planners and policy makers. However, he raised the question that, although quantitative measures of social preferences for health states were of undoubted importance, the establishment of proper numerical 'weights' to the different functional states previously investigated, was the "weakest link in the chain". He compared three instruments which had been previously used to elicit health states - the *standard gamble technique*, the *time trade-off technique*, and *category scaling*. He showed that all three methods were usable for measuring the value of health states, but that respondents found the category scaling technique to be the most difficult to deal with, despite it being technically, the easiest "interview" to administer. Torrance[189] felt that with adequate sample sizes, both the "standard-gamble", and "time-trade-off" techniques were reasonably reliable and valid instruments for measuring the utility of health states. He showed that the standard gamble and time trade-off techniques were comparable with each other (coefficient of correlation r=0.98) and concluded that the time trade-off method of evaluating health states amongst the general public was comparable to the established standard gamble procedure.

Later, Sackett and Torrance[190] actually measured the utility of various health states as perceived by the general public. They successfully applied a series of 'scenarios' describing the physical, social and emotional characteristics, limitations and duration, of different health states, to a random sample of the general public. By doing so they determined the social utility of the health states. The ten health states they used were described to the respondents as being suffered for one, two or three time-durations. They used a time trade-off technique described in detail in a previous paper[191] and completed 246 home interviews with the sample (82% completion rate: 16% refused interview and 2% broke off the interview). It was shown that short illnesses were always preferable to long ones - i.e. the mean daily utility for a health state fell as the duration of time in the health state lengthened. They concluded it was indeed possible to measure the utility of health states in the eyes of the general public, and that respondent's age, sex and social class only minimally effect the utility assigned to a health state. They also stated that actually being "in" one of the health states described, affected the respondent's assessment of its utility. In their discussion, the authors encouraged other researchers to attempt to measure health state utilities with samples from a general population, rather than by using 'samples of convenience' such as those found at hospitals or other tertiary care centres[192].

The conclusion was that the results could be used in decision-making about health programmes.

In 1979 Pliskin *et al*[193] promulgated the idea that it is preferable to use a systematic, decision analysis approach to medical decision-making, rather than the rough, intuitive and implicit weighting of prolonged survival, against quality of life, currently used by many physicians. Their study concerned the decisions involved when deciding between kidney dialysis and transplantation, and the decision whether to provide coronary artery bypass graft surgery for patients with coronary artery disease. The authors indicated that quality of life could be seen as being related to either relief of symptoms, or return to social activity, such as family life, sexual function and employment. These authors went further in their attempts to place a quantitative measure on health than the attempt which will be made in this thesis to assess dental health state utilities. They not only assessed utility for health states, but also gave this as a function of life-years. Development of tooth utility as a function of life-years requires more detailed research than will be addressed in this thesis.

Pliskin *et al*[193] concluded that although assessment of health state utilities may be useful on a public health planning basis, individual preferences for quality of life rather than longevity, are important when physicians are choosing an optimal treatment for an individual patient. Their lengthy mathematical proof of this theorem indicates that the conventional wisdom of individualising treatment planning is, indeed, valid.

In 1982, the validity of previous measures of health state utilities was questioned[194]. McNeil *et al* raised the possibility that the way information and 'scenarios' were presented to the patient, would influence the preferences which the individual purported to hold. These authors presented data concerning the results of surgery and radiation therapy to patients, students and physicians. They asked respondents to imagine they suffered from lung cancer and required them to choose between surgery and radiation therapy on the basis of life-expectancy data. The study also varied the scenarios, in respect of whether or not the different treatments were specifically identified and whether or not the outcomes were framed in terms of the probability of dying, or framed in terms of the probability of living. Results showed that all respondents preferred surgery to radiation, but only when the procedure was explained specifically, when specific life expectancies were given, and when the outcome was discussed in terms of living, rather than dying.

It is tempting to believe from these data, that since utilities are so sensitive to the way a problem is presented, it is pointless to measure them. However, the authors[194] concluded that these effects are apparent whether or not the physician explicitly or implicitly derived patients' preferences. Further efforts to quantity the utility of varying health states are therefore clearly needed.

Tversky and Kahneman[195], in 1981, explored further the effects of the framing of a decision on the psychology of choice. These authors were aware that the 'frame' or wording which a decision-maker adopts, is dictated by how he/she formulates the problem, and also by the norms, habits and personal characteristics of the decision-maker. They likened alternative framing of decision problems to alternative perspectives of a visual scene. Just as changes of visual perspective can alter the apparent size of objects, changes in the way options are framed can reverse a person's preference between the options. For example, the displeasure associated with losing a sum of money is much greater than the pleasure associated with winning the same amount[196,197]. The authors reported that the effects of framing are large and systematic, and that they occur whether or not the outcomes framed concern loss of human life or choices about money. They therefore concluded that the "framing" issue raises significant questions.

In the same year, Eraker and Sox[198] sought to characterize patients' preferences in choosing therapy decisions. They showed that if a patient is asked to choose between two drugs with equivalent positive effects, (one with two possible (uncertain) outcomes, and the other with a single (certain) outcome), patients will most commonly choose a drug with a lesser but definite, effect rather than choose a drug from which they might have no benefit, but, equally, might gain a very large and favourable effect. When the problem of adverse drug effects was examined, the authors observed an opposite trend, i.e. rather than prefering a certain and moderate adverse drug effect, most patients were willing to risk a possible severe drug effect in order to have a chance of expecting no adverse reaction. This paper supported the notion that, if medical/dental therapy involves hazards or "unpleasantness" it is essential that the decision-maker understands how patients feel about decisions affecting their health. The authors[198] suggested that patient attitudes towards risk were worthy of further research in order that they can be incorporated into formal decision-making. The most important point raised in this paper was that, to be able to elicit patients' preferences, the decision maker must have a **comprehensive** knowledge of the benefits and liabilities of treatments. This is of great importance and relevance to dentists choosing restorative treatments for patients.

The problems of framing are further considered by Byeth-Marom[199] who published a paper which indicated just how vague perceptions of verbal, rather than numerical probabilities are. She found wide variations, even among professional forecasters, in the interpretation of terms such as "poor chance", "doubtful", "perhaps", "reasonable to assume", "likely" etc. These results confirmed an earlier study by Lichtenstein and Newman[200] and have enormous implications for medical/dental decision-makers.

Patients cannot be expected to make rational decisions about treatment preferences on the basis of outcomes descriptions such as "you might need to have the tooth extracted in 3-4 years". Given that the information on many dental treatment outcomes is available, any probabilities presented should be given numerically. Byeth-Marom[199] indicated that verbal probabilities are frequently used because (a) there is an underlying feeling that there is an uncertainty inherent in the prediction of outcomes, or (b) because verbal rather than numerical expressions are preferred as they allow one to defend one's prediction in retrospect (because verbal expressions are so vague). The author therefore suggests there are some phrases which should be assiduously avoided when eliciting people's treatment preferences, e.g. "good chance" as it confuses the *strength* of the probability and the *desirability* of the associated outcome. The author is emphatic that all forecasts of probabilities should be made in a numerical, rather than verbal manner. Such a demand in health-care fields implies that physicians and dentists must firstly participate in continous clinical audit in order that the numerical probabilities of various outcomes are known, **and** that they must acquaint themselves thoroughly with the results of such audit, in order that they may present these realistically to their patients.

In 1982, Llewellyn-Thomas *et al*[201] reported on the use of the standard gamble technique[202] for measuring patients' health values. They stated that the standard gamble technique "is widely regarded as the reference method for measuring values for health"[203]. Their study consisted of seeking the opinions of a panel of 64 individuals about the relative values of several health states. They utilised rating scales and three types of standard gamble techniques in their study. The raters gave highly reproducible results in replicate gambles. An important conclusion in terms of the thesis presented here was that the "worst health outcome" used in the standard gamble technique, should be presented as the worst possible outcome relating to the clinical situation considered, rather than death being used as the worst outcome, when it is not a possibility.

Decision theorists[204], applied psychologists[205,206], and health care specialists[207] have examined both the theory and procedures underlying the development of mulitiattribute health indices.

Boyle and Torrance[208] reviewed the literature relating to the development of *Multiattribute Health Indexes*. They pointed out that such indices of health had not been truly recognised in applied health care research. This has now changed to some extent, except perhaps in the context of applied **dental** health-care research - a situation which it is hoped this thesis will rectify.

Further work considering the influence of framing, medium and rater variables has been reported by O'Connor *et al*[209]. They confirmed that, among 216 university nursing students, preferences for alternative cancer drug treatments were not significantly dependent on the sex, age or professional status of the respondents, nor on the medium used (computer vs pen+paper questionnaires.) However they confirmed Tversky and Kahneman's[195] work concerning the importance of the way the decision problems were framed. The word "survive" was found to account for much of the framing bias. Although this was an early work on the problems of patient decision-making, the authors conclude that it was vital to avoid negative words and phrases when describing treatment outcomes to patients. Again this empirical proof of a conventional wisdom can be taken to indicate the validity of the technique. Clinicians do, and should continue to, present outcome data in a realistic, but hopeful manner.

In 1985 Kirshner and Guyatt[210] attempted to describe all the methodological considerations which should be made when developing and assessing health indices. The burden of the authors' theme was that the method of constructing an index should depend on the purpose for which it was to be used. They divided potential uses for health indices into three broad categories i.e. discrimination, prediction and evaluation. The authors suggest, therefore, that those who wish to develop new instruments for measuring quality of life or health, should tailor the items chosen for inclusion, and the subsequent questionnaire content, according to the primary purpose of their research. This, therefore, is precisely the methodology used in the current thesis, as there is no existing dental health index which has confirmed reliability, validity and responsiveness for the purpose of eliciting a differential measure of utility of various "tooth states".

Torrance[211] reviewed the use of health state utilities for economic appraisal. He implied that although health state utilities could be measured, this was usually done with some imprecision. However, the author points out that this was not an insurmountable problem, as large sample sizes and careful use of sensitivity analysis can reduce the inherent imprecision. Torrance felt that health state measurement was an important tool for economic evaluation of health care programmes and commended the approach to health care planners. He felt that contributions to the research pool in the field were important, not only for the direct research benefits in each particular study, "but also for the increased knowledge and improved techniques which come from further shared experiences".

Evaluation of oral health states has not been given the amount of interest which measurement of general health status has received, nor has it benefitted from serious research into the matter.

This problem was first brought to attention by Nikias *et al*[212] in 1978. These authors were the first to highlight that there was a requirement for an index of oral health states which would draw together the disease processes acting in the mouth, i.e. aspects of pain, function, well-being and psychological, social and behavioural dimensions, into a single meaningful concept which could be expressed numerically. They did however recognise that the requirement of such an index would vary according to its intended use.

Following this early work, Nikias *et al*[213] collected oral health data from 1,290 adults. They attempted to incorporate this data into eight distinct categories of an oral health profile, by employing criteria which related the categories' significance to oral health and its relevance to the adult population. Each category was then assigned four grades. However, the attempt to reduce the data into a simple numerical index of oral health met with limited success as the minimum number of profiles to which the data could be reduced was 42. The authors then pursued their objective of developing an oral health index, by asking panels of dental professionals to rank profiles of oral health status. This was done using the method of scaling known as "equal appearing intervals". The overall index of oral health status was based on the mean rank, assigned to each oral health profile by 29 dentists. However, again, the study met some difficulties, as the dentists involved reported great difficulty with the ranking process, and showed considerable variation in their assessment of clinical oral states. The final study reported by these authors used a paired preference technique to enable 12 judges to rank oral health status profiles. The authors felt this methodology allowed the development of an index which would be useful for both outcome assessment in quality control programmes, and cost effectiveness analysis, and could therefore be utilised in the planning and evaluation of dental care delivery programmes. Basically, Nikias *et al's*[212,213] work translates 42 patient profiles into a concept akin to the Index of Well-Being developed for measuring general health[185,214].

Marcus *et al*[215,216] pointed out that the available indices of oral health had an inherent weakness, in that they were not mathematically based. These authors also used a pair-preference methodology (analogous, although not directly comparable to the von Neumann Morgenstern standard gamble technique). In their study, 12 dentists acted as the "judges" on 232 patient cases. They found that three variables could capture almost all the information which influenced the ranks finally given to the oral health status profiles. The authors felt their findings allowed mathematical weighting of the factors which contributed to oral health, and that the results offered a mechanism by which programmes and decisions affecting dental public health could be rationalised. It must be realised however, that all these methods of developing oral health indices were based on dentists' clinical judgments of the health of the individual profiles which they examined. This, although useful, does not assist in the search for a numerical expression of how the population assess their oral well-being. Nonetheless Cushing *et al*[217] have made an attempt to measure the social impact of dental disease. They emphasised that the assessment of an individual's dental status should include some measurement of the social and psychological impact of dental disease. The aim of their study was to develop indicators of the impact of dental disease, as experienced by the population, in terms of the pain, anxiety and dysfunction which it caused. They found that almost three-quarters of their sample had experienced social and psychological impacts as a result of dental disease, the most common of these being discomfort. For this reason, it is imperative, that in the search for utility measures of oral health, the possibility of the advent of pain is taken into account. These authors also found that, apart from in relation to eating problems, DMFT was **NOT** well correlated with measures of impact. Therefore measuring disease using the traditional DMFT index does not measure the importance of dental disease to the population. Most importantly, in Cushing *et al's* study[217] there were no differences in the impact which dental disease had on peoples' lives, between those who attended, and those who did not attend a dentist regularly. It therefore appears that the dental care system in the UK does not significantly influence the prevalence of dental problems as perceived by the patient - although regular attendance may influence the individual's state of health, as perceived by the profession.

Finally, in recent years, two studies have been published which attempt to measure the significance of dental health problems, in relation to general health. In 1988 Westert *et al*[218] questioned 109 subject on their general health status. All were known to have experienced dental problems in the six months prior to the study. Only seven, however, spontaneously mentioned dental problems. The results of the study showed that the impact of most dental problems on a person's feelings of health were small. The findings therefore indicate that the assessment of oral health utilities should be separated from the measurement of general health.

However Reisine[219] disputed the results of the two studies reviewed above. She suggested that the impact of dental health on a person's quality of life was important and pervasive, and she cites as evidence for this conclusion, the fact that as much as one third of the employed population cannot function at work for limited periods because of dental problems. Reisine[219] suggests that, although quality of life is difficult to measure, dental problems with pain, aesthetics and function contribute significantly to a person's quality of life.

It can be seen from this review of the literature pertaining to oral health utilities, that little work has contributed significantly to a satisfactory measurement of social preferences for dental health states. Further effort towards the development of such an instrument is crucial if dental service planning is to be rationalised.

2.3 CONCLUDING COMMENTS

This chapter has reviewed the literature which will provide data for the decision analysis to be presented in chapter 9. It has also examined studies which have previously utilised the theories which are to be developed in this thesis.

The changing prevalence of dental disease is of importance as it affects the performance of a diagnostic test. The rates and probabilities of progression, and also of regression of lesions, determine the utility of the outcomes of decisions not to restore perceived caries, just as the probability of restoration failure affects the utility of outcomes of decisions to restore lesions.

Signal detection theory enables separation of the visual task of detecting a carious lesion, from the decision process which leads to a tooth being restored.

Dental radiography has been reviewed as its use is continuing to increase, without, it must be said, rigorous evaluation of its use and usefulness as a diagnostic tool. As caries declines the benefit of radiography is concomitantly reduced. The costs of radiography increase with high usage and the reduced efficacy of the test when disease prevalence is low.

The review also emphasises that radiography is not particularly accurate as a predictor of carious lesion depth. This implies that the test is best evaluated in relationship to the outcomes which occur as a result of it use, rather than its ability to detect a lesion of a particular depth.

The process of decision analysis has been reviewed, along with a synopsis of the uses to which it may be put, in order that the context of using this process within dentistry can be appreciated. Recent changes which have taken place in the UK, with respect to the organisation of dental health care, encourage the idea that the responsibility to show a test or intervention is required, lies with the dentist. Unless treatment decision-making is subjected to decision analysis in which all the possible outcomes are described and quantified, such 'proofs' of need are impossible to achieve.

The variation in treatment planning among dentists has been reviewed, as this thesis will attempt to explain these differences in terms of factors other than 'errors'.

Finally,the difficult subject of the evaluation of health outcomes has been reviewed, as it is upon the theories and methods developed in medical fields that the assessment of dental health outcomes undertaken in this thesis lies.

In each of these fields, the volume of literature is enormous. This review has attempted only to highlight key-note papers and recent advances in each field.

CHAPTER 3. AIMS, RESEARCH QUESTIONS AND THEORETICAL BACKGROUND

3.1 INTRODUCTION

This chapter details the aims and research questions which are to be addressed. Chapter 1 has introduced the concepts underlying the research questions, and has delineated the approach which has been taken to the problem of dental treatment decision-making. The specific methodologies of the studies which were undertaken to explore the research questions are detailed in the relevant chapters (Chapters 5-9). This chapter, however, examines the theoretical background underlying the specific methodologies used.

3.2 AIM

The aim of this study is to ascertain which radiographic thresholds for dental treatment decisions maximise the health benefits and minimise the health costs of treating approximal carious lesions.

3.3 OBJECTIVES

The objectives of this thesis are to answer the following research questions.

1. Given a specific radiographic appearance of an approximal lesion, what is the probability of the progression of the lesion within a given time period? (Chapter 2).

2. Given a Class II restoration, what is the probability that the restoration will require to be replaced within a given time period? (Chapter 2).

3. What materials must be used in order to construct simulated dentitions to give a radiographic appearance similar to that of a bitewing radiograph taken *in vivo* ? (Chapter 4).

4. What is the relationship between the visual appearance of an approximal surface, the microscopic appearance, and a dentist's interpretation of the radiographic appearance of a tooth? (Chapter 5).

5. How "accurate" are the treatment decisions made about carious lesions by practising dentists on the basis of bitewing radiography? (Chapter 5).

6. Does a dentist's rating of the "costs" (in health terms) of a treatment decision affect his treatment threshold? (Chapter 6).

7. How and why do treatment decisions for approximal lesions, which are based on bitewing radiography, vary between individual dental practitioners? (Chapter 7).

8. What values do the general population place on the outcomes of dental disease and its treatment? (Chapter 8).

9. What are the public health implications of dentists' decisions about treatment for approximal lesions? (Chapter 9).

3.4 THEORETICAL BACKGROUND

3.4.1 The probability of lesion progression

The progression of carious lesions requires longitudinal population studies. These are reviewed in Chapter 2 (Section 2.1.1, Table 2.1). Such studies allow conclusions to be drawn as to the maximum and minimum expected rates of carious lesion progression and analysis of these results provides a means whereby the 'average' results from all the studies together is assumed to represent the most accurate estimate of the lesion progression rate.

3.4.2 The probability of Class II restoration failure

When an approximal carious lesion is deemed to be in need of restorative care, the treatment of choice is the placement of a two, or three-surface amalgam filling. As highlighted in Section 1.2, this type of restoration is prone to failure. It is therefore, necessary to take this failure rate into account when examining the outcomes of dentists' treatment decisions. Studies examining the failure of Class II restorations are detailed in Chapter 2. (Section 2.1.4 and Table 2.2.)

3.4.3 Simulating bitewing radiographs

The radiographic properties of a bitewing film affect the diagnostic and decision-making abilities of an observer. It was therefore necessary to determine a material which had a similar radiodensity to bone, using microdensitometric analysis of test and control radiographs. A further requirement of the films to be used in this research, was that they should generate radiographic "noise", such as that produced by the periodontal membrane, lamina dura, etc. in a bitewing film taken "*in vivo*". Finally, the x-ray beam used in the production of simulated bitewing films should undergo the same beam alteration as the x-rays used when taking radiographs of patients.

These parameters required several studies before bitewing radiographs with appropriate properties were produced (see Chapter 4).

3.4.4 The relationship between the "actual" state of a tooth surface and dentists' interpretations of the radiographic appearance of the tooth

If dentists undertake treatment decisions about dentitions which are represented on simulated bitewing radiographs, it is possible to subsequently evaluate these decisions, if a 'gold standard' diagnosis of the true condition of the tooth surface is available. This may be achieved, either by visually inspecting the tooth surface for evidence of caries, or by examining sectioned teeth microscopically.

Using such 'gold-standards', it is possible to determine the number of lesions correctly deemed to be in need of treatment (true positives (TP)), the number of caries-free tooth surfaces which were to be restored (false positives (FP)), the number of caries-free surfaces which are to be left unrestored (true negatives (TN)), and finally the number of carious lesions which are left untreated (false negatives (FN)).

The true positive percentage (TP%) is:

TP% = Number of truly carious surfaces restored x 100, divided by the number of carious teeth restored plus the number of carious surfaces left untreated.

Likewise:
FP% = Number of sound surfaces restored x 100, divided by the number of sound surfaces left unrestored plus number of sound surfaces restored.

True negative and false negative percentages can be calculated in the same way. The sensitivity and specificity can be calculated from the data laid out in a 2 x 2 table as shown in Section 1.3.1.

An ROC curve is constructed by plotting the rate of true positive and false positive decisions at a number of treatment thresholds. ROC curves therefore allow the accuracy of a test at any given threshold of diagnosis to be ascertained (see Figure 1.2).

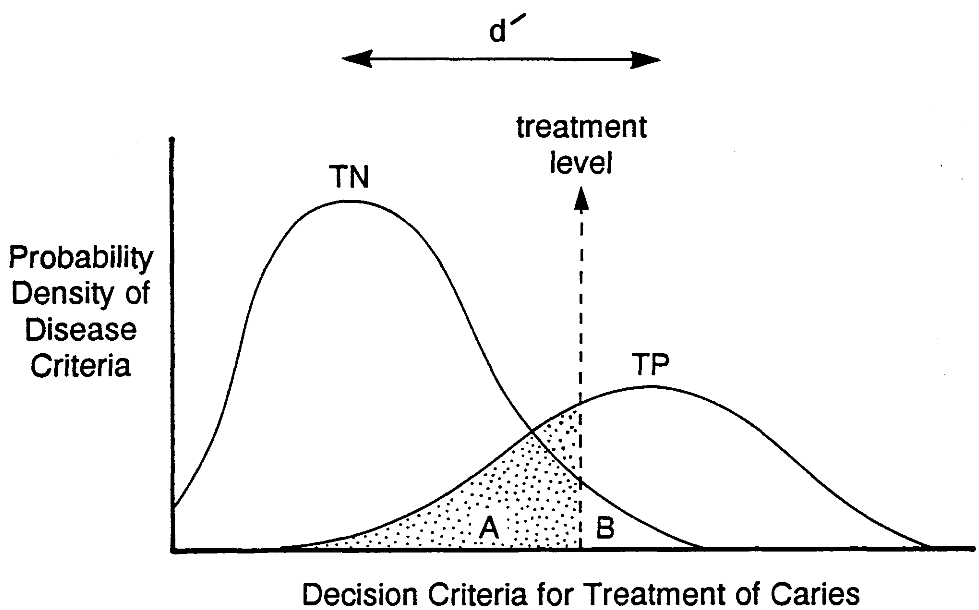
In Figure 3.1 area A represents false negative treatment decisions and area B false positive treatment decisions, and by shifting the diagnostic criteria the relative size of these areas is increased or decreased. The ROC curve is a method of plotting these changes as the diagnostic threshold is altered.

3.4.5 The accuracy of dentists' decisions

The "accuracy" of dentists treatment decisions may be evaluated using measures such as sensitivity and specificity as described in Section 1.3.1.

Figure 3.1

Diagram indicating the detectability of a carious lesion



A more meaningful measure. of "accuracy" is the *'predictive power'* of positive and negative decisions. The predictive power of a positive treatment decision is the proportion of all tooth surfaces designated for restoration, which are truly carious.

This value will alter depending on the 'gold-standard' criteria used to define "truly carious". It will also depend upon the definition of a positive treatment decision. The predictive power of a positive test is also known as the posterior probability of disease. Its value is related to the prevalence or "prior probability" of disease.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \qquad = \text{true positive rate}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$1 - \text{Specificity} = \frac{FP}{TN + FP} \qquad = \text{false positive rate.}$$

$$\text{predictive value (positive)} = \frac{TP}{TP + FP} \qquad = \text{posterior probability.}$$

$$\text{Prevalence of disease} = \frac{TP + FN}{TP + FN + FP + TN} \qquad = \text{prior probability.}$$

3.4.6 Dentists' values in relation to treatment thresholds

The setting of treatment thresholds are dependent. not only on the dentists' knowledge of the natural history of a lesion, but also on the importance he attaches to always treating carious lesions and never treating caries-free teeth.

Therefore the proportions of TP, TN, FP and FN decisions made by a dentist who is acting at a given treatment threshold are mathematically related to the 'value' placed upon each by a practitioner.

It can be shown that the TN:TP ratio derived from a dentist's decisions is related to the slope of the ROC curve at that point on the curve. The relationship is determined by the dentist's views as to the costs associated with 'false' decisions, and the benefits associated with 'correct' decisions.

The ROC curve generated by a dentist allows computation of each dentists views about missing, and overtreating lesions, without asking the dentist to explicitly state his attitudes.

However, by comparing a dentist's stated attitudes concerning the costs and values of treatment decisions, to the values derived from his ROC curve, it is possible to determine whether or not a dentist makes treatment decisions according to his stated attitudes. In other words if the implicitly and explicitly derived values are incongruent, then the dentist must be holding mistaken views as to how a pathology is represented on a radiograph.

3.4.7 The variations in treatment decisions between individual dentists, and the reasons underlying them

For each possible pair of dentists involved in this study a 2 x 2 table of the decisions can be constructed

		Dentist A	
		No Treatment	Treatment Required
Dentist B	No Treatment	a	b
	Treatment Required	c	d

Such a table classifies each of the treatment decisions made by two dentists into four categories. In cell (a), both dentists agree that the surface does not require treatment, in cell (d) both agree the surface does require treatment. In cell (b), dentist A assigns the surface for treatment, while dentist B does not, and in cell (c), dentist B assigns the surface for treatment while dentist A does not.

From this table, the proportion of agreement can be calculated as $\frac{a+d}{a+b+c+d}$

Also, Cohen's Kappa - as described in Chapter 1, can be calculated. This gives an assessment of overall agreement which takes into account the agreement which could have arisen by chance.

In the study reported in Chapter 6 the Kappa value was calculated for each dentist pair. The Kappa score between pairs of dentists holding similar and dissimilar views on treatment philosophies could then be compared, in order that the source of disagreement could be determined. If a pair of dentists seek to fill the same depth of lesion, and yet are found to disagree profoundly when actually making treatment decisions, the source of the discrepancy must either be their treatment threshold, or simple errors in interpretation of the radiograph.

Overall, Kappa scores were calculated for all dentists and for each tooth surface in order that the tooth surfaces causing greatest discrepancies between dentists could be determined.

3.4.8 The values placed on dental health states by the population

Chapters 4-7, and the studies described therein are concerned with dentists, their treatment-decisions, and how their value systems might affect the treatment which they offer to their patients. However it is unlikely that a dentist can correctly weight a true positive or true negative decision appropriately unless he knows how his patient feels about the outcomes of dental disease and dental treatment.

The study described in Chapter 8 was carried out in order that a 'value' could be placed on a number of dental health states. The worst outcome for an unhealthy tooth is a need for extraction, whilst the best health outcome which might result from a dentist's decision, is that a tooth which is perfectly sound remains in that state. The value of these outcomes to the individual may affect the treatment thresholds held by dentists. Therefore it is important to attempt to ascertain their value. This was carried out by using both a visual analogue scale and a standard gamble technique.

The standard gamble can be considered to be a method of measuring the risks an individual is prepared to take in order to avoid poor health outcomes and increase the probability of favourable health outcomes.

3.4.9 The public health implications of dentists' treatment decisions

The first step in applying decision analysis involves identifying all possible choices and all potential outcomes of a clinical decision. For example, what will happen to a tooth if it is left unrestored, and what will be its fate if it is filled?

A decision tree consists of a sequence of nodes, each of which describes the choices available and the chances of particular outcomes. Thus, by following a particular route down the tree, when a particular stratagem is employed, it is possible to calculate the likelihood of each outcome. The relative values of each of the outcomes in the case of the current research, were ascribed according to the population utilities derived as described in Section 3.4.8. Hence, a score of 0.5 might be assigned to a well-filled tooth, on a scale where a sound unfilled tooth is defined as 1.0 and a tooth requiring extraction is defined as zero. The utility of each "CHANCE" node is calculated as the weighted average of the outcomes which arise from it, where the weights are the probabilities that each outcome will occur. The best choice at a DECISION node is to take the strategy with the maximum expected value of the outcomes which follow it.

In the following example (Figure 3.2) a square, or DECISION node represents the dentist's choice of treatment. From this stem, branches representing the strategies taken by the dentist, lead to chance nodes, represented by circles. At the circles the branch divides into sub-branches, one leading to each possible end-state for the tooth. Therefore the ends of the branches represent all the ultimate possible consequences of a dentist's decision about a tooth surface.

In the example on the following page, the probability that an unrestored tooth will continue to develop caries is 0.5, and that the caries will remain static is also 0.5. The probability of restoration failure is 0.2 and the probability that the filled tooth will remain sound is 0.8. Hence, it is possible to analyse the decision of whether a tooth should be restored. From this example tree it can be seen that the expected utility (i.e. the value of the outcomes which can be expected from adopting particular strategy) of a decision to restore a tooth is

= (0.2 x 0.1) + (0.8 x 0.6)

= 0.5

and the utility of a decision to leave the tooth unrestored is

= (0.5 x 0.4) + (0.5. x 1.0)

= 0.7

Since 0.7 is the greater, in this situation the tooth should be left unrestored.

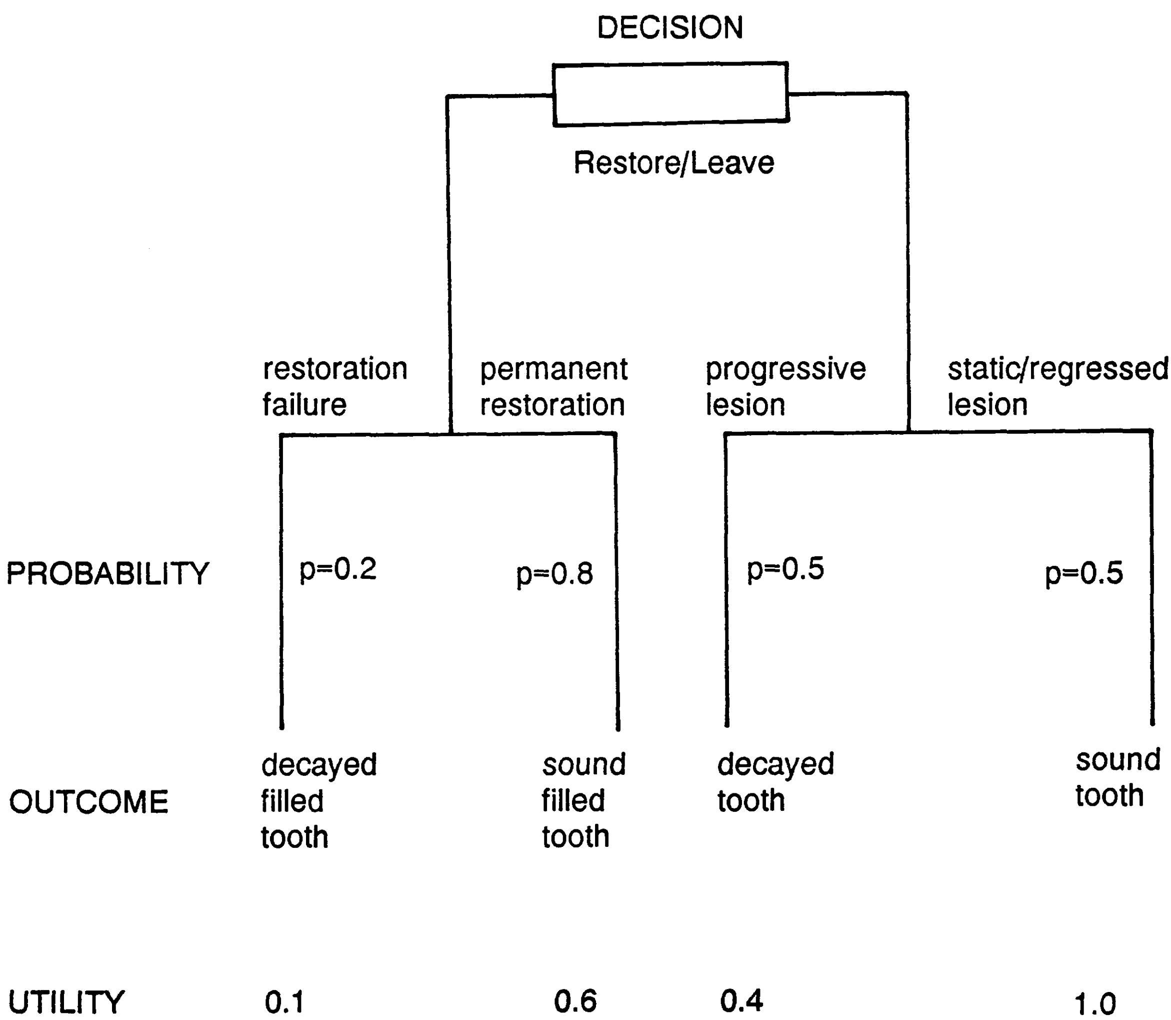
The full benefit of constructing the restore/leave unrestored decision tree is not, however, obtained by simply using the model to determine the optimal management strategy. The most important advantage of constructing the tree is that it has the capacity to allow variation in the assigned values in a systematic fashion. This enables one to determine whether the optimal decision changes when circumstances are altered. For example, the optimal decision may vary when dentists' decision thresholds change, or water is fluoridated, or disease is less prevalent or perhaps when the longevity of restorations is increased. This type of variation can be accounted for by performing sensitivity analyses for each set of external factors.

3.4.10 Overview

In order to achieve the overall stated aim of this research, it has been necessary to utilise methodologies and theoretical concepts which may, at times appear to be somewhat disparate.

Figure 3.2 Example of Decision Tree

(The probabilities and utilities shown here are for the purpose of example only)



This chapter has aimed to describe in outline the theoretical background which allowed the research questions to be answered. The detail of the methodology employed in each part of the research is described in the relevant chapters, but a brief and simple synopsis was held to be necessary in order that an overview of the research could be gained before detail is added.

The information gathered in this thesis was used to structure a decision tree. This analysis aimed to integrate the ROC analysis with the outcomes which could be expected from the decisions made. It is important to note that the tree was structured with the depths of lesions which were actually present given a particular treatment decision. This was because the lesion depths were determined after the decision had been made. This is a different arrangement to the figure shown on the previous page which was constructed to show the principles underlying any decision tree, and not the exact form used later in the thesis. Thus, in the main analysis, the depths of lesions are conditional upon the decision. It would also be possible to construct a tree in which the decision was conditional upon the depth of lesion. However the aim was, in both the ROC and decision analyses, to examine the outcomes contingent upon the dentists' current practices, rather than examining how information concerning lesion depth affected treatment decision making.

CHAPTER 4. CONSTRUCTION OF PHANTOM MODELS AND SIMULATING BITEWING RADIOGRAPHS

4.1 INTRODUCTION

Several studies have examined the effectiveness of bitewing radiography as a diagnostic tool[117-119], but those which used simulated bitewings, as opposed to those being taken in patients, are limited. Clearly, if the sensitivity and specificity of a diagnostic system are to be tested, it is necessary that a 'gold standard' of lesion depth can be established. This requires that the bitewings are of teeth *ex vivo*, in order they can be subjected to microscopic examination, subsequent to the dentists' examination of the radiographs.

It is known that the radiographic properties of tissues surrounding the lesions in question affect the observer's ability to detect pathological lesions[83,84]. It is important therefore, that in any study investigating dental practitioners' treatment criteria for approximal caries, that the films simulating the clinical situation in question, must have the same radiographic properties and radiographic density as those taken *in vivo*. This chapter therefore describes the construction of the models of which the bitewing radiographs used in the studies were taken.

Pitts[223] employed simulated bitewings in several studies, but unfortunately he does not specify the materials used to mimic bone or soft tissues, nor does he describe in detail their radiographic properties. Mileman[79] also simulated bitewing radiographs. He used perspex to represent the effect of soft tissues on x-ray attenuation, but did not attempt to simulate the bone surrounding the teeth, using only impression materials to mount them.

The information content of a radiograph is influenced by the passage of x-rays through various tissues. It is therefore important that any study designed to examine the diagnostic capabilities of radiographs, simulates natural conditions as far as possible. The diagnostic information available from radiographs is also influenced by the technical conditions under which they are taken. The tube current (mA) and the exposure time (sec) both affect the quantity of radiation emitted, and the tube voltage (kV) influences both radiation quality and quantity. Radiation quantity affects the density of the film, whereas quality affects contrast. The bitewing radiographs used for this study thus required several properties.

It was necessary that the radiographs appeared to be similar to those taken *in vivo* in order that the dental practitioners' visual search of the films resembled those in normal practice. The conditions required for this to be so are:

1. The tube current (mA), exposure time (sec) and tube voltage should be identical to those used when taking a normal bitewing radiograph in a patient.

2. The materials surrounding the teeth and those intervening between the x-ray tube and the teeth should have, as far as possible, the same beam attenuating properties as patients tissues.

4.2 PREVIOUS WORK ON SIMULATION OF BITEWINGS

Most studies in which phantom materials are used to simulate the radiographic properties of human tissues are concerned with dosimetry. It is unfortunate that in many of these the construction of the phantoms used is referred to in vague terms which do not allow the reader to assess the extent to which the materials used simulated the real-life situation. For example, Oishi and Parfitt[221] refer to a material which "simulated soft tissue", without ever specifying the composition of the substance used.

In 1935 Hodge *et al*[222] showed that aluminium had the characteristics, when placed in an x-ray beam, of dense cortical bone, while Wuehrmann and Curby[223] used ivory. Later studies have moved away from these solid materials and have used simple plaster mixtures[224]. Arnold[225], in 1983, employed a real skull to mount the teeth for his investigation, thus obviating the problem of simulating bone.

The tissues other than bone and teeth which needed to be simulated were the surrounding soft tissues i.e. the periodontal space, the gingivae and the cheek. Water is the major constituent of soft tissues and is often mentioned as an adequate substitute for radiographic purposes. Its use presents severe practical problems but nevertheless has been used in the past. Webber *et al*[226] made a human face-shaped container and filled this with water, whilst Hedin and Halse[227] placed water in a holder and placed this between the object and the x-ray source. With this latter method, the beam attenuating properties of the container must also be considered. The practical difficulties encountered when using a liquid phantom medium can be minimised by using solid material of equivalent radiographic density such as that reported in 1949 by Jones and Raine[228].

Paraffin has also been used as a soft tissue simulator[229,230] whilst others[231] have used a mixture of wax, paraffin and rosin. Wax alone is particularly useful in inaccessible areas such as the periodontal space and interdentally[225,232]. Lastly, perspex and other plastics have been shown to be suitable soft tissue simulators[233].

4.3 METHODOLOGY

Several stages were involved in the construction of the phantom radiographs:

1. Determination of a material which had the same radiodensity as bone, for the mounting of teeth.

2. The construction of models.

3. Determination of a soft tissue simulator.

4.3.1. Test 1. Determination of a material with similar radiographic density to bone

Method

Studies were undertaken in which blocks of various materials considered to have potential as a bone replacement material were radiographed.

Several standard radiographic views of a real human mandible were then taken (Figure 4.1). A fully immersed impression of this entire edentulous mandible was then taken using an impression material. Models of this mandible were then cast (see Figures 4.2(a)(b)(c)) using various combinations of self-curing plastic and bonemeal, as these were materials assessed from the first stage radiographs as having the closest radiographic resemblance to bone. The combinations of plastic and bonemeal explored were 1:1, 2:1 and 3:1 ratios.

These models were then radiographed using the same technical specifications as those used to radiograph the real mandible. Each set of radiographs (Figures 4.3(a)(b)(c)) were then subjected to microdensitometric analysis.

For each of the seven "intraoral" radiographs from each simulated mandible, and from the control (real mandible), an area of 0.8 x 0.6 mm was randomly selected. For each corresponding area in each radiograph the microdensity of the test and control films were recorded. The differences between the model and real radiographs were compared using Students t test.

Results

Shown in Table 4.I are the microdensitometric measurements for each radiograph from each of the real and simulated mandibles. The data reveal that there were no significant differences between the radiographs from the model ($p > 0.05$) with a ratio of 2:1 plastic

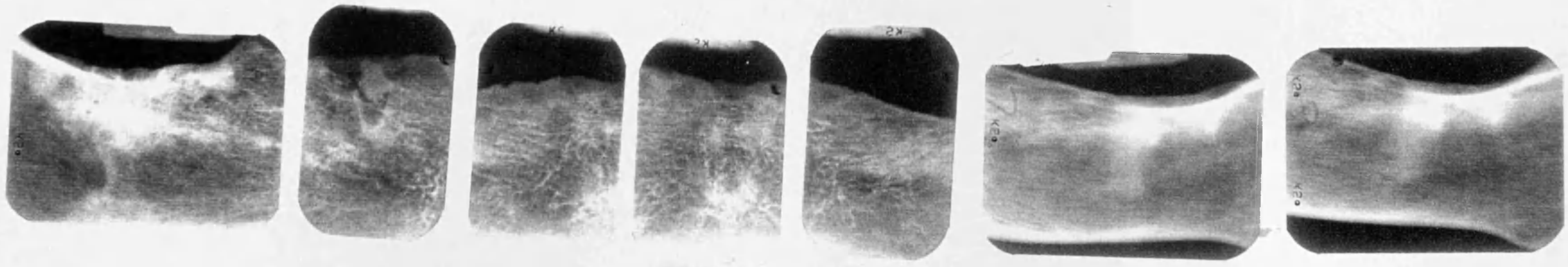
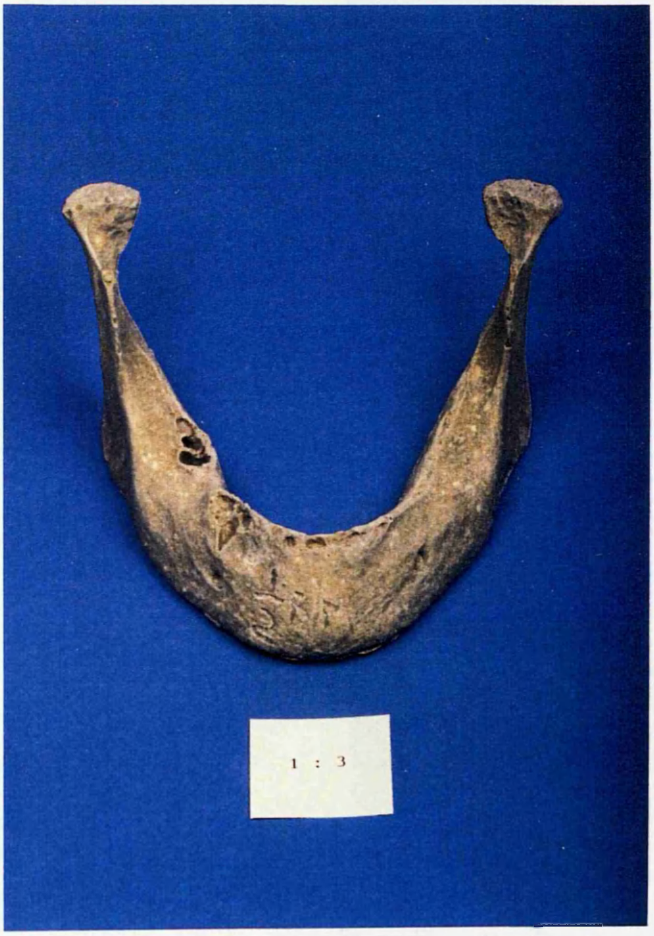


Figure 4.1 Radiographs of edentulous mandible

Figure 4.2(a) Simulated mandible 1:1 ratio acrylic : bonemeal

Figure 4.2(b) Simulated mandible 1:2 ratio acrylic : bonemeal

Figure 4.2(c) Simulated mandible 1:3 ratio acrylic : bonemeal



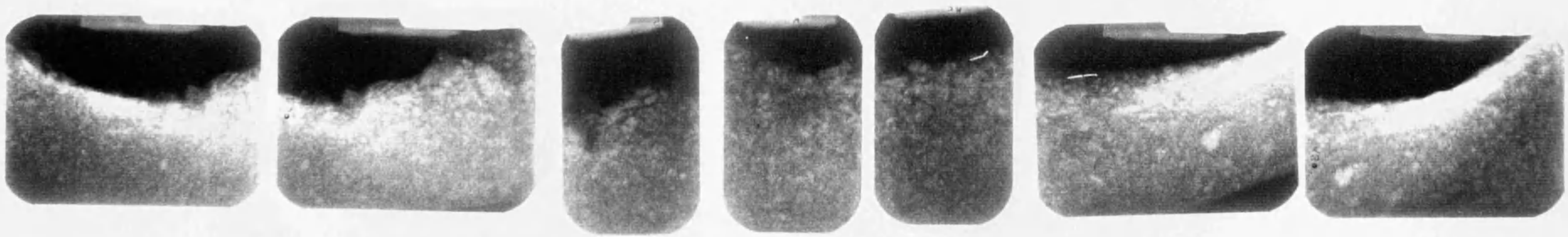


Figure 4.3(a) "Intra-oral" radiographs of simulated mandible 1:1 ratio acrylic : bonemeal

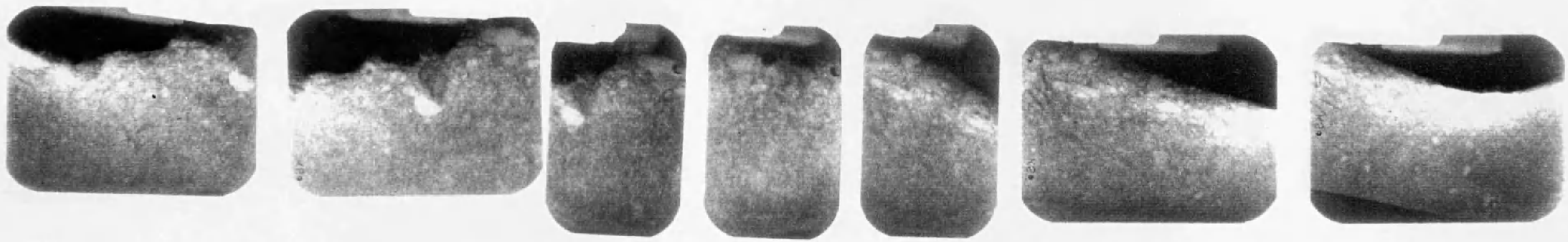


Figure 4.3(b) "Intra-oral" radiographs of simulated mandible 1:2 acrylic : bonemeal

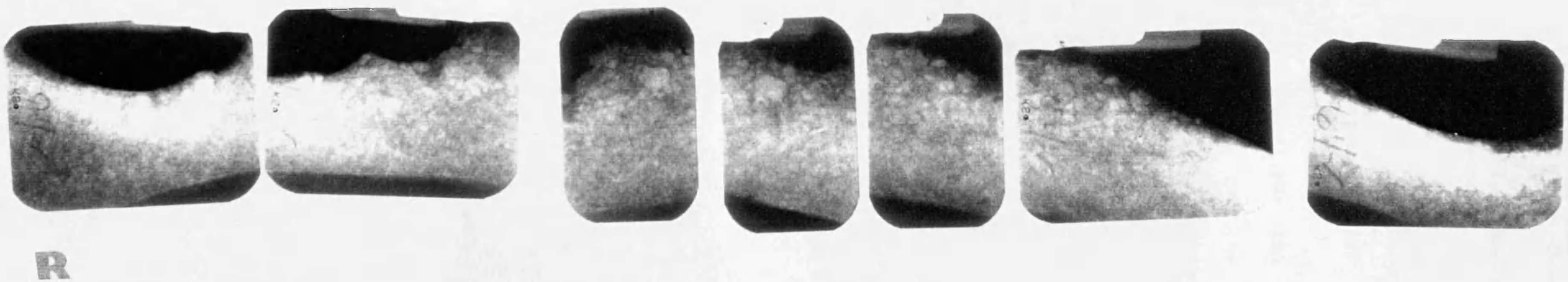


Figure 4.3(c) "Intra-oral" radiographs of simulated mandible 1:3 ratio acrylic : bonemeal

Table 4.1 Average grey level per pixel in test and control radiographs

	1	2	3	4	5	6	7	Mean	T value	Difference control vs test
Control	400	440	450	470	470	520	420	452		-
1:3 Bone/acrylic	310	290	310	380	280	300	350	317	6.81	$p<0.05$
1:2 Bone/acrylic	350	430	450	490	390	460	370	420	1.35	$p>0.05$
1:1 Bone/acrylic	440	640	500	750	735	580	770	630	3.48	$p<0.05$

to bonemeal when compared to the radiograph of a human mandible, whilst the other models (plastic:bonemeal ratio, 1:1 and 3:1) were significantly different ($p < 0.05$) from the real radiograph according to microdensitometric criteria. Thus it was decided that of the material combinations examined, a ratio of 2:1 plastic to bonemeal, gave the best radiodensitometric simulation of real bone. This was therefore the combination of materials used in the construction of the simulated bitewings.

4.3.2 Test 2. Construction of the models for the bitewing radiographs

Method

A volunteer had bitewing radiographs taken of their right and left dentition. The same volunteer then had upper and lower alginate impressions of the dentition taken. These impressions were poured in stone and the resulting models used to construct multiple wax casts. The posterior teeth were removed from these and replaced by "biteblocks" (Figure 4.4) into which extracted teeth were mounted. For this part of the study cariously involved extracted teeth were used as the extent of cavitation was unimportant. During the mounting of the teeth, care was taken to simulate correctly both occlusal level and contact points. The wax mounting was then carved to a level which would be similar to that seen in a healthy 30-year-old patient. The wax was then replaced by the plastic/bonemeal mixture using a traditional flasking/packing procedure. However, prior to packing, the roots of the teeth were coated with a thin coating of acrylic using a paint brush, in order that the periodontal space could be simulated. The acrylic-covered roots were then painted with a calcium hydroxide suspension in order that a radio-opaque simulation of the lamina dura was created. It is important to note that at all times the teeth were kept moist in an attempt to avoid dehydration and subsequent splitting or cracking. The appearance of the models prior to being radiographed may be seen in Figure 4.5.

Conclusion

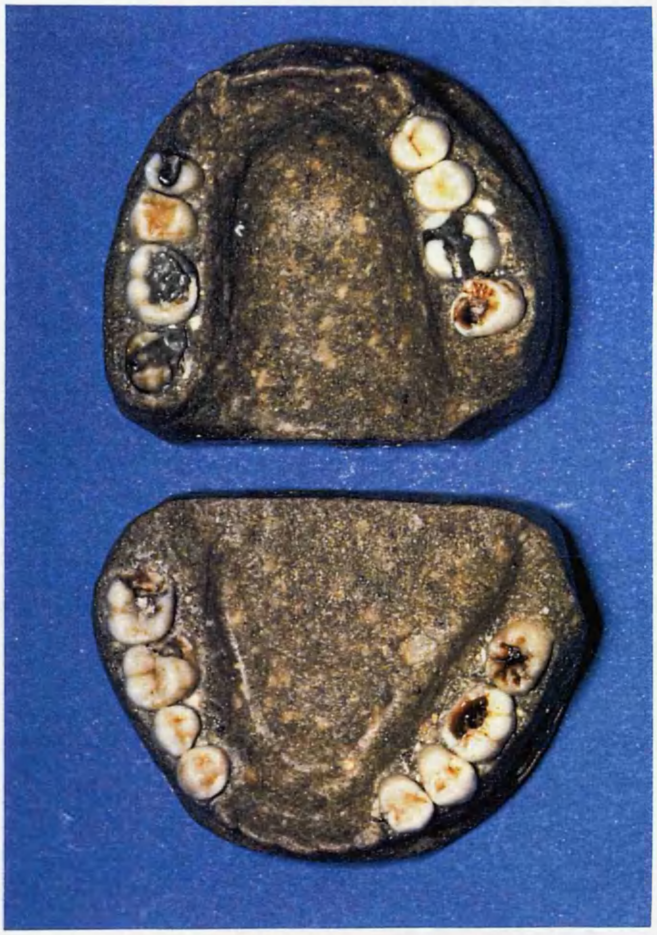
It was concluded that using the methodology described, a bitewing radiograph which adequately simulated those taken in patients could be created.

4.3.3 Test 3. Determination of a soft tissue simulator

Water has already been discussed with regards to its properties as a soft tissue simulator. However it was decided that for the purposes of this study, the practical problems of interposing water between the x-ray tube and the models could not be overcome.

Figure 4.4 Wax mounts for placement of extracted teeth

Figure 4.5 Models prior to radiography



Method

Using a standardised technique, the right and left cheek thicknesses of twenty members of the general public attending Glasgow Dental Hospital were measured.

The patients were positioned so that the mandible was parallel to the floor. Points of micrometer screw gauge callipers were then placed on the buccal mucosa opposite the contact point of the lower second premolar and the first molar. The patient was then asked to hold a piece of card similar to the "tab" of a bitewing radiograph between the teeth in order to simulate the tooth and soft tissue relationships pertaining during bitewing radiography. The callipers were then closed until the first signs of blanching of the cheek were seen. The first subject was measured several times in order to ensure standardisation of, and familiarity with, the technique. Only when reproducible results were obtained did the investigator proceed to the main part of the study, in which the cheek thicknesses of twenty individuals were measured.

Results

The mean cheek thickness of the twenty persons sampled was 14.7 mm (SD - 4.04). The mass attenuation coefficients of several tissues including perspex are shown in Table 4.2. This table shows that in the range of 60-80 kV photon energy, (i.e. the range usually produced by the x-ray tube during bitewing radiography) the mass attenuation coefficients of perspex are very similar to that of muscle tissue.

4.4 CONCLUSIONS

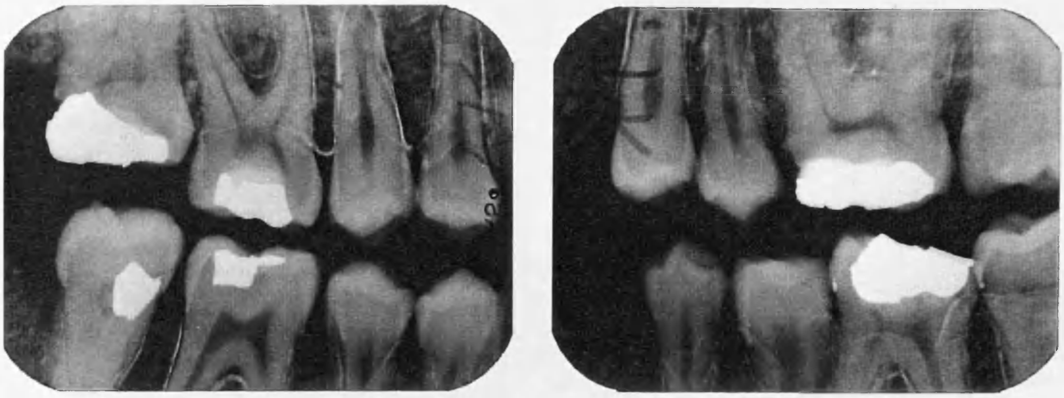
If perspex was to be used as a soft tissue simulator, it was concluded that a slab 15 mm thick would give a reasonable representation of the radiographic beam attenuation of the cheek soft tissues.

The fifteen pairs of bitewing radiographs used in the study are reproduced in Figure 4.6. It is clear they are not perfect replicas of normal radiographs. However, the purpose of the simulation was to produce a radiodensity surrounding the teeth to be examined, similar to that of an normal radiograph. This was in order that the radiographic "noise" during the rating scale experiment (Chapter 5) affected the treatment decision task, in the same way as the structures surrounding human teeth produce radiographic "noise" when a dentist examines a routine bitewing radiograph.

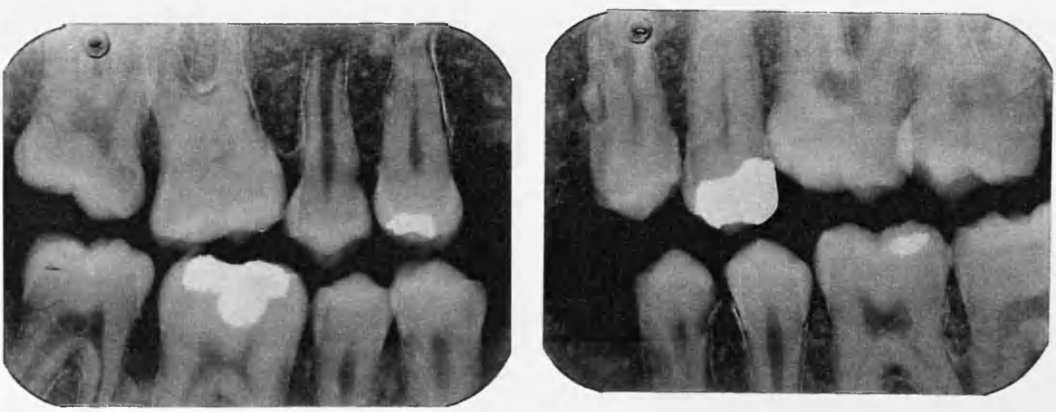
Table 4.2 Mass attenuating coefficients of various materials

Kilo voltage of x-ray	Polystyrene	Perspex	Water	Air	Muscle
15	0.0737	0.105	0.161	0.155	0.164
20	0.0420	0.0547	0.0773	0.0747	0.0790
30	0.0258	0.0295	0.0364	0.0343	0.0368
40	0.0217	0.0233	0.0264	0.0244	0.0265
50	0.0198	0.0206	0.0225	0.0206	0.0224
60	0.0187	0.0192	0.0205	0.0187	0.0204
80	0.0173	0.0175	0.0183	0.0166	0.0182
100	0.0162	0.0164	0.0171	0.0154	0.0169
150	0.0144	0.0145	0.0150	0.0135	0.0149

Figure 4.6 The simulated radiographs



Dentition 1

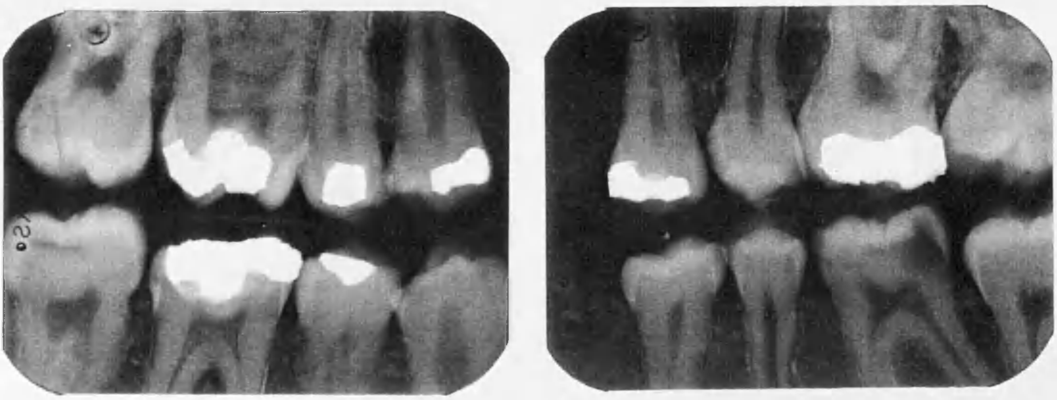


Dentition 2

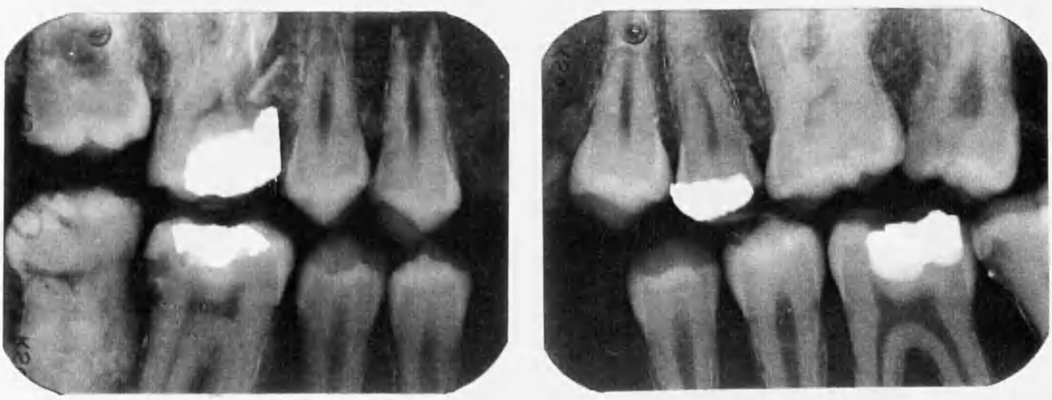
Dentition 3

Figure 4.6 (ctd) The simulated radiographs

Dentition 4

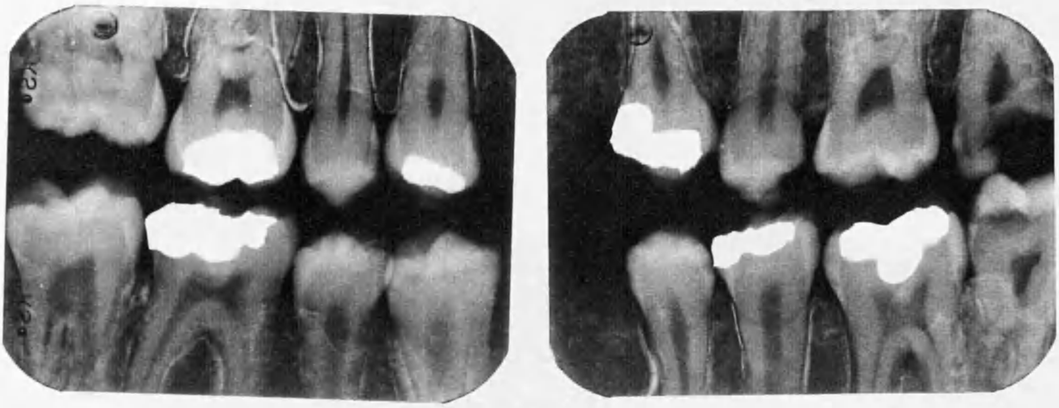


Dentition 5

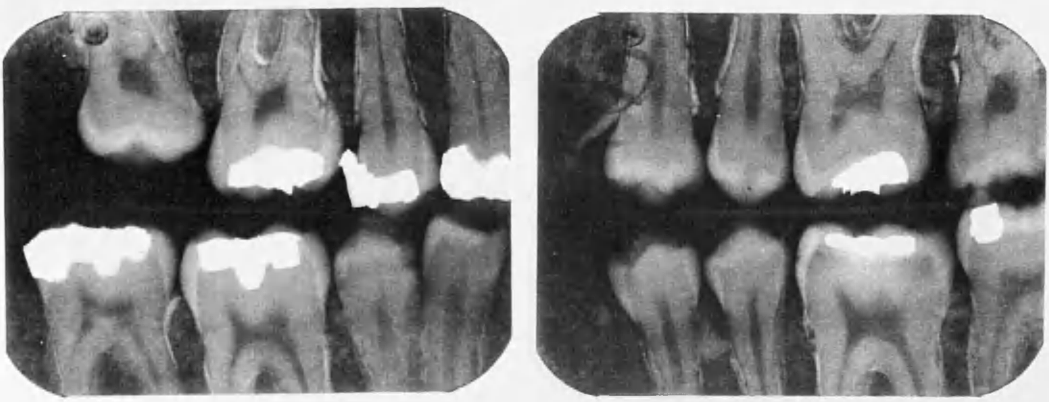


Dentition 6

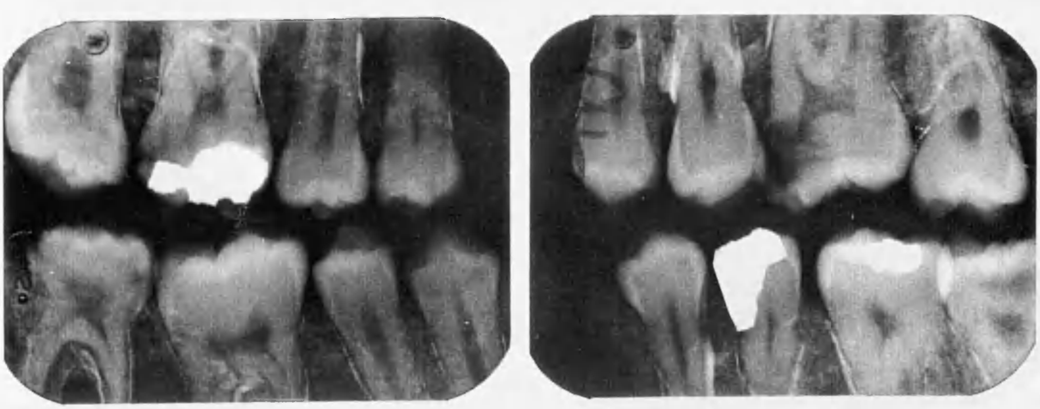
Figure 4.6 (ctd) The simulated radiographs



Dentition 7



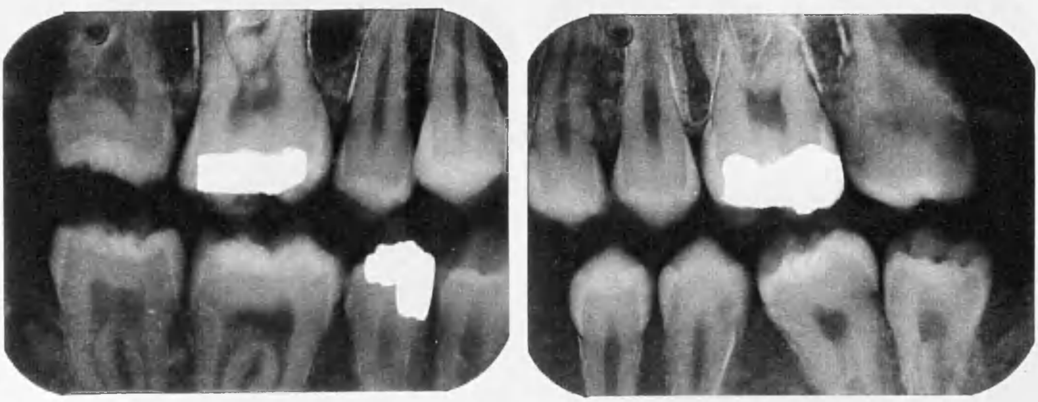
Dentition 8



Dentition 9

Figure 4.6 (ctd) The simulated radiographs

Dentition 10

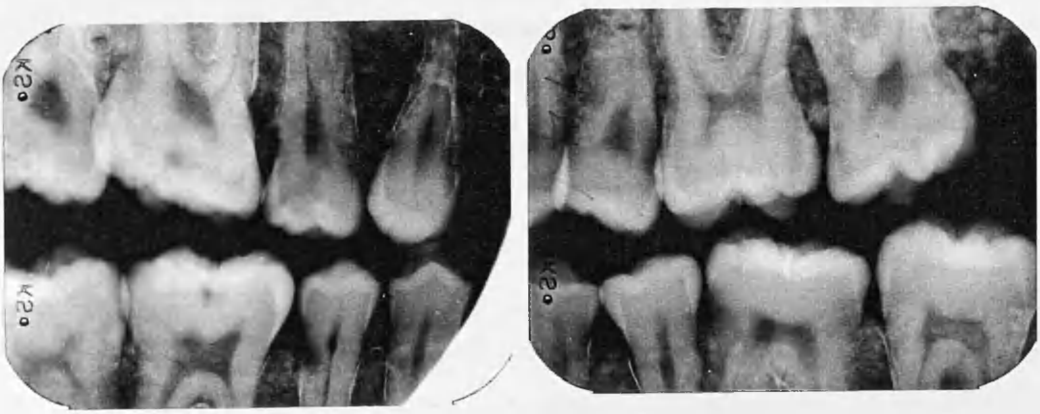


Dentition 11

Dentition 12

Figure 4.6 (ctd) The simulated radiographs

Dentition 13



Dentition 14

Dentition 15

CHAPTER 5. VALIDATION OF DENTISTS' TREATMENT DECISIONS

5.1 INTRODUCTION

Previous work[177,234] has been undertaken to validate dentists' treatment decision-making, but these studies concentrated on the reliability of the dentists' decisions in epidemiological studies, rather than on the diagnosis and subsequent treatment of caries undertaken by practising clinicians. This paucity of validation research is an important omission as the 'dental check-up' is accepted by dental health educationalists[235], dentists, and a considerable proportion of the general public[39,40], as an important method of preventing untoward sequelae of dental disease. Dental professionals must be certain that the treatment criteria which they adopt reliably predict teeth requiring restoration and, perhaps more importantly, enable them to identify correctly teeth which do not require intervention.

In the 1970's, any evidence of approximal enamel caries on a radiograph was considered to warrant restorative intervention[236]. More recently, authors have suggested that radiographic evidence of lesions half-way through dentine may not be sufficient to indicate treatment need[121]. Such scientific debate may engender confusion in the mind of the practitioner. Dentists who have not adapted to changing patterns of disease[27-33], may be over-investigating or over-treating patients, to the cost of both government and patients, in both monetary and health terms.

5.2 AIMS

The study described in this Chapter aims to describe the accuracy of dentists' treatment decisions in terms of their sensitivity, specificity and predictive power. A subsidiary aim was the construction of receiver operator characteristic curves describing the dentists' patterns of decision-making.

5.3 METHOD

5.3.1 Materials

The study required simulated bitewing radiographs which produced patterns of radiographic "noise" and "signals" similar to normal bitewing radiographs. In order to do this, two hundred and forty extracted molar and premolar teeth were mounted in models which simulated human dentitions (see Chapter 4). The teeth were selected to represent a range of carious lesions, from complete health, through early enamel and dentine lesions, to gross cavitation. The models were then radiographed in such a way as to ensure that the quality and quantity of radiation passing through the 'mock' dentition was in every way similar to that in normal bitewing radiography. This methodology had been developed and radiographically validated in tests which have been described in Chapter 4.

5.3.2 Data collection

Of the 134 full-time practising NHS general dental practitioners in Glasgow, 20 were randomly selected from Greater Glasgow Yellow Pages Telephone Directory. They were written to by the investigator and asked to join the study panel (Appendix I). Two refused and were substituted by the next dentist listed in the sampling frame, until 20 participants were identified. Each dentist was then visited by the investigator and was asked to examine the 30 simulated bitewing radiographs (15 pairs) in a method identical to that which he utilised during his day-to-day practise of dentistry.

The dentists were then asked to indicate whether or not they would choose to restore the approximal surfaces of each tooth, from the mesial aspect of the second permanent molar, to the distal aspect of the first premolar. Each dentist thus made a total of 360 treatment decisions as there were 12 approximal surfaces in each of the 30 bitewing radiographs. The dentist was asked to grade the certainty of each decision on a six point scale (see Appendix II) in order to allow ROC curves to be constructed.

5.3.3 Conducting a rating scale experiment

To gain information about an observer's criteria for decisions to restore teeth, one requires values for the probability of true positive and false positive decisions, at differing levels of a dentist's certainty that restoration was the treatment of choice.

Two methods are available to do this. Firstly it is possible to construct ROC curves by conducting a series of yes/no (restore/leave unrestored) decisions, whilst asking the observer (the dentist) to vary his bias, (criteria) from task to task. The second method, is to ask the observer to vary his selection criteria by expressing the certainty of his decision. This second technique is used here as it is the most efficient with regard to the number of decisions needed to obtain an ROC curve.

With regard to the rating method employed (see Appendix II) a six point scale was used. In ROC analysis there is one fewer available points for the ROC curve than the number of rating categories. There are no hard and fast rules for choosing a maximum number of categories[237], but the following points were borne in mind when designing this rating scale experiment.

Firstly, although a greater number of rating categories gives more points for the ROC curve, observers cannot usually deal consistently with many categories of "certainty"[238,239]. One therefore has to trade the diminishing returns caused by this problem, against the advantages of achieving increased numbers of points on the curve[240].

The greater the number of categories in a rating scale, the more likely it is that some will not be used by the observers. This can give rise to problems if the observer fails to use the categories at the end of the scale. In general, the greater the number of categories in a rating scale, the greater the number of decisions which must be made in order to ensure that the false positive, and false negative values are accurately recorded. McNicol[240] suggests that 4 to 10 categories are sufficient. He also indicates that combining adjacent categories during analysis is acceptable.

Secondly, it was important that the dentists involved in this rating scale experiment understood its meaning. Despite precautions, observers in rating scale tasks can forget what they have been told, or can be overwhelmed by the demands of the task and become confused. Fortunately, this study had advantages over other signal detection studies as the observers were dealing with a familiar, rather than an unfamiliar, treatment decision task.

5.3.4 Visual validation

After the models had been radiographed in preparation for the general dental practitioners' examination, the teeth were removed from them (Figure 5.1) and classified visually into four groups according to criteria first reported by Rugg-Gunn[50] (see Table 5.1). All teeth deemed to be sound were labelled according to the "dentition" from which they came, their tooth position, and their mesial and distal surfaces were identified. These teeth were then examined by a second independent examiner who was asked to classify them according to the same criteria.

When the two examiners disagreed, re-examination of the surface was followed by discussion until agreement was reached. Equivocal lesions were always scored as the severest score deemed to be appropriate by either examiner. The visual appearance of each approximal tooth surface was recorded. Subsequently, the teeth were mounted for sectioning (Figure 5.2). From each of these teeth 300 μ thick longitudinal sections were cut (Figure 5.3).

5.3.5 Microscopic validation

The sections were then mounted in water and viewed under the lowest power (x20) of a microscope.

Lesion depths were then classified according to the criteria shown in Table 5.2 and the lesion depth recorded (see Figures 5.4(i-iv) for examples).

Figure 5.1 Removal of teeth from model

Table 5.1 Visual criteria for caries

Score	Description
0	Sound
1	White spot with surface shine
2	White/brown spot with loss of surface shine
3	Cavity < 1/2 mm diameter
4	Cavity > 1/2 mm diameter
5	Amalgam - sound
6	Amalgam with defects
7	Amalgam with caries

Figure 5.2 Teeth mounted and labelled for sectioning

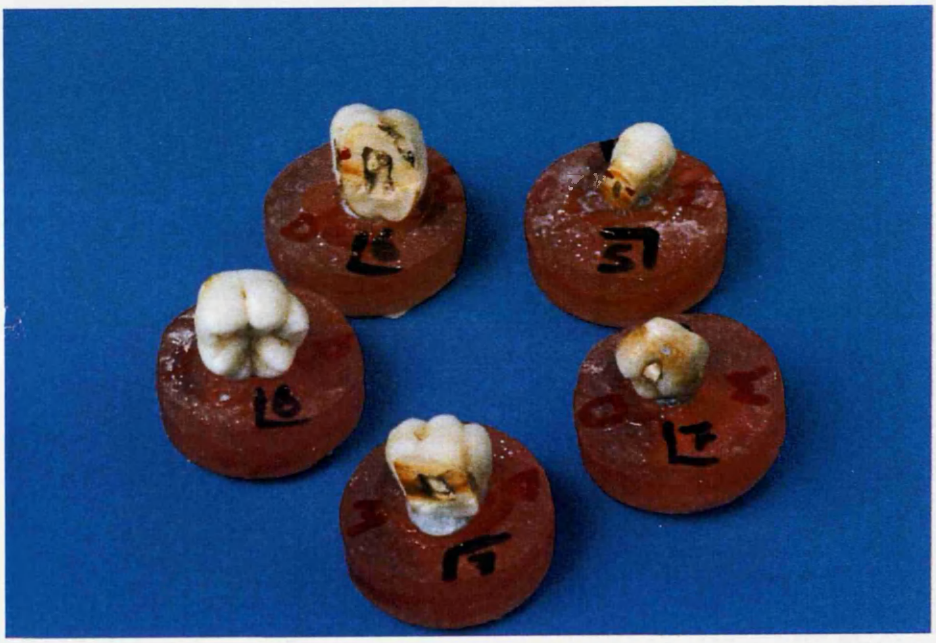


Figure 5.3 Sectioning of teeth

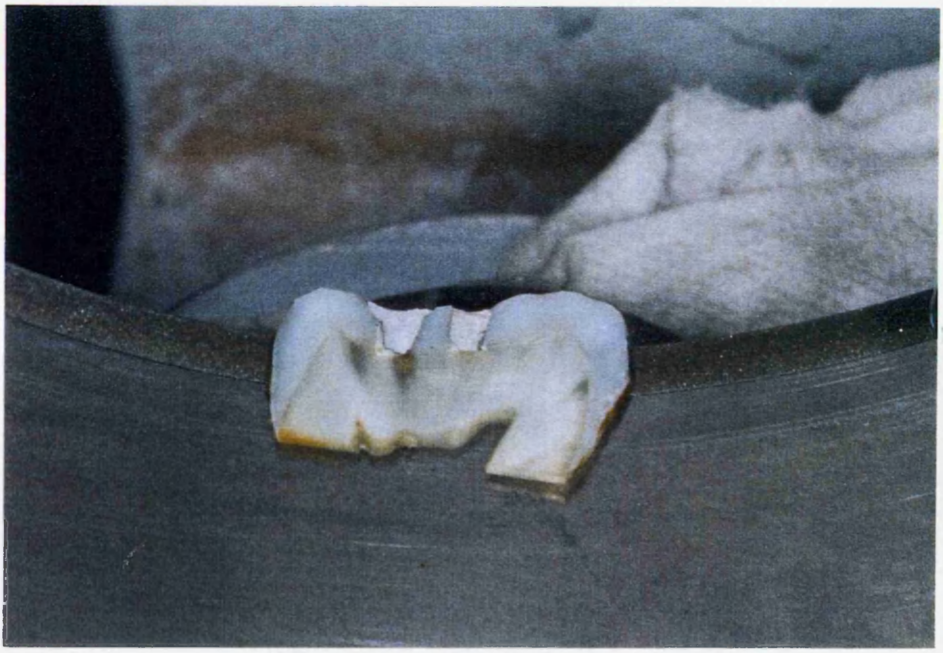
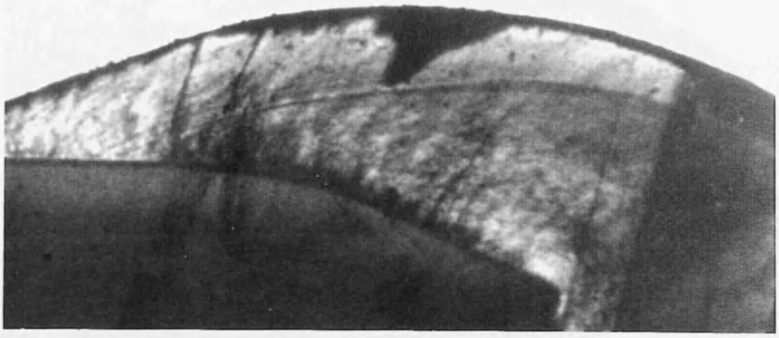


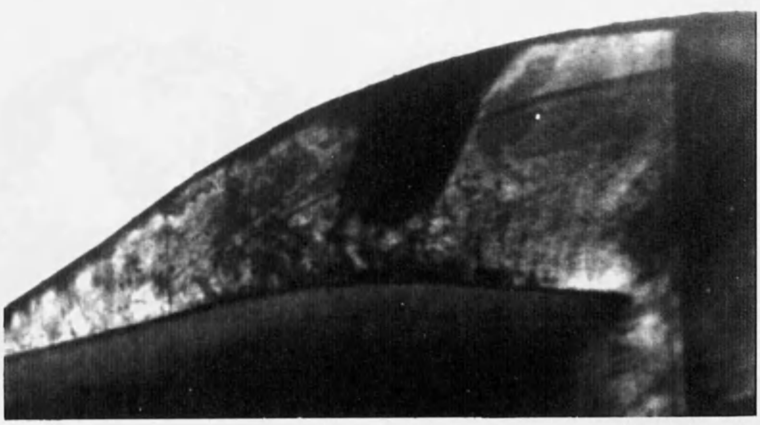
Table 5.2 Microscopic criteria for caries

Score	Description
0	Sound
1	Lesion limited to outer 1/2 of enamel
2	Lesion more than 1/2 way through enamel but not penetrating dentine
3	Lesion into dentine but not more than 1/2 way to pulp
4	Lesion into dentine and more than 1/2 way to pulp
5	Restored surface without faults
6	Restored surface with faults
7	Restored surface with caries (with or without faults)
9	Missing data

Figure 5.4

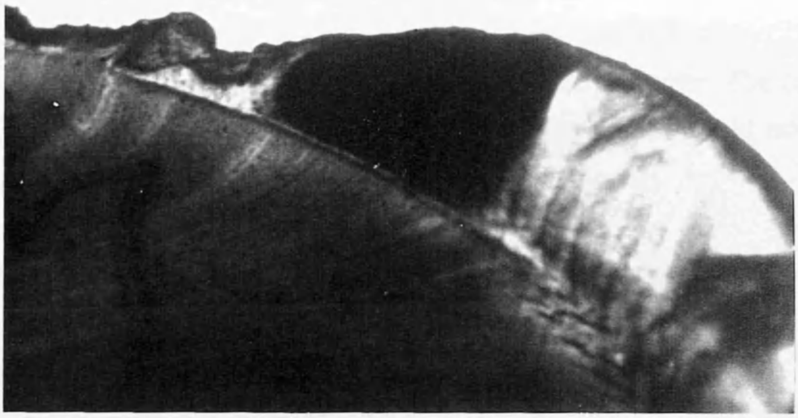


(i) carious lesion in outer enamel

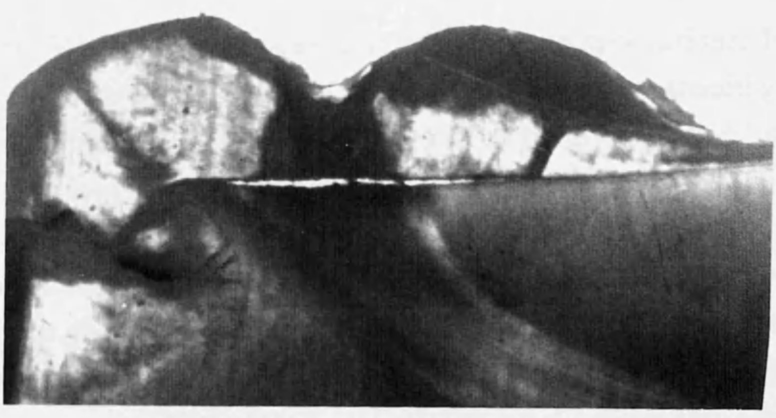


(ii) carious lesion extending into inner enamel

Figure 5.4 (ctd)



(iii) caries extending to, but not beyond ADJ



(iv) caries extending into dentine

5.3.6 Dentist intentions

One month after the rating scale experiment had been conducted, the dentists were asked by questionnaire (Appendix III) to state the depth of carious lesion which they intended to restore.

5.4 RESULTS

5.4.1 Disease prevalence

Table 5.3 shows, on the left hand side, the prevalence of visually detectable lesions amongst the 360 tooth surfaces examined. Of these 68% (244) of the surfaces were sound and 15% (56) of the tooth surfaces had visual evidence of caries (shiny, and non-shiny, white or brown spots), but showed no sign of cavitation. A further 12% (43) of the surfaces had cavities, most of these being less than 0.5 mm diameter. The remaining 7% (16) of the tooth surfaces had been restored. One tooth surface could not be visually inspected due to damage to the tooth on removal from the models.

The right hand side of Table 5.3 shows the prevalence of microscopic lesions in the approximal surfaces of the teeth used in the test radiographs. Of these surfaces 14% (52) had enamel lesions (outer and inner half), and 12% (43) had dentine lesions. Four percent of the surfaces (16) were not microscopically examined as eight teeth shattered during sectioning.

5.4.2 Comparison of visual and microscopic appearances

Data in Table 5.4 show the comparison between microscopically and visually detected lesions. The Table shows, that of the tooth surfaces which had appeared on visual inspection to be sound, seven (3%) were found microscopically to have outer enamel lesions with one (0.4%) having an inner enamel lesion and one (0.4%) a lesion extending into dentine. Of the visually sound tooth surfaces 1.6% were lost to microscopic follow-up as the tooth shattered during sectioning.

Regarding the fifteen tooth surfaces with glossy white spot lesions, 10 (66%) were microscopically restricted to enamel and one had a lesion extending into dentine. However four (26%) of these visually apparent early lesions showed no evidence of caries when examined microscopically.

Of the 41 surfaces which had lost surface shine but were not seen to be cavitated on visual inspection, four (10%) were microscopically sound and 23 (56%) of these surfaces exhibited enamel lesions, whilst 8 (20%) had lesions extending into dentine.

Table 5.3 The prevalence of visually detected and microscopically detected caries lesions in the test dentitions

	VISUAL N	VISUAL %		MICROSCOPIC N	MICROSCOPIC %
Sound	244	68	Sound	246	68
White spot with surface shine	15	4	Lesion in outer 1/2 enamel	23	6
White/brown spot lesion with loss of surface shine	41	11	Lesion in inner 1/2 enamel but not reaching dentine	29	8
Cavitation (<1/2 mm diameter)	36	10	Lesion into dentine (less than 1/2 way to pulp)	39	11
Cavitation (>1/2 mm diameter)	7	2	Lesion more than 1/2 way to pulp	4	1
Restoration, otherwise sound	9	2	Restored surface without faults	2	1
Restoration with defects	2	2	Restored surface with faults	0	0
Restoration with 2^{o} caries	5	1	Restored surface with caries (with or without faults)	1	0
Missing data	1	0	Missing data	16	4
TOTAL	360	100		360	100

Table 5.4 Comparison of visually and microscopically detected carious lesions

VISUAL CRITERIA	MICROSCOPIC CRITERIA: Sound	Lesion outer 1/2 enamel	Lesion inner 1/2 enamel	Lesion into dentine	Lesion more then 1/2 way in pulp	Restored surface w/o faults	Restored surface w caries	Missing	Total
Sound	231	7	1	1				4	244
White spot w surface shine	4	6	4	1					15
White/brown spot w loss of surface shine	6	9	14	8				4	41
Cavitation <1/2 mm	1		9	24	1			1	36
Cavitation >1/2 mm				3	3			1	7
Restoration otherwise sound	4		1			2	1	1	9
Restoration w defects								2	2
Restoration w 2^{o} caries		1		2				2	5
Missing								1	1
Total	246	23	29	39	4	2	1	16	360

The majority of surfaces with cavities less than 0.5 mm at their largest diameter had dentine lesions (25/36 = 68%). However, one of these cavities appeared microscopically to be sound, whilst 9 (25%) had caries extending into the inner half of the enamel.

Large cavities (>0.5 mm) were observed on visual inspection of seven tooth surfaces. and all of these which could be microscopically examined exhibited dentine lesions. The above results are detailed in Tables 5.3 and 5.4.

5.4.3 Dentists' treatment decisions

The treatment decisions (according to the rating scale in Appendix II) made by each dentist are shown in Table 5.5. It can be seen from this table that the dentists made decisions, and utilised the rating scale in an idiosyncratic manner. e.g. the range of number of decisions to definitely restore a tooth was 6 to 64 (mean = 29), whilst decisions to definitely leave a tooth unrestored were even more varied, with the numbers of definitely negative treatment decisions ranging from 0 to 335 (mean = 225).

Some dentists, (most notably dentist 11, and to a lesser extent dentists 1 and 18) showed a reluctance to use the extreme ends of the rating scale, in that they avoided making very definite treatment decisions. On the other hand, dentists 6 and 20 rarely equivocated in their decisions.

The depth of lesion at which each dentist intended to restore teeth is detailed in Table 5.5, by the use of superscripts a, b and c to indicate the groupings of dentists. Group (a) implies that the dentist felt that inner enamel lesions required restoration, group (b) implies intention to restore lesions extending through the full thickness of the enamel. and group (c) implies restoration, only of lesions which extended into dentine. The Table shows that the mean number of definitely positive decisions made by dentists who wished to fill deep enamel and dentine lesions group (a) was 21 whilst the number of definitely positive decisions made by dentists in group (b) was 36. Finally the mean number of definitely positive decisions made by those who intended to restore only dentine lesions (group c) was 25. It can be seen from this table that there was no trend towards a lower number of positive decisions among the dentists who wished to fill only deep lesions. There was a trend towards greater numbers of definite negative treatment decisions amongst practitioners who wished to fill only deep lesions but these differences were not statistically significant. (Gp A vs Gp B, $t = 0.15$; Gp A vs Gp C, $t = 0.23$; Gp B vs Gp C, $t = 0.16$) ($p>0.05$).

5.4.4 Treatment decisions in relation to visual appearance of tooth surface

Data in Table 5.6 show the distribution of the 7200 treatment decisions in relation to the visual appearance of tooth surfaces. Of the total number of decisions to definitely restore

Table 5.5 Ratings given by each of 20 dentists to the 360 approximal surfaces viewed in the test radiographs

RATING SCALE SCORES GIVEN BY EACH DENTIST
(Bold type indicates extremes of range)

DENTIST	1	2	3	4	5	6
1[b]	22	**55**	**86**	90	103	4
2[c]	15	2	2	5	1	**335**
3[b]	32	12	26	8	10	272
4[b]	47	13	41	35	79	145
5[c]	24	35	45	25	65	166
6[a]	34	5	3	2	1	315
7[b]	11	7	10	9	11	312
8[b]	36	13	39	17	5	250
9[a]	22	12	18	**0**	1	307
10[c]	35	11	18	5	11	280
11[a]	6	25	31	**204**	94	**0**
12[c]	34	**4**	30	2	**0**	290
13[c]	7	6	20	2	14	311
14[c]	34	23	26	30	58	189
15[c]	31	27	35	23	36	208
16[c]	30	5	13	1	2	309
17[b]	**64**	12	44	6	9	225
18[c]	14	18	44	76	**207**	1
19[b]	26	15	12	0	16	291
20[b]	50	11	4	1	1	293
Mean	29	16	27	27	36	225

Score 1 = Definitely restore tooth
Score 2 = Probably restore tooth
Score 3 = Possibly restore tooth
Score 4 = Possibly leave tooth unrestored
Score 5 = Probably leave tooth unrestored
Score 6 = Definitely leave tooth unrestored

Mean no. of definitely positive treatment decisions by dentists in group a =
mean definitely positive = 21
mean definitely negative = 207

Mean no. of definitely positive and negative treatment decisions by dentsts in group b =
mean definitely positive = 36
mean definitely negative = 224

Mean no. of definitely positive and negative treatment decisions by dentists in group c =
mean definitely positive = 25
mean definitely negative = 232

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.6 Comparison of visually detected caries in the dentists' treatment decisions

DENTISTS' TREATMENT DECISION	VISUAL APPEARANCE OF SURFACE: Sound	White spot with surface spot shine	White/brown with loss of surface shine	Cavitation <1/2 mm	Cavitation >1/2mm	Restored (surface sound)	Restoration with defects	Restoration with 2° caries	Total
1 Definitely restore	190	32	68	152	63	15	16	38	574
2 Probably restore	132	13	50	77	13	10	3	13	311
3 Consider restore	300	38	79	83	16	18	3	10	547
4 Consider leave unrestored	363	24	61	55	9	15	4	10	541
5 Probably leave unrestored	516	24	79	67	6	20	2	10	724
6 Definitely leave unrestored	3379	169	483	286	33	102	12	39	4503
Total (%) of all tooth surfaces examined visually	4880 (68%)	300 (4%)	820 (11%)	720 (10%)	140 (2%)	180 (3%)	40 (<1%)	120 (2%)	7200

a tooth surface (574), one third (190) were made about sites which on visual inspection were completely sound. However, of all the decisions to definitely leave a tooth unrestored (4503), 75% were made concerning tooth surfaces which were visually sound. The trend is clear, that greater numbers of decisions to restore are made with increasing visual evidence of caries.

5.4.5 Treatment decisions in relation to microscopic appearance of tooth surface

The distribution of microscopic lesions within each diagnostic rating category is shown in Table 5.7. Of 514 decisions to definitely restore a tooth surface, 34% (194) were made about surfaces which were microscopically sound. However, 48% (274) of these definitely positive decisions were made about surfaces which had microscopic evidence of lesions into dentine. Of the 4503 decisions to definitely leave a tooth unrestored, 75% (3389) were made of sites which were microscopically sound, although (6%) 279 of these definitely negative decisions concerned surfaces shown to have lesions into dentine.

For lesions into dentine, but less than half way to the pulp, the distribution of treatment decisions is bimodal, whilst for all other lesion depths the trend in the distribution of treatment decisions is as would be expected. i.e. the deeper the lesion the more likely a dentist was to wish to restore the surface.

5.4.6 "Correct" treatment decisions

Tables 5.8(a), (b) and (c) indicate the numbers and proportions of 'correct' treatment decisions made by each practitioner according to three validating criteria. The Tables show the sum of the number of 'correct' positive and negative treatment decisions. Decisions are considered to be "positive" at three levels, i.e. using rating 1, ratings 1-3 and ratings 1-5 as "positive decisions". In Table 5.8(a) a decision to restore a tooth surface is considered to be 'correct' if the caries was seen microscopically to extend into dentine, and a decision to leave unrestored is considered to be 'correct' if the tooth had no lesion, or if a lesion was confined to enamel. In the following Tables, 5.8(b) and (c) the number of 'correct' decisions are shown when it is considered that a tooth surface "should" be restored when the lesion was in inner half of enamel or in dentine (Table 5.8(b)). In Table 5.8(c) any lesion, even if only in the outer enamel is considered to be in need of restoration. Therefore, in that Table a treatment decision was considered 'correct' if there was any microscopic evidence of a carious lesion.

Using the mean values, the highest number of correct treatment decisions were made (87%) when lesions into dentine were considered as the correct point of intervention, and when the strictest diagnostic criteria were held, i.e. when only a 'definite' decision to

Table 5.7 Comparison of microscopically detected caries to the dentists' treatment decisions

DENTISTS TREATMENT DECISION	MICROSCOPIC CRITERIA Sound	Lesion outer 1/2 enamel	Lesion inner 1/2 enamel	Lesion into dentine less than 1/2 way to pulp	Lesion more then 1/2 way to pulp	Restored surface without faults	Restored surface with caries	Missing data	Total
Definitely restore	194	31	27	244	30	5		43	574
Probably restore	137	21	41	83	10	5		14	311
Possible restore	311	36	71	92	12	7	2	16	547
Possibly leave unrestored	361	34	55	42	6	5	4	34	541
Probably leave unrestored	528	41	57	58	4	2	1	33	724
Definitely leave unrestored	3389	297	329	261	18	16	13	180	4503
Total no. and % of all tooth surfaces examined microscopically (6880)	4920 72%	460 7%	580 8%	780 11%	80 1%	40 <1%	20 <1%	320	7200

Table 5.8(a) Number and proportion of 'correct' treatment decisions using caries into dentine as the validating criterion

Dentist	Rating 1 as positive treatment decisions N	%	Rating 1-3 as positive treatment decisions N	%	Rating 1-5 as positive treatment decisions N	%
1[b]	288	84	210	61	48	14
2[c]	304	88	301	88	301	88
3[b]	302	88	287	83	276	80
4[b]	298	87	263	77	177	52
5[c]	306	89	256	74	194	56
6[a]	302	88	301	88	300	87
7[b]	303	88	300	87	287	83
8[b]	296	86	266	77	250	73
9[a]	297	86	283	82	282	82
10[c]	305	89	290	84	282	82
11[a]	298	87	288	84	44	13
12[c]	304	88	287	83	287	83
13[c]	296	86	299	87	291	85
14[c]	304	88	273	79	202	59
15[c]	296	86	280	81	225	65
16[c]	304	88	302	88	301	88
17[b]	287	83	254	74	243	71
18[c]	286	83	265	77	45	13
19[b]	307	89	297	87	291	85
20[b]	303	88	295	86	293	85
Mean	299	87%	280	81%	231	67%

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	299	291	209
	S.D.	2.65	9.29	142.9
Group b	Mean	298	272	233
	S.D.	7.29	30.3	84.1
Group c	Mean	301	284	236
	S.D.	6.62	16.45	83.5
	a vs b	t = 0.33 NS	t = 1.60 NS	t = 0.28 NS
	a vs b	t = 0.58 NS	t = 0.91 NS	t = 0.32 NS
	b vs c	t = 0.75 NS	t = 1.01 NS	t = 0.08 NS

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.8(b) Number of proportion of 'correct' treatment decisions using caries into inner enamel or dentine as the validating criterion

Dentist	Rating 1 as positive treatment decisions N	%	Rating 1-3 as positive treatment decisions N	%	Rating 1-5 as positive treatment decisions N	%
1[b]	276	80	111	33	77	22
2[c]	275	80	274	80	274	80
3[b]	277	81	276	80	271	79
4[b]	277	81	254	74	184	54
5[c]	277	81	243	71	203	59
6[a]	277	81	278	81	277	81
7[b]	274	80	273	79	266	77
8[b]	271	79	259	75	243	71
9[a]	268	78	266	77	265	77
10[c]	278	81	277	81	273	79
11[a]	269	78	271	79	73	21
12[c]	279	81	279	81	270	79
13[c]	273	79	274	80	272	79
14[c]	279	81	258	75	203	59
15[c]	269	78	261	76	230	67
*16[c]	279	81	283	82	282	82
17[b]	266	77	247	72	248	72
18[c]	269	78	252	73	74	22
19[b]	278	81	280	81	276	80
20[b]	265	77	220	64	110	32
Mean	274	76%	257	71%	218	61%

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	271	272	205
	S.D.	4.93	6.03	114.5
Group b	Mean	273	240	209
	S.D.	5.13	55.6	77.6
Group c	Mean	275	267	231
	S.D.	4.12	13.8	66.7
	a vs b	t = 0.49 NS	t = 1.59 NS	t = 0.06 NS
	a vs b	t = 1.26 NS	t = 0.85 NS	t = 0.38 NS
	b vs c	t = 1.03 NS	t = 1.33 NS	t = 0.62 NS

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.8(c) Number of proportion of 'correct' treatment decisons using any carious lesion as the validating criterion

Dentist	Rating 1 as positive treatment decisions N	%	Rating 1-3 as positive treatment decisions N	%	Rating 1-5 as positive treatment decisions N	%
1[b]	257	75	112	33	100	29
2[c]	252	73	253	74	255	74
3[b]	258	75	259	75	256	74
4[b]	256	74	245	71	193	56
5[c]	254	74	234	68	200	58
6[a]	258	75	261	76	260	76
7[b]	251	73	252	73	247	72
8[b]	254	74	246	72	234	68
9[a]	249	72	249	72	250	73
10[c]	259	75	262	76	260	76
11[a]	246	72	254	74	17	5
12[c]	260	76	253	74	253	74
13[c]	252	73	257	75	257	75
14[c]	256	74	243	71	196	60
15[c]	250	73	228	66	223	65
16[c]	260	76	264	77	263	77
17[b]	247	72	240	70	249	72
18[c]	248	72	243	71	97	28
19[b]	259	75	263	77	259	75
20[b]	297	86	283	82	281	82
Mean	256	71%	245	68%	218	60%

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	251.0	254.7	176
	S.D.	6.24	6.03	138
Group b	Mean	259.8	237.5	227
	S.D.	15.51	52.5	57.3
Group c	Mean	254.6	248.6	222.7
	S.D.	4.45	12.38	53.8
	a vs b	t = 1.35 NS	t = 0.91 NS	t = 0.63 NS
	a vs b	t = 0.91 NS	t = 1.13 NS	t = 0.58 NS
	b vs c	t = 0.94 NS	t = 0.58 NS	t = 0.17 NS

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

restore was considered as a positive treatment choice, and all other decisions were deemed to be negative.

The decreasing numbers of 'correct' treatment decisions in tables 5.6(b) and 5.6(c) are due to the decreasing numbers of negative decisions made about shallow lesions, which, according to the validating criteria used in the construction of these tables, "should" have been restored.

Of the 360 decisions made by each dentist, taking both positive and negative decisions into account, over 80% of the decisions were correct when dentine caries was used as the 'gold standard'. There were no significant differences in the mean numbers of 'correct' decisions made by the groups of dentists who intended to restore at these different lesion depths ($p > 0.05$).

5.4.7 Sensitivity and specificity of dentists' treatment decisions

Tables 5.9 (a, b and c) indicate the sensitivity and specificity of the dentists' decisions using three gold standards (lesions in dentine, inner enamel and dentine, and any lesion, respectively). Sensitivity, in this context is the proportion of all truly diseased tooth surfaces which are identified as such and deemed to require restoration. It can be seen that in all three tables the sensitivity of the treatment decisions increased as the 'cut-off' level for a positive decision is lowered to include all decisions other than a definitely negative decision. The mean sensitivity declined when the validating criteria, or 'gold standard' was made more relaxed i.e. when the lesion does not have to extend so deeply before restoration was considered to be necessary. (See Tables 5.9(b) and (c)).

Specificity is the proportion of all truly sound tooth surfaces which are identified as such and deemed not to require restoration. The specificity of the dentists' treatment decisions tended to decrease as the 'cut-off' point for what was considered to be a positive treatment decision was lowered.

The Tables indicate wide variation in both the sensitivity and specificity of individual dentist's decisions at each treatment threshold. For example, when only the strictest threshold (rating 1) is considered as a positive decision (Table 5.9a), dentists' sensitivities ranged from 0.09 (dentist 13) to 0.57 (dentist 20). The mean sensitivity being 0.26. The specificities ranged from 0.88 (dentist 17) to 0.99 (dentists 2,7,11 and 13) with a mean specificity of 0.96.

There were no differences in the mean sensitivities and specificities achieved by dentists who stated they intended to restore teeth at different depths of lesions (denoted by superscripts a, b and c) ($p > 0.05$).

Table 5.9(a) Sensitivity (Sens) and specificity (Spec) of treatment decisions using caries into dentine as the validating criterion

Dentist	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisions	
	Sens	Spec	Sens	Spec	Sens	Spec
1[b]	0.27	0.90	0.75	0.59	1.00	∞
2[c]	0.18	0.99	0.18	0.95	0.25	0.94
3[b]	0.38	0.95	0.61	0.86	0.68	0.82
4[b]	0.47	0.92	0.68	0.78	0.91	0.45
5[c]	0.32	0.97	0.64	0.76	0.86	0.52
6[a]	0.36	0.95	0.43	0.94	0.45	0.93
7[b]	0.14	0.99	0.29	0.96	0.36	0.90
8[b]	0.32	0.94	0.55	0.81	0.39	0.74
9[a]	0.21	0.96	0.39	0.89	0.39	0.88
10[c]	0.43	0.95	0.59	0.88	0.66	0.84
11[a]	0.05	0.99	0.57	0.88	1.00	∞
12[c]	0.41	0.95	0.59	0.87	0.61	0.87
13[c]	0.09	0.99	0.34	0.95	0.43	0.91
14[c]	0.41	0.95	0.59	0.82	0.73	0.57
15[c]	0.27	0.95	0.66	0.80	0.77	0.64
16[c]	0.34	0.96	0.46	0.93	0.55	0.92
17[b]	0.52	0.88	0.75	0.74	0.80	0.69
18[c]	0.11	0.94	0.43	0.82	1.00	∞
19[b]	0.36	0.97	0.48	0.91	0.66	0.87
20[b]	0.57	0.93	0.64	0.89	0.64	0.88
Mean	0.26	0.96	0.53	0.85	0.66	0.74

		Rating 1 as positive		Rating 1-3 as positive		Rating 1-5 as positive	
		Sens	Spec	Sens	Spec	Sens	Spec
Group a	Mean	0.207	0.967	0.463	0.903	0.613	0.905
	S.D.	0.155	0.02	0.095	0.032	0.336	0.035
Group b	Mean	0.379	0.935	0.594	0.817	0.680	0.764
	S.D.	0.140	0.04	0.154	0.117	0.2260	0.159
Group c	Mean	0.284	0.961	0.498	0.864	0.651	0.776
	S.D.	0.131	0.018	0.160	0.069	0.226	0.171
a vs b		t=1.68	t=1.81	t=1.69	t=1.9	t=0.32	t=2.17
a vs c		t=0.78	t=0.41	t=0.45	t=1.31	t=0.18	t=1.97
c vs b		t=1.43	t=1.86	t=1.26	t=0.99	t=0.26	t=0.14

All differences non-significant

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.9(b) Sensitivity (Sens) and specificity (Spec) of treatment decisions using caries into inner enamel or dentine as the validating criterion

	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisions	
Dentist	Sens	Spec	Sens	Spec	Sens	Spec
1[b]	0.17	0.66	0.66	0.60	1.00	0.02
2[c]	0.11	0.12	0.12	0.98	0.03	0.97
3[b]	0.26	0.49	0.49	0.89	0.58	0.85
4[b]	0.32	0.55	0.55	0.79	0.79	0.47
5[c]	0.19	0.49	0.49	0.76	0.78	0.54
6[a]	0.25	0.30	0.30	0.95	0.31	0.94
7[b]	0.08	0.19	0.19	0.96	0.27	0.91
8[b]	0.22	0.48	0.48	0.83	0.52	0.76
9[a]	0.12	0.32	0.32	0.90	0.32	0.89
10[c]	0.27	0.47	0.47	0.90	0.53	0.36
11[a]	0.03	0.43	0.43	0.89	1.00	∞
12[c]	0.27	0.44	0.44	0.88	0.45	0.88
13[c]	0.06	0.23	0.23	0.95	0.33	0.92
14[c]	0.27	0.45	0.45	0.83	0.64	0.58
15[c]	0.18	0.53	0.53	0.82	0.70	0.66
16[c]	0.23	0.38	0.38	0.94	0.40	0.93
17[b]	0.37	0.60	0.60	0.75	0.71	0.72
18[c]	0.08	0.37	0.37	0.83	1.00	∞
19[b]	0.08	0.41	0.41	0.92	0.49	0.89
20[b]	0.40	0.47	0.47	0.90	0.49	0.90
Mean	0.20	0.42	0.42	0.86	0.57	0.73

		Rating 1 as positive		Rating 1-3 as positive		Rating 1-5 as positive	
		Sens	Spec	Sens	Spec	Sens	Spec
Group a	Mean	0.133	0.350	0.350	01.913	0.543	0.915
	S.D.	0.11	0.070	0.070	0.032	0.396	0.035
Group b	Mean	0.237	0.481	0.481	0.830	0.606	0.690
	S.D.	0.123	0.142	0.142	0.116	0.223	0.307
Group c	Mean	0.184	0.387	0.387	0.877	0.540	0.730
	S.D.	0.084	0.133	0.133	0.0723	0.282	0.226
a vs b		t=1.35	t=2.04	t=2.04	t=1.85	t=0.26	t=2.02
a vs c		t=0.73	t=0.61	t=0.73	t=1.21	t=0.01	t=0.30
c vs b		t=1.03	t=1.41	t=1.41	t=0.98	t=0.54	t=2.21

All differences non-significant

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.9(c) Sensitivity (Sens) and specificity (Spec) of treatment decisions using any carious lesion as the validating criterion

	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisons	
Dentist	Sens	Spec	Sens	Spec	Sens	Spec
1[b]	0.16	0.98	0.63	0.81	1.00	0.02
2[c]	0.08	0.98	0.10	0.74	0.15	0.97
3[b]	0.22	0.96	0.41	0.79	0.48	0.88
4[b]	0.27	0.93	0.49	0.80	0.77	0.48
5[c]	0.15	0.97	0.45	0.78	0.70	0.54
6[a]	0.21	0.96	0.26	0.77	0.27	0.94
7[b]	0.06	0.99	0.16	0.75	0.23	0.91
8[b]	0.20	0.95	0.42	0.79	0.47	0.76
9[a]	0.12	0.96	0.27	0.76	0.28	0.87
10[c]	0.23	0.96	0.40	0.79	0.46	0.87
11[a]	0.02	0.98	0.35	0.78	0.18	∞
12[c]	0.23	0.96	0.37	0.78	0.38	0.88
13[c]	0.05	1.00	0.21	0.76	0.29	0.92
14[c]	0.21	0.95	0.39	0.78	0.57	0.57
15[c]	0.15	0.95	0.46	0.79	0.62	0.66
16[c]	0.20	0.97	0.31	0.78	0.32	0.94
17[b]	0.30	0.88	0.54	0.81	0.70	0.73
18[c]	0.07	0.97	0.35	0.77	1.00	-
19[b]	0.19	0.97	0.34	0.79	0.41	0.89
20[b]	0.34	0.94	0.40	0.79	0.42	0.91
Mean	0.17	0.96	0.37	0.78	0.49	0.76

		Rating 1 as positive		Rating 1-3 as positive		Rating 1-5 as positive	
		Sens	Spec	Sens	Spec	Sens	Spec
Group a	Mean	0.117	0.967	0.2933	0.770	0.243	0.937
	S.D.	0.050	0.0115	0.4493	0.0100	0.0551	0.065
Group b	Mean	0.2175	0.9500	0.4240	0.7912	0.360	0.697
	S.D.	0.088	0.0346	0.140	0.0189	0.246	0.310
Group c	Mean	0.1522	0.968	0.338	0.7744	0.499	0.794
	S.D.	0.0708	0.0156	0.117	0.0159	0.256	0.175
a vs b		t=1.62	t=1.21	t=2.28	t=2.41	t=1.42	t=2.07
a vs c		t=1.68	t=0.90	t=0.92	t=0.57	t=0.12	t=1.98
c vs b		t=1.78	t=1.36	t=1.36	t=1.97	t=0.50	t=0.77

All differences non-significant

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

5.4.8 Predictive power of positive and negative treatment decisions

The predictive power of a positive treatment decision is the proportion of all tooth surfaces designated for restoration which are truly carious.

Data in Table 5.10(a) indicate the predictive power of positive treatment decisions when "truly carious" is defined as a lesion extending into dentine. In this table, it can be seen that, as the threshold for a "positive" decision is lowered (i.e. ratings 2-5 included as positive decisions), the predictive power declines. It can also be seen that as the predictive power of a positive treatment decision decreases, the predictive power of a negative decision increases.

The range of predictive power of positive decisions was 0.33 to 0.66. For example of the tooth surfaces designated by dentist 11 as definitely requiring treatment, only one third had a lesion extending into dentine. Likewise, of the tooth surfaces designated by dentist 13 as definitely requiring restoration, two-thirds had lesions extending into dentine.

Overall, approximately half of the surfaces (53%) designated as definitely requiring restoration actually had lesions into dentine.

If all the tooth surfaces considered to possibly, probably or definitely require restoration are included, a mean of 39% actually had dentinal caries.

If all decisions other than a decision to definitely not restore a surface are considered to be positive decisions, only 32% of the positive decisions were made about tooth surfaces which actually had a dentinal lesion.

With regard to negative decisions, (decisions to definitely leave a tooth unrestored), a mean of 95% were made regarding tooth surfaces which did not have caries into dentine. The range was 91% to 100%.

The predictive powers of positive and negative treatment decisions when the 'gold standard' was inner enamel **AND** dentine lesions, are shown in Table 5.10(b). The data show that a mean of 58% of lesions designated as definitely requiring restoration had lesions in the inner half of enamel or dentine (Table 5.10(b)). As in table 5.10(a), as the treatment threshold is lowered. i.e. if dentists restore all teeth which they consider possibly, probably, and definitely require restoration, 49% of the surfaces restored would have lesions which extended beyond the outer enamel.

Table 5.10(a) Predictive power (PP) of positive and negative treatment decisions using caries into dentine as the validating criterion

	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisions	
	PP +ve	PP -ve	PP +ve	PP -ve	PP +ve	PP -ve
1[b]	0.57	0.89	0.21	0.94	0.13	1.0
2[c]	0.66	0.89	0.53	0.91	0.52	0.90
3[b]	0.53	0.91	0.40	0.94	0.36	0.95
4[b]	0.47	0.92	0.31	0.94	0.20	0.97
5[c]	0.63	0.91	0.28	0.93	0.21	0.96
6[a]	0.53	0.91	0.51	0.92	0.50	0.92
7[b]	0.67	0.89	0.50	0.91	0.35	0.91
8[b]	0.44	0.90	0.29	0.92	0.26	0.93
9[a]	0.43	0.89	0.33	0.91	0.33	0.91
10[c]	0.58	0.92	0.42	0.94	0.38	0.94
11[a]	0.33	0.88	0.40	0.94	0.13	∞
12[c]	0.56	0.92	0.40	0.94	0.40	0.94
13[c]	0.67	0.88	0.48	0.91	0.40	0.92
14[c]	0.56	0.92	0.33	0.93	0.21	0.93
15[c]	0.43	0.90	0.33	0.94	0.24	0.95
16[c]	0.58	0.91	0.52	0.93	0.51	0.93
17[b]	0.39	0.93	0.30	0.95	0.28	0.96
18[c]	0.36	0.88	0.26	0.91	0.02	1.00
19[b]	0.64	0.91	0.47	0.93	0.43	0.95
20[b]	0.53	0.94	0.45	0.94	0.44	0.94
Mean	0.53	0.91	0.39	0.93	0.32	0.95

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	0.43	0.413	0.320
	S.D.	0.10	0.0907	0.185
Group b	Mean	0.530	0.366	0.306
	S.D.	0.961	0.103	0.109
Group c	Mean	0.559	0.394	0.321
	S.D.	0.103	0.101	0.163
	a vs b	t = 1.49 NS	t = 0.74 NS	t = 0.12 NS
	a vs b	t = 1.92 NS	t = 0.30 NS	t = 0.01 NS
	b vs c	t = 0.60 NS	t = 0.57 NS	t = 0.22 NS

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.10(b) Predictive power (PP) of positive and negative using caries into inner enamel or dentine as the validating criterion

Dentist	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisions	
	PP +ve test	PP -ve test	PP +ve test	PP -ve test	PP +ve test	PP -ve test
1[b]	0.62	0.81	0.31	0.87	0.21	1.00
2[c]	0.66	0.80	0.60	0.81	0.57	0.81
3[b]	0.86	0.83	0.54	0.87	0.50	0.88
4[b]	0.54	0.83	0.41	0.87	0.29	0.89
5[c]	0.64	0.82	0.36	0.85	0.31	0.89
6[a]	0.60	0.82	0.59	0.83	0.58	0.84
7[b]	0.67	0.80	0.54	0.82	0.44	0.82
8[b]	0.50	0.82	0.47	0.86	0.37	0.85
9[a]	0.43	0.80	0.45	0.83	0.44	0.83
10[c]	0.61	0.83	0.55	0.86	0.51	0.87
11[a]	0.33	0.79	0.50	0.85	0.21	1.00
12[c]	0.63	0.83	0.49	0.85	0.49	0.86
13[c]	0.67	0.80	0.55	0.82	0.51	0.84
14[c]	0.63	0.83	0.42	0.85	0.29	0.86
15[c]	0.46	0.81	0.44	0.87	0.36	0.90
16[c]	0.65	0.82	0.64	0.85	0.62	0.85
17[b]	0.46	0.84	0.39	0.88	0.42	0.90
18[c]	0.43	0.80	0.37	0.83	0.21	∞
19[b]	0.64	0.82	0.59	0.85	0.54	0.87
20[b]	0.62	0.85	0.56	0.87	0.57	0.87
Mean	0.58	0.82	0.49	0.85	0.42	0.88

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	0.453	0.5133	0.410
	S.D.	0.137	0.0709	0.187
Group b	Mean	0.614	0.4762	0.417
	S.D.	0.123	0.0983	0.124
Group c	Mean	0.598	0.491	0.430
	S.D.	0.089	0.100	0.141
	a vs b	t = 1.78 NS	t = 0.69 NS	t = 0.06 NS
	a vs b	t = 1.72 NS	t = 0.42 NS	t = 0.17 NS
	b vs c	t = 0.30 NS	t = 0.31 NS	t = 0.19 NS

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 5.10(c) is a comparable table to 5.10(a) and (b), when it is hypothesised that **all** carious lesions of any depth "should" be restored. Again, the predictive powers of positive decisions are increased, but the predictive powers of negative decisions decline.

However, although the predictive power of practitioners' positive decisions increase when the gold standard includes shallower lesions, dentists' ability to predict accurately surfaces which do not require restoration is lower. The predictive powers of decisions not to restore are consistently lower in Tables 5.10(b and c) than in Table 5.10(a), at all treatment cut-off criteria.

There were no significant differences in the positive and negative predictive powers between the groups of dentists who intended to restore at different depths of lesions (denoted by superscripts a, b and c). This statement is true regardless of the 'gold standard' used, or the treatment threshold in question.

5.4.9 True positive/negative and false positive/negative treatment decisions

The ratios of true positive/negative and false positive/negative treatment decisions for each dentist are detailed in Appendix IV, at each rating scale level, and for each microscopic validating criteria.

The Tables in Appendix IV indicate the considerable variations between dentists' treatment decisions. They also show that the proportion of true positive decisions increase as the treatment threshold is lowered whilst the proportion of true negative decisions declines. Likewise, as the treatment threshold is lowered the proportion of false positive decisions increases, and false negative decreases. These lengthy Tables are included as an Appendix as they constitute the cells of the 2x2 decision matrices, upon which the further analyses in this Chapter are based.

5.4.10 Areas beneath ROC curves

The ROC curve generated by the pooled data from twenty dentists, who made in total, 7200 treatment decisions is shown in Figure 5.5). The 'gold standard' criterion used in the construction of this figure was a carious lesion into dentine. The Figure shows that when using rating 1 (only teeth deemed to definitely require restoration are considered as positive decisions), the dentists achieved an overall true positive rate of close to 30%. At this level, the false positive rate is low. The Figure demonstrates that if dentists hold low thresholds for the evidence they require before filling a tooth i.e. they restore all tooth surfaces other than those which they feel definitely do not require restoration, the true positive rate is increased to 68%, but at the expense of a false positive rate of 32%. From this Figure it is clear that the optimum criteria to be used depends on the value placed on

Table 5.10(c) Predictive power (PP) of positive and negative treatment decisions using any carious lesion as the validating criterion

Dentist	Rating 1 as positive treatment decisions		Rating 1-3 as positive treatment decisions		Rating 1-5 as positive treatment decisions	
	PP +ve test	PP -ve test	PP +ve test	PP -ve test	PP +ve test	PP -ve test
1[b]	0.71	0.75	0.39	0.81	0.71	-
2[c]	0.67	0.74	0.67	0.74	0.67	0.75
3[b]	0.64	0.76	0.58	0.79	0.55	0.81
4[b]	0.59	0.77	0.49	0.80	0.36	0.84
5[c]	0.64	0.75	0.43	0.78	0.37	0.82
6[a]	0.67	0.76	0.63	0.77	0.65	0.77
7[b]	0.67	0.73	0.58	0.75	0.49	0.75
8[b]	0.59	0.75	0.49	0.79	0.43	0.79
9[a]	0.52	0.74	0.51	0.76	0.52	0.76
10[c]	0.67	0.76	0.61	0.79	0.58	0.81
11[a]	0.33	0.72	0.55	0.78	0.06	-
12[c]	0.69	0.76	0.54	0.78	0.54	0.78
13[c]	0.83	0.73	0.65	0.76	0.60	0.77
14[c]	0.63	0.76	0.34	0.78	0.47	0.78
15[c]	0.53	0.74	0.45	0.79	0.41	0.82
16[c]	0.73	0.76	0.68	0.78	0.66	0.78
17[b]	0.36	0.77	0.46	0.81	0.50	0.84
18[c]	0.50	0.73	0.47	0.77	0.28	-
19[b]	0.72	0.76	0.65	0.79	0.58	0.79
20[b]	0.70	0.79	0.62	0.79	0.64	0.80
Mean	0.62	0.75	0.55	0.78	0.50	0.79

		Rating 1 as positive	Rating 1-3 as positive	Rating 1-5 as positive
Group a	Mean	0.507	0.563	0.7700
	S.D.	0.170	0.061	0.0100
Group b	Mean	0.623	0.533	0.8029
	S.D.	0.118	0.088	0.0315
Group c	Mean	0.654	0.538	0.7887
	S.D.	0.099	0.122	0.0253
	a vs b	t = 1.08 NS	t = 0.65 NS	t = 2.48 NS
	a vs b	t = 1.42 NS	t = 0.48 NS	t = 1.76 NS
	b vs c	t = 0.60 NS	t = 0.10 NS	t = 0.95 NS

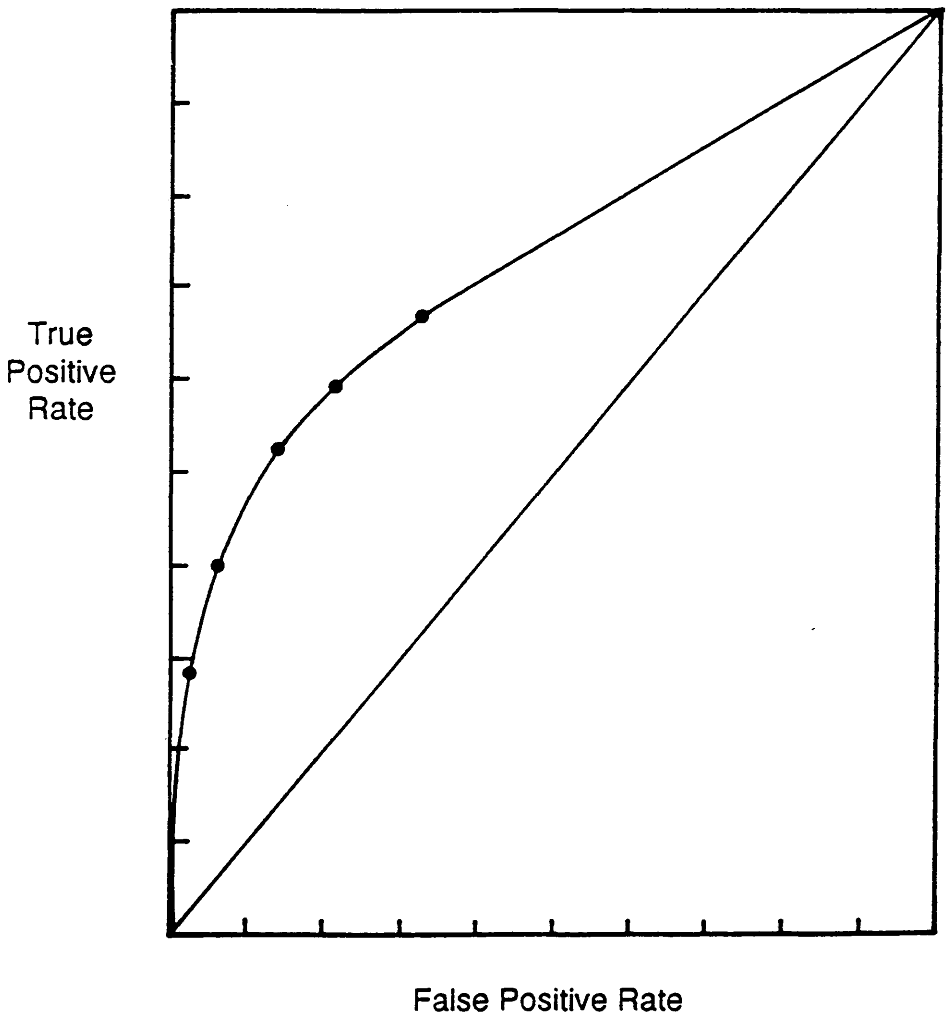
a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Figure 5.5

Receiver Operator Characteristic Curve generated by pooled data (7200 decisions) with caries into dentine as the validating criterion



true positive and false positive decisions, because as the true positive rate declines the false positive rate is inevitably increased.

The ROC curves generated by each individual dentist are shown in Appendix V where a curve is drawn for each dentist and for each microscopic 'gold standard'.

The results of the ROC analysis are detailed in Table 5.11, and show that, for six dentists (marked NS), alteration of the validating criteria did not significantly alter the dentists' discriminatory ability. This suggests that for these dentists, the certainty of decision to restore is not affected by the depth of the lesion. Here the two subjects denoted with 'X' are significantly more likely to distinguish between sound and diseased surfaces when the lesion extends into the inner enamel **OR** dentine. The 14 dentists denoted 'Y' show greater suspicion of lesions into dentine when compared with lesions which penetrate only the outer enamel.

Finally the two dentists denoted 'Z' view surfaces with dentine lesions with greater suspicion than they view surfaces with lesions extending through the enamel, but not penetrating dentine.

The majority of the participants show an ability to distinguish between inner enamel/dentine lesions and lesions in the outer enamel (those marked 'Y') while only 2 dentists demonstrated significant discrimination between inner enamel and dentine lesions. The two dentists who showed greater suspicion of inner enamel than outer enamel lesions, also treated dentine lesions with greater suspicion than inner enamel lesions (those marked 'Z').

Data in Table 5.11 are of interest as they indicates the ability of dentists to discriminate between lesion depths. This does not necessarily imply the dentist will treat deeper lesions but that he will regard them with greater suspicion than shallower ones. If however, his treatment philosophy is a restorative one, he may well decide to treat all lesions, despite any uncertainty he may have.

5.5 DISCUSSION

The results show that of the 344 approximal surfaces for which microscopic validation was available, 96 (28%) had some evidence of caries. This prior probability of disease represents a compromise between the 18% prevalence of decayed and filled approximal surfaces in 15-16 year old children reported by Dummer *et al*[67], and the prevalence or prior probability required for the rating scale experiment, which requires sufficient numbers of signal and noise events to occur in order to construct ROC curves.

Table 5.11 Areas under the ROC curve generated by each dentist

Column A = Lesion into dentine considered 'abnormal' in validating test

Column B = Lesion into inner 1/2 of enamel or into dentine considered 'abnormal' in validating test

Column C = Lesion in any part of enamel or dentine considered 'abnormal' in validating test

Dentist	Areas under ROC curves (see Appendix V)			Sig. Diff.
	A	B	C	
1[b]	0.74	0.69	0.67	N S
2[c]	0.61	0.57	0.59	X ,Y
3[b]	0.72	0.72	0.67	N S
4[b]	0.81	0.72	0.70	X ,Y
5[c]	0.78	0.70	0.66	Y
6[a]	0.70	0.63	0.96	Y, Z
7[b]	0.64	0.59	0.57	N S
8[b]	0.70	0.65	0.63	N S
9[a]	0.64	0.61	0.59	N S
10[c]	0.77	0.71	0.59	Y
11[a]	0.74	0.68	0.64	Y
12[c]	0.75	0.67	0.63	Y
13[c]	0.68	0.62	0.61	N S
14[c]	0.74	0.67	0.62	Y
15[c]	0.76	0.75	0.67	Y
16[c]	0.74	0.67	0.63	Y
17[b]	0.78	0.72	0.69	Y
18[c]	0.65	0.61	0.82	Y, Z
19[b]	0.78	0.70	0.65	Y
20[b]	0.77	0.70	0.67	Y

X denotes sig. diff. ($p<0.05$) between A + B
Y denotes sig. diff. ($p<0.05$) between A + C
Z denotes sig. diff. ($p<0.05$) between B + C

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

It is clear that the visual criteria used to detect caries slightly underestimated the prevalence of microscopic carious lesions. If cavitation is taken as the visual criterion for 'disease positive', and lesion into dentine as the microscopic criteria for 'disease positive', 289 surfaces were considered sound by both methods and 33 carious by both methods. It appears from the results that if cavitation is seen, the probability of dentinal caries is 71%. Conversely the probability of cavitation, given a lesion into dentine is 33/43 = 0.77 or 77%. From these figures it can be seen that the correlation between a dentine lesion and cavitation formation and vice versa, is not perfect. Which of the criteria is more important as an indicator of treatment need is open to speculation.

From the variations seen between dentists in their usage of the rating scale, it seems firstly, that there is little consensus as to when a lesion becomes 'suspicious', and secondly, that some dentists are much more reluctant about taking definite decisions than others. Since it has been established in the review of the literature that bitewing radiographs can only give an indication of the probability of disease, this attitude is perhaps defensible. However, it is also possible that the experimental situation in which the dentists in the study undertook the treatment decisions affected their ability to commit themselves to a definite decision.

The results indicate that the differences in treatment decisions seen between dentists are not due solely to them holding differing criteria as to when to consider restoration. The proportion of positive ratings did not increase systematically with lesion depth. Causes of variation between dentists, other than that due to being more treatment/prevention orientated or by misunderstanding the terms "probably" and "possibly", are explored in chapters 6 and 7.

The trends apparent in Table 5.4, indicate that the dentists' decisions, based on radiographic evidence, did, to some extent, relate to the extent of visually evident caries. This would lead one to believe that the rating scale was being utilised correctly, but that the dentists did not always successfully correlate the radiographic appearance with the presence or absence of cavitation.

It is notable that the dentists' stated criteria as to the depth of lesions they intended to restore did not significantly affect their ability to correctly identify lesions of the depth they wished to treat. Likewise their intentions did not affect the sensitivity, specificity nor predictive power of their decisions, even when differing validation criteria were used.

The distribution of dentists' ratings of treatment need in relation to the microscopic appearance of carious lesions shows that the probability of a positive treatment decision increased with the depth of lesion.

It is not universally accepted that treatment should be instituted only when a lesion reaches dentine. Therefore the analyses in this chapter include validation of the decisions using enamel only lesions as 'gold standards'. Some of the dentists in the study felt that treatment should be instituted before lesions reached dentine (see superscripts) and therefore it was considered necessary to examine whether they were 'correct' more frequently when the depth of lesion which they favoured as an indicator of treatment need, was present. This was not shown to be the case, as the Tables show that all dentists were 'correct' less often when the validating criteria included enamel lesions, and there were never any statistically significant variations between dentists who held differing views as to when a lesion 'should' be restored.

The requirements of a diagnostic test, in terms of sensitivity and specificity, are dependent on whether it is considered more important to avoid unnecessarily treating sound teeth, or more important not to miss diseased teeth. The mean sensitivity values ranged from 0.31 - 0.66 (Table 5.9(a)) and would imply that a considerable number of diseased teeth are not being filled. However, although ideally, sensitivity and specificity should both be high, they are often indirectly proportional to each other. Therefore, although the dentists in this study often failed to restore teeth with lesions into dentine (in 37-69% of cases overall) they would restore teeth which did not have dentine lesions comparatively rarely (5-26%).

The most clinically relevant probabilities mentioned in this analyses are the predictive values. These alter as the prevalence of disease changes, and therefore, the disease prevalence of 28% in this study will increase the predictive power of a positive test beyond that which would pertain when bitewing radiographs from a normal population are read, given that in the UK the prevalence of approximal lesions is lower than 28%. This is a direct consequence of Bayes Theorem, in that the more prevalent a disease is, the probability of being correct when it is detected, increases. This explains why the predictive powers of decisions to restore are always lower when the gold standard is restricted to the more uncommon lesions (those penetrating dentine). As the validating criterion is enlarged to include all lesions, even the shallowest, then the probability of a dentist being 'correct' when he assigns a surface for treatment automatically increases. Therefore it is important not to wrongly interpret these Tables as indicators that dentist 'diagnose' enamel lesions 'better' than deeper ones.

The Tables also show that as a dentist becomes more willing to restore teeth - i.e. if he takes a "when in doubt, fill" attitude - the proportion of teeth he prescribes treatment for, which actually have disease, declines. Likewise, the proportion of teeth he leaves unrestored, which are in reality sound, increases.

With regard to the area under the ROC curves, such a figure represents merely a numerical categorisation of the probability than a randomly selected diseased surface is correctly rated (in the rating scale experiment) as being in greater need of restoration than a randomly selected caries-free surface.

The ROC data reveals that the discriminatory ability of dentists varies according to lesion depth, but not in a consistent fashion. Most dentists showed the highest discrimination between dentine lesions and outer enamel lesions although dentist 6 was discriminating between 'normal' and 'abnormal' almost perfectly, if the task was to discriminate between teeth which had any type of carious lesion, and those which did not.

Several practitioners, did not treat shallower lesions with greater suspicion than deep lesions, resulting in the areas under all three ROC curves (with 3 validating criteria) being similar. This implies that a deeper lesion does not increase the probability that they would prescribe treatment for the tooth. However, the majority of dentists had an increased probability of correctly deciding a tooth did or did not require treatment, when lesions extended into inner enamel or beyond. However, the depth of lesion which they **said** they intended to restore, did not relate well to their ROC curves, implying that the radiological appearance was not the only factor taken into account regarding their decisions.

These results show that dentists' intentions relating to treatment do not predict reliably the sensitivity, specificity or predictive powers of their decisions. It is clear from the ROC analyses that factors other than the dentists' intentions affect their discriminatory powers.

CHAPTER 6. VALUES AFFECTING DENTISTS' TREATMENT DECISIONS

6.1 INTRODUCTION

The probability of occurrence of the four possible outcomes of a dentist's treatment decision (true positive (TP); true negative (TN); false positive (FP); false negative (FN)) can be used to study the relative weight that a dentist attaches to correct decisions. More importantly the data can be used to measure the value the dentist attaches to errors. The ROC curves generated from the data for each participant (Appendix V), can be used to investigate whether their treatment decisions are consistent with their stated values. By selecting an operating point on the curve, and measuring the slope of the ROC curve at that point, the ratio of the costs and benefits implicit in each decision can be derived. By using the ratio of true positive and false positive rates at this point, an expression can also be derived to investigate the dentist's views of the relative costs of 'mistaken' decisions.

6.2 QUALITATIVE ASSESSMENT OF VALUES AFFECTING DENTISTS' TREATMENT DECISIONS

6.2.1 Method

In this first part of the investigation into the values dentists hold concerning the treatment decisions they make, two convenience samples of practitioners were interviewed, in order to identify issues that might affect treatment decision criteria. During these interviews the issues of sensitivity and specificity were raised and the dentists asked to accept that the setting of treatment criteria required one to decide: either whether it is more important always to avoid treating a tooth when it did not require restoration, or whether it is more important always to restore carious lesions. The dentists were also asked to raise any issues that influence the criteria they apply when deciding whether to restore a tooth. The investigator noted any issues that the dentists appeared to consider relevant.

6.2.2 Results

The discussions were free-ranging and participant led. Many dentists appeared to have difficulty in accepting that the setting of a diagnostic criterion involved a trade-off between true positive and false negative rates. The following comments were considered relevant to where the dentist sets his intervention threshold.

> *"I don't sleep at night if I have opened up a tooth which was sound.... I feel as if I'm not a good dentist"*

> *"It isn't a matter of people knowing my mistakes, it's the feeling that I'm not very good"*

"Mistakes make me feel bad, whatever type of mistakes I make"

"No-one ever knows if I've made a mistake, apart from me"

"My patients expect me to treat them.... imagine what they'd say if they came back after six months with toothache, when I'd seen them before and told them they had no problems"

"There is a temptation to not treat if you think the patient will last until the next visit with no problems; patients don't particularly like fillings you know!"

"You get a name for yourself as being a 'filler'"

"The people who come to the dentist care about their teeth - I suppose they care if I get it right then, don't they"

"If you're getting it wrong a lot, you'd get a name for yourself"

"Don't forget people are paying somehow or other for any treatment given - you're wasting the world's money if you treat unnecessarily"

"We can't expect people to keep paying for mistakes"

"Well we're expected to get it right aren't we"

"Dentists are respected really, people wouldn't believe in us if they thought we didn't actually know what was right and what was wrong"

"The dental profession has had a rough ride recently - all this stuff about cavities getting bigger each time you fill - we have to watch what we're doing"

"Since I've been working I've learnt that it causes much more trouble nursing things - the new sealant restorations have made it much easier...."

"Well, you don't want to waste time filling teeth when they don't need it - the money doesn't justify the time you'd waste"

"The money would definitely not affect my diagnosis.... no, definitely not.... it's not that much, not worth it"

"I suppose if you didn't get paid to put a filling in you might think harder about it... it wouldn't change the way I decide through - I never fill teeth if I think it's okay to leave them.... but then most of my patients will come back before any problems, even if I had nursed something - which I don't think I do, very often"

"I'm here to look after people... it's sort of ethics really, to want to get it right"

"Ethically speaking, I suppose it's worse to fill unnecessarily than miss something"

"I don't think the odd mistake is the end of the world, except perhaps if you're into the ethical side of it"

It is impossible to report here the full range of responses, but these selected examples cover some aspect of all the comments made. From these pilot discussions, ten broad issues were identified which might affect the setting of treatment criteria.

These were:

1. Self esteem
2. Patients' preferences
3. Dentist's reputation with colleagues
4. Patients' values
5. Dentists' reputation with patients
6. Professional values
7. Professional responsibility
8. Fees
9. Ethics
10. Societal values

These issues were then used to construct a questionnaire (Appendix VI).

6.3 QUANTITATIVE ASSESSMENT OF VALUES AFFECTING DENTISTS' TREATMENT DECISIONS

6.3.1 Pilot Study

A pilot study was then undertaken using a convenience sample of five dentists to test the acceptability and comprehension of the developed questionnaire. No attempt was made to ascertain the validity of the questionnaire. The pilot study resulted in rewording the questionnaire and adding the "equal value" category. The final questionnaire may be seen in Appendix VI.

6.3.2 METHOD

Twenty randomly selected dentists (the same respondents who undertook the rating scale exercise) were asked to consider the issues listed. They were then asked to indicate how each issue might influence their treatment planning. The dentists then completed the rating scale experiment (detailed in Chapter 5) to make it possible to relate their responses to the questionnaire to their performance when making treatment decisions.

6.3.3 Analysis and Results

The numbers of dentists responding "A" (more important **always** to fill carious teeth), the numbers responding "B" (more important **never** to unnecessarily fill sound teeth) or "C" (avoiding errors of any kind is of equal importance), for each of the issues identified in the qualitative research as being relevant to treatment decisions are indicated in Table 6.1.

With the exception of the issues of self-esteem and professional responsibility it was also shown that the majority of subjects weighted errors equally (column 'c'). For the remaining issues with the exception of "reputation with patients", a majority reported it was more important never to unnecessarily fill sound teeth than always to fill carious teeth. This trend was most notable with regard to dentists' feelings of self-esteem and their patients' preferences.

Data in Table 6.2 indicate that when all issues were considered together three dentists (Nos **5**, **12**, and **13**) placed greater value on treating caries (sensitivity) than avoiding overtreatment (specificity). Other dentists placed varying degrees of 'weight' upon the avoidance of unnecessary treatment - most notably dentist number 11, who placed all the weight on this choice.

The slope of the ROC curve and the likelihood estimate was calculated for each dentist at three of his operating points.

It can be shown that the ratio of true positive to false positive decisions is related to the weight dentists place on errors, and the value they place on correct decisions[241]. i.e. treatment criteria are shifted according to dentists' views of the importance of the TP, FN, FP, TP outcomes of their decisions.

Table 6.1 Dentists' responses regarding issues which may affect their treatment decision-making

(Sample size = 20)

ISSUE		A More important to always fill carious teeth	B More important to never unnecessarily fill sound teeth	C Avoiding erros of either kind of equal importance
1.	SELF-ESTEEM	2	11	7
2.	PATIENTS' PREFERENCES	0	10	10
3.	REPUTATION WITH COLLEAGUES	2	7	11
4.	VALUE TO PATIENT	2	7	11
5.	REPUTATION WITH PATIENTS	4	3	13
6.	VALUE TO SOCIETY	1	8	11
7.	VALUE TO PROFESSION	2	5	13
8.	PROFESSIONAL RESPONSIBILITY	3	9	8
9.	PROFESSIONAL FEE	4	6	10
10.	ETHICAL CONSCIENCE	1	7	12

Table 6.2 Individual dentists' views concerning issues which may affect their treatment decision-making

DENTIST	A More important to always fill carious teeth	B More important to never unnecessarily fill sound teeth	C Avoiding errors of either kind of equal importance
1	-	3	7
2	-	5	5
3	2	2	6
4	-	-	10
5	5	-	5
6	-	1	9
7	-	7	3
8	-	4	6
9	-	4	6
10	-	7	3
11	-	10	0
12	7	2	1
13	6	4	0
14	1	6	3
15	1	4	5
16	1	1	8
17	1	2	7
18	2	7	1
19	1	2	7
20	1	1	8

$$\text{Slope of ROC at given interventive threshold} = \left(\frac{p(TN)}{p(TP)} \times \frac{(V_{TN}) + (C_{FP})}{(V_{TP}) + (C_{FN})}\right)$$

where	V_{TN}	=	the value of true negative outcomes
	C_{FP}	=	the costs of false positive outcomes
and	V_{TP}	=	the value of true positive outcomes
	C_{FN}	=	the costs of false negative outcomes

Pursuing this argument

$$\underline{\text{Slope of ROC}} = \left(\frac{p(TP)}{p(TN)} \times \frac{(V_{TN}) + (C_{FP})}{(V_{TP}) + (C_{FN})}\right)$$

Thus, the ratio of the value of true negatives and costs of false positives, to the value of true positives plus the costs of false negative treatment decisions, can be derived by examining the slope of the curve and the $\frac{TP}{TN}$ ratio at each operating level.

The values of the $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ ratios calculated from each dentist's ROC curve are shown in Table 6.3(i) for the data derived when caries into dentine was used as the validating criterion and rating (1) was considered as the level at which the dentist would restore teeth. The value of the ratio could not be computed for dentists who generated ROC curves with empty cells, - empty cells being produced when one or other of the rating scale categories was unused by the dentist.

The value of the ratios, for each dentist with the same validating criteria as above, but using the TP|TN|FP|FN ratios generated when the dentists held looser criteria for a decision to restore - i.e. ratings 1-3 in the rating scale experiment were considered as decisions to restore, are shown in Table 6.3(ii).

The information in Table 6.3(iii) is similar, but uses the ratios generated when the dentists would restore every surface other than those which they categorised as "definitely leave unrestored". The tables 6.3(i-iii) show the true positive, true negative and false positive rates of the dentists' treatment decisions. It also shows the computed ratios of $\frac{TP}{TN}$ and the slope of the curve at this level of intervention, (otherwise known as the likelihood ratio $\frac{TP}{FP}$).

Table 6.3(i) Dentists' true positive, true negative and false positive rates, and derived values of treatment decisions, with *rating 1* considered as a decision to treat, and dentinal caries as the validating criterion

Dentist	TRUE POSITIVE RATE	TRUE NEGATIVE RATE	1-Spec FALSE POSITVE RATE	$\frac{TP}{TN}$	Slope of ROC	$\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$
1	0.27	0.90	0.10	0.30	25.0	7.5
2	0.18	0.99	0.01	0.18	2.2	0.4
3	0.38	0.95	0.05	0.40	2.5	1.0
4	0.47	0.92	0.08	0.50	5.0	2.5
5	0.32	0.97	0.03	0.33	10.0	3.3
6	0.36	0.95	0.05	0.38	20.0	7.6
7	0.14	0.99	0.01	0.14	3.6	0.5
8	0.32	0.94	0.06	0.34	5.0	1.7
9	0.21	0.96	0.04	0.22	6.0	1.3
10	0.43	0.95	0.05	0.45	20.0	9.0
11	0.05	0.99	0.01	0.05	10.0	0.5
12	0.41	0.95	0.05	0.43	50.0	21.5
13	0.09	0.99	0.01	0.09	4.0	0.4
14	0.41	0.95	0.05	0.43	5.0	2.2
15	0.27	0.95	0.05	0.28	10.0	2.8
16	0.34	0.96	0.04	0.35	3.6	1.26
17	0.52	0.88	0.12	0.59	4.3	2.5
18	0.11	0.94	0.06	0.11	4.9	0.54
19	0.36	0.97	0.03	0.37	5.3	2.0
20	0.57	0.93	0.07	0.61	3.3	2.0

Table 6.3(ii) Dentists' true positive, true negative and false postive rates, and derived values of treatment decisions, with *ratings 1-3* considered as a decision to treat, and dentinal caries as the validating criterion

Dentist	TRUE POSITIVE RATE	TRUE NEGATIVE RATE	1-Spec FALSE POSITIVE RATE	$\frac{TP}{TN}$	Slope of ROC	$\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$
1	0.66	0.60	0.40	1.1	1.3	1.4
2	0.12	0.98	0.02	0.12	0.2	0.03
3	0.49	0.89	0.11	0.55	0.2	0.1
4	0.5	0.79	0.21	0.70	1.3	0.9
5	0.49	0.76	0.24	0.64	2.0	1.28
6	0.30	0.95	0.05	0.31	10.0	3.1
7	0.19	0.96	0.04	0.20	2.4	0.48
8	0.48	0.83	0.17	0.58	1.3	0.8
9	0.32	0.90	0.10	0.36	0.9	0.3
10	0.47	0.90	0.10	0.52	2.2	1.1
11	0.43	0.89	0.11	0.48	1.0	0.5
12	0.44	0.88	0.12	0.5	1.5	0.8
13	0.23	0.95	0.05	0.24	1.7	0.4
14	0.45	0.83	0.17	0.54	1.0	0.5
15	0.53	0.82	0.18	0.65	0.8	0.5
16	0.38	0.94	0.06	0.40	2.6	1.0
17	0.60	0.75	0.25	0.80	1.1	0.9
18	0.37	0.83	0.17	0.45	1.2	0.5
19	0.41	0.92	0.08	0.45	4.0	1.8
20	0.47	0.90	0.10	0.52	1.0	0.5

Table 6.3(iii) Dentists' true positive, true negative and false postive rates, and derived values of treatment decisions, with ratings 1-5 considered as a decision to treat, and dentinal caries as the validating criteria

Dentist	TRUE POSITIVE RATE	TRUE NEGATIVE RATE	1-Spec FALSE POSITIVE RATE	$\frac{TP}{TN}$	Slope of ROC	$\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$
1	1.00	-	-	-	-	-
2	0.25	0.94	0.06	0.27	0.05	0.01
3	0.68	0.82	0.08	0.83	0.05	0.04
4	0.91	0.45	0.55	2.02	0.37	0.74
5	0.86	0.52	0.48	1.65	0.47	0.78
6	0.45	0.93	0.07	0.48	1.0	0.48
7	0.36	0.90	0.10	0.40	0.7	0.28
8	0.39	0.74	0.26	0.53	0.69	0.37
9	0.39	0.88	0.12	0.44	0.7	0.31
10	0.66	0.84	0.16	0.79	0.4	0.32
11	1.00	∞	-	-	-	-
12	0.61	0.87	0.13	0.70	0.45	0.31
13	0.43	0.91	0.09	0.47	0.65	0.31
14	0.73	0.57	0.43	1.28	0.35	0.45
15	0.77	0.64	0.36	1.00	0.5	0.50
16	0.55	0.92	0.08	0.60	0.5	0.30
17	0.80	0.69	0.31	1.16	0.5	0.58
18	1.00	∞	-	-	-	-
19	0.66	0.87	0.13	0.76	0.4	0.30
20	0.64	0.88	0.12	0.73	0.4	0.30

Application of the equation-slope x $\frac{p(TP)}{p(TN)} = \frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ gives the values shown in the extreme right hand column of the Tables 6.3(i-iii). The right hand column of the Table shows that many of the values for $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ were greater than 1.0. This implies that if dentists only restored teeth which they felt definitely required restoration they are acting in a way which weights the value of a true negative decision, plus costs of a false positive decision, more highly than the value of a true positive plus the costs of a false negative decision.

The values implicit in restorative decisions, when all teeth for which a restoration is "considered" would actually be restored, are given in Table 6.3 (ii). The table shows that the implicit value of true negative plus the cost of false positives, decreases in relation to the value of true positive and false negative treatment decisions, when the criterion for treatment is lowered.

Reducing the strictness of the criteria at which a dentist would restore a surface, so that all ratings apart from a decision definitely **not** to restore are considered as positive decisions, increases the sensitivity of treatment decisions and thus the true positive rate increases. The computed values of true negative, and costs of false positive decisions are shown in the right hand columns of Tables 6.3(i-iii). When the data in these Tables are compared, it can be seen that the ratios shown in the right hand column decrease, as the decision criteria are reduced in strictness with the median value of the ratio falling to 0.31 in Table 6.3(iii).

From the table 6.3(i) it can be seen that for dentist 1, $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ is equal to 7.5. This implies that for treatment decisions made at this cut-off level, this dentist rates the value of a true negative plus the associated cost of a false positive treatment decision 7.5 times more highly than he rates the value of a true positive decision plus the cost of a false negative treatment decision. By knowing this weighting, and comparing it to his beliefs, the dentist might wish to reduce the strictness of his decision criteria, as this dentist stated that $V_{TN} + C_{FP}$ was only just more valuable to him than $V_{TP} + C_{FN}$.

One can dissect out the components of the equation further by making certain assumptions. If one assumes that the costs of a false positive treatment decision (an unnecessarily treated tooth) are equivalent to the cost of false negative (a tooth with caries into dentine which is left untreated), the equation resolves so that the value in the last column of the Table is the weight the clinician places on true negative decisions in relation to true positives, i.e. if the assumption is correct, then dentist number 1 must consider that he is prepared to trade 7.5 true positive decisions for one true negative. Similar derivations can be made for each dentist.

Conversely if the $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ ratio is less than 1, then the dentist, when operating at the given criteria, values true positive decisions plus the costs of false negatives more highly than the value of true negative plus costs of false positives. i.e. he is happier to accept the consequences of overtreatment than the consequences of undertreatment. Taking the same assumption as previously (that the costs of false positive diagnoses are equivalent to the costs of false negatives), a dentist whose ratio is less than 1, values a true positive diagnosis more highly than a true negative: such a dentist will risk treating some teeth unnecessarily in order to try to ensure that all carious teeth are assigned to the "for treatment" category. Data in Table 6.4 compares the data from Tables 6.3(i-iii) regarding the computed values for the values of TN and costs of FP, in relation to the values of TP and costs of FN, for each dentist, using three cut-off points of rating scale. It also shows the dentists' reported ratios for this value, as measured by the questionnaire (Appendix VI). In this Table, if the calculated value is greater than the reported value, (seen in the right hand column) the dentists who intervene at this level of certainty are undertreating in relation to their stated attitudes. If the calculated value is lower than the reported value, the dentists would be overtreating in relation to their reported values.

It can be seen in Table 6.4 that only three dentists (Nos. 5,12,13) **claimed** to hold an attitude, where $V_{TP} + C_{FN}$ was considered to be more important than $V_{TN} + C_{FP}$, yet their treatment decisions and the implicit weightings were broadly similar to those of other dentists. Most weighted their decisions in a way which was not congruent with their reported attitudes when they held 'rating 3' as their cut-off criteria.

It can also be seen in Table 6.4 that when a "definitely restore" decision is used as the supposed cut-off criteria at which restorative treatment would actually take place, six dentists (marked * in the left hand column) would be overtreating in relation to his stated attitudes.

If "might consider restoration" to "definitely restore" (i.e. ratings 1-3) are considered as positive treatment decisions (Table 6.4, second column) fourteen participants would be overtreating in relation to their stated attitudes. If dentists were very "treatment orientated", that is if any decision other than "definitely leave unrestored" would lead to treatment of the surface in question, all but one of the subjects would be overtreating in relationship to reported attitudes.

As a whole, in this Table (6.4), any values less than 1 imply that the dentist's attitude to treatment is such that when operating at the treatment threshold shown, he rates the sum

Table 6.4 Derived values for $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ ratio at 3 operating levels and the comparable reported values

DENTIST	When rating 1 is taken as the operating level $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$	When ratings 1-3 is taken as the operating level $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$	When ratings 1-5 is taken as the operating level $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$	Reported $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$
1	7.5	1.4	-	1.3
2	*0.4	*0.03	*0.01	2.0
3	1.0	*0.1	*0.64	1.0
4	2.5	*0.9	*0.74	1.0
5	3.3	1.28	*0.78	0.5
6	7.6	3.1	*0.48	1.11
7	*0.5	*0.48	*0.28	3.3
8	*1.7	*0.8	*0.37	2.5
9	*1.3	*0.33	*0.31	2.5
10	9.0	*1.1	*0.32	3.3
11	*0.5	*0.5	-	3.0
12	21.5	0.8	*0.31	0.29
13	*0.4	*0.4	*0.31	0.66
14	2.2	*0.5	*0.45	2.25
15	2.8	*0.5	*0.50	1.5
16	1.26	1.0	*0.30	1.0
17	2.5	*0.9	*0.58	1.13
18	*0.54	*0.5	-	2.67
19	2.0	1.8	*0.30	1.13
20	2.0	*0.5	*0.30	1.0

* denotes calculated $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ is of lower value than the reported $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ value.

of the value of true positive decisions and costs of false negatives, as being greater than the sum of the value of true negative decisions plus the cost of false positive decisions. i.e. he feels that sensitivity is of more importance when making treatment decisions, than specificity.

6.4 DISCUSSION

Working on the assumption that all dentists would consider that a lesion into dentine "should" be restored, various conclusions can be derived from the results.

The decision 'rules' employed by dentists clearly do not maximize 'correct' decisions. If this were so the maximally effective decision rule would apply where $\frac{TP}{FP}$ was greatest. i.e. at the point on the ROC curve nearest to the top left hand corner, where the slope of the curve is close to 1. However, different rewards and costs can be attached to each outcome of a decision (FP, FN, TP, TN).

Dentists may weight the value of a true negative decision plus the cost of a false positive decision highly, i.e. they would much prefer to risk missing a few lesions (as the costs associated with doing so are high) rather than risk reducing their level of true negative decisions. A dentist will therefore "undertreat" or "overtreat" according to the importance he attaches to sensitivity and specificity.

It can be argued that the assumption that V_{TN} is equivalent to V_{TP} is untenable. Such an assumption implies that the value of a totally sound unrestored surface is equivalent to a treated surface. If this assumption were correct then all surfaces would be restored, regardless of the evidence for caries. Such an assumption is considered unrealistic, thus this line of argument will not be pursued.

The explanation of the results, and the one which coincides with dentists' reported attitudes, is that dentists purposefully weight their decisions according to their views about the importance of sensitivity and specificity. i.e. they intentionally do not operate at the maximum ratio as their reported values show that 'correct' decisions (TP and TN) are not the most important consideration in the decision to restore a surface. The "best" point on the curve for the dentist to operate, is therefore, the point at which the probability of TP,TN,FP,FN generated by his decisions, are congruent with his views about the importance of sensitivity and specificity.

One question which must be considered when analysing the results of this study is the validity of the questionnaire examining the dentists' treatment preferences. Although the document was developed via a qualitative issue-identification methodology, it is unclear as to whether or not the questionnaire covered a sufficiently broad number of issues. Also it is possible that dentists might "answer-to-please" due to recent media attention and government enquiries about overtreatment by dentists. It is assumed that this was not the case as the dentists were assured of confidentiality and were invited to complete the questionnaire at their leisure, unobserved by the investigator.

CHAPTER 7. VARIATION BETWEEN DENTISTS' TREATMENT DECISIONS

7.1 INTRODUCTION

It is clear from the review of the literature and the two preceding Chapters, that individual dentists' treatment decisions vary widely. This variation may stem from differing perceptions of the radiographic image, differing views as to the appropriate stage at which a lesion should be restored, differing views as to how a given depth of lesion is represented on a radiograph, and differing attitudes concerning specificity and sensitivity.

This Chapter aims to measure the differences in treatment planning between dentists, and to determine whether this variation is random, or systematic.

7.2 METHOD

From the data gathered in the rating scale experiment, each individual dentist's decisions were compared to those made by every other dentist. The comparison was examined using the Kappa statistic, (see Table 7.1 for description). The numbers and proportion of agreements between each dentist pair were also examined.

Following this analysis, the proportion of dentist pairs agreeing on treatment/non treatment, for each individual tooth surface were examined. In this way, surfaces causing the greatest discrepancies in treatment planning could be determined.

The radiographic appearance of the tooth surfaces can be seen in Figure 4.6.

7.3 RESULTS

Of the 20 dentists in the study, three (dentists 6, 9, 11) stated their intention was to restore all lesions extending into the inner half of the enamel but which had not reached the amelo-dentinal junction. In this Chapter these dentists will be denoted by a superscript '(a)'. A further eight dentists indicated that lesions extending to but not beyond the amelo-dentinal junction should be treated restoratively and they will be denoted with superscript '(b)'. The remaining nine dentists felt that restorative treatment should only be instituted when the lesion had penetrated dentine (dentists denoted with superscript '(c)').

If all decisions made by each individual dentist for each of the 360 tooth surfaces are compared to all the decisions made by other dentists for that surface, a total of 68400 pairwise comparisons are possible. Of these, 34656 decisions (51%) were in agreement.

The Kappa values for the agreements between dentist pairs for each category of the rating scale are shown in Table 7.2. For decisions to definitely restore a tooth surface, the Kappa

Table 7.1 The derivation and meaning of the Kappa statistic

		Dentist 1		
		Leave unrestored	Restore	Total
Dentist 2	Leave unrestored	a	c	a+c
	Restore	b	d	b+d
		a+b	c+d	a+b+d+c+d

p_o = observed agreement (a+d)

p_e = agreement expected by chance = sum of expected agreement on no-treatment (a+c) x(a+b) and expected agreement on treatment (b+d) x (c+d)

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e}$$

$$= \frac{(a+d) - [(a+c)(a+b)] + [(b+d)(c+d)]}{1 - [(a+c)(a+b)] + [(b+d)(c+d)]}$$

$$Z = \frac{\text{Kappa}}{\text{S.E. (Kappa)}}$$

Table 7.2 Kappa statistic for agreement between dentists and probability of second observer agreeing with first, for each category of rating scale

Rating	Number of decisions	Kappa	Z	Probability of 2nd observer agreeing with 1st
1	574	0.459	37.8*	0.502
2	311	0.081	9.3*	0.121
3	547	0.053	4.5*	0.125
4	541	-0.002	-0.18	0.073
5	724	-0.009	-0.63	0.093
6	4503	0.196	3.3*	0.699

* indicates Kappa is significantly different from zero ($p<0.01$)

value for dentist pairings was 0.459. The dentists pairings generated Kappa values of 0.081 or less when ratings 2-5 are considered, and the Kappa values for decisions to "definitely leave the tooth unrestored" (rating 6) gave a Kappa value of 0.196. The negative figures for the Kappa statistic for ratings 4 and 5, in Table 7.2, indicate that the dentist pairs disagreed more frequently than would be expected by chance when they used these categories. There were 1265 occasions on which ratings 4 and 5 were used. Therefore, the dentists were using these categories in different ways. In the case of a decision to definitely restore a tooth, the probability of a second observer agreeing with the first was 0.502. For decisions to definitely leave a tooth unrestored the chance of a second observer's opinion concurring with that of a first observer was 0.699 (Table 7.2). The probability of a second observer agreeing with a first, for ratings 2-5, ranged between 0.073 and 0.125.

The data were dichotomised to positive and negative decisions using rating 1 as a decision to restore, and ratings 2-6 as decisions to leave a tooth unrestored. There were 574 decisions to restore, and 6626 to leave the tooth surface unrestored and the Kappa score for agreements between dentist pairs, using this method of dichotomising the data, was equal to 0.459. The probability of a second observer agreeing with a first for a decision to leave a tooth unrestored was 0.957, whilst for a decision to restore, the probability of a second observer agreeing with the first was 0.502.

There were 190 possible pairings of dentists and Tables 7.3 (i-iii) show the Kappa values achieved between each possible pair of dentists, using different rating scale cut-off levels to define positive and negative treatment decisions. The highest Kappa value (0.734) is seen between dentists 8 and 10 in Table 7.3(i). The lowest Kappa value was 0.00 (seen between several dentist pairings in Table 7.3(iii) (where ratings 1-5 constitute a positive decision, and rating 6 a negative). Dentist No 11 never showed a Kappa score of greater than zero in this Table (7.3(iii)) as he never utilised rating 6. Dentist 18 and dentist 1 also showed extremely low levels of agreements with other dentists when only rating 6 was considered as a negative treatment decision.

Each of the Tables 7.3(i-iii) indicates (with superscripts a, b and c.) the depth of lesions at which each dentist intends to restore teeth (see first paragraph of 'Results'). Thus, Table 7.3(i) shows that dentists 10 and 3 achieved a Kappa score of 0.687 for their agreement, although they had different views as to the depth of lesion which "should" be restored. Dentist 10 felt that lesions should be definitely into dentine before restorations are placed, whilst dentist 3 felt that the lesion need only have penetrated the full thickness of enamel before it should be restored. Similarly dentists 10 and 8 who achieved a score of Kappa = 0.734, (Table 7.3(i)) held differing views as to when a lesion should be restored, and yet they demonstrated the highest level of agreement achieved between any pair of dentists.

Table 7.3(i) Kappa values for agreement between dentists, for each dentist pairing (Rating 1 as a positive decision, Rating 2-6 as negative decisions)

	1[b]	2[c]	3[b]	4[b]	5[c]	6[a]	7[b]	8[b]	9[a]	10[c]	11[a]	12[c]	13[c]	14[c]	15[c]	16[c]	17[b]	18[c]	19[b]	20[b]
1[b]	-																			
2[c]	0.403																			
3[b]	0.561	0.211																		
4[b]	0.478	0.346	0.505																	
5[c]	0.582	0.486	0.420	0.521																
6[a]	0.537	0.415	0.566	0.598	0.514															
7[b]	0.463	0.442	0.293	0.274	0.553	0.371														
8[b]	0.515	0.312	0.513	0.660	0.384	0.652	0.308													
9[a]	0.516	0.403	0.441	0.447	0.443	0.576	0.526	0.515												
10[c]	0.640	0.363	0.687	0.671	0.466	0.599	0.362	0.734	0.450											
11[a]	0.267	0.171	0.242	0.164	0.247	0.228	0.459	0.216	0.413	0.222										
12[c]	0.614	0.415	0.633	0.626	0.551	0.578	0.418	0.557	0.576	0.631	0.177									
13[c]	0.325	0.160	0.232	0.195	0.368	0.269	0.545	0.255	0.254	0.262	0.452	0.269								
14[c]	0.498	0.372	0.533	0.542	0.514	0.545	0.418	0.494	0.498	0.663	0.228	0.643	0.269							
15[c]	0.451	0.401	0.600	0.542	0.399	0.577	0.352	0.622	0.614	0.600	0.249	0.679	0.239	0.510						
16[c]	0.421	0.459	0.435	0.523	0.480	0.657	0.413	0.633	0.586	0.544	0.200	0.520	0.247	0.451	0.624					
17[b]	0.309	0.226	0.504	0.502	0.346	0.558	0.254	0.450	0.386	0.434	0.116	0.441	0.139	0.415	0.393	0.448				
18[c]	0.417	0.102	0.311	0.285	0.280	0.383	0.296	0.407	0.475	0.373	0.488	0.294	0.364	0.383	0.413	0.328	0.151			
19[b]	0.643	0.408	0.588	0.592	0.699	0.637	0.520	0.542	0.643	0.625	0.293	0.564	0.344	0.564	0.524	0.574	0.480	0.368		
20[b]	0.484	0.293	0.508	0.607	0.436	0.598	0.327	0.632	0.514	0.615	0.153	0.624	0.183	0.463	0.489	0.553	0.584	0.235	0.593	

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 7.3(ii) Kappa values for agreement between dentists for each dentist pairing (Rating 1-3 as a positive decision, Rating 4-6 as negative decisions)

	1[b]	2[c]	3[b]	4[b]	5[c]	6[a]	7[b]	8[b]	9[a]	10[c]	11[a]	12[c]	13[c]	14[c]	15[c]	16[c]	17[b]	18[c]	19[b]	20[b]
1[b]	-																			
2[c]	0.102																			
3[b]	0.298	0.253																		
4[b]	0.281	0.177	0.522																	
5[c]	0.277	0.188	0.506	0.434																
6[a]	0.239	0.381	0.498	0.406	0.376															
7[b]	0.125	0.387	0.334	0.179	0.240	0.464														
8[b]	0.270	0.212	0.434	0.577	0.377	0.470	0.238													
9[a]	0.267	0.313	0.548	0.524	0.412	0.609	0.388	0.546												
10[c]	0.332	0.279	0.670	0.574	0.481	0.539	0.366	0.520	0.631											
11[a]	0.272	0.208	0.499	0.415	0.401	0.509	0.353	0.449	0.625	0.519										
12[c]	0.286	0.311	0.588	0.503	0.473	0.533	0.415	0.446	0.641	0.573	0.493									
13[c]	0.193	0.423	0.523	0.307	0.296	0.539	0.588	0.399	0.488	0.590	0.438	0.447								
14[c]	0.192	0.249	0.495	0.476	0.518	0.460	0.256	0.486	0.433	0.481	0.424	0.473	0.365							
15[c]	0.301	0.198	0.519	0.535	0.421	0.409	0.221	0.520	0.501	0.588	0.471	0.468	0.358	0.414						
16[c]	0.230	0.370	0.537	0.435	0.388	0.670	0.387	0.485	0.536	0.537	0.508	0.510	0.488	0.495	0.423					
17[b]	0.237	0.121	0.484	0.460	0.392	0.358	0.149	0.437	0.330	0.420	0.389	0.425	0.260	0.465	0.424	0.397				
18[c]	0.218	0.138	0.485	0.427	0.427	0.441	0.284	0.369	0.528	0.469	0.481	0.462	0.358	0.411	0.360	0.402	0.326			
19[b]	0.238	0.277	0.580	0.533	0.422	0.649	0.409	0.470	0.632	0.562	0.576	0.673	0.528	0.426	0.460	0.551	0.354	0.428		
20[b]	0.255	0.248	0.590	0.599	0.445	0.597	0.361	0.612	0.623	0.632	0.608	0.585	0.489	0.458	0.566	0.571	0.456	0.445	0.656	

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Table 7.3(iii) Kappa values for agreement between dentists for each dentist pairing (Rating 1-5 as positive decision, Rating 6 as negative decision)

	1[b]	2[c]	3[b]	4[b]	5[c]	6[a]	7[b]	8[b]	9[a]	10[c]	11[a]	12[c]	13[c]	14[c]	15[c]	16[c]	17[b]	18[c]	19[b]	20[b]
1[b]	-																			
2[c]	0.002																			
3[b]	0.007	0.276																		
4[b]	0.19	0.096	0.277																	
5[c]	0.026	0.89	0.359	0.442																
6[a]	0.003	0.498	0.486	0.176	0.197															
7[b]	0.003	0.231	0.325	0.062	0.107	0.395														
8[b]	0.002	0.206	0.459	0.334	0.310	0.427	0.176													
9[a]	0.004	0.377	0.522	0.209	0.215	0.622	0.344	0.487												
10[c]	0.006	0.286	0.612	0.313	0.286	0.571	0.269	0.589	0.607											
11[a]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000										
12[c]	0.005	0.379	0.580	0.220	0.268	0.559	0.436	0.432	0.619	0.563	0.000									
13[c]	0.004	0.434	0.584	0.173	0.196	0.609	0.440	0.372	0.498	0.618	0.000	0.470								
14[c]	0.020	0.129	0.379	0.274	0.419	0.238	0.141	0.382	0.251	0.377	0.000	0.330	0.285							
15[c]	0.007	0.134	0.396	0.380	0.374	0.264	0.097	0.397	0.282	0.404	0.000	0.337	0.292	0.390						
16[c]	0.004	0.449	0.483	0.181	0.216	0.640	0.356	0.438	0.528	0.548	0.000	0.516	0.559	0.239	0.281					
17[b]	0.013	0.150	0.471	0.332	0.341	0.330	0.150	0.477	0.285	0.452	0.000	0.430	0.307	0.404	0.393	0.337				
18[c]	0.397	0.000	0.002	0.008	0.006	0.001	0.001	0.002	0.001	0.002	0.000	0.001	0.001	0.005	0.004	0.001	0.003			
19[b]	0.005	0.313	0.602	0.216	0.234	0.587	0.361	0.437	0.607	0.569	0.000	0.634	0.516	0.266	0.379	0.522	0.369	0.001		
20[b]	0.005	0.299	0.566	0.247	0.285	0.622	0.310	0.596	0.601	0.633	0.000	0.558	0.550	0.323	0.415	0.556	0.433	0.001	0.619	

a - denotes dentists who feel it is appropriate to restore approximal carious lesions before the lesion has reached the amelo-dentinal junction

b - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion has reached the amelo-dentinal junction but before it penetrates dentine

c - denotes dentists who feel it is appropriate to restore approximal carious lesions when the lesion extends into dentine

Examiner 6 had a widely disparate treatment philosophy from examiner 10, yet the decisions made by these two dentists achieved a Kappa score of 0.599 when the pairwise comparison of treatment decisions was made. Similarly, dentist nine, who was one of the three who felt that a lesion need not have extended through the full thickness of the enamel before it required restoration, showed high agreement with almost all other dentists.

The proportion of the 190 dentists pairings which fell into each Kappa range 0-0.2, 0.21-0.4, 0.41-0.6 and 0.61-0.8 is shown in Figure 7.1. No dentist pairing achieved a Kappa value greater than 0.8, a level of agreement which represents "good agreement" (i.e. Kappa > 0.8).

When the proportions of dentist pairs showing high and low Kappa scores are examined, and when rating 1 is considered to be a positive treatment decision, and 2-6 a negative decision, 68% of the dentist pairings showed moderate or substantial agreement (K>0.4). This Figure (7.1) also shows that when ratings 1-3 were considered to be a positive decision, 65% of the pairings showed a Kappa score greater than 0.4. Finally, it can be seen in Figure 7.1 that when ratings 1-5 were considered as positive decisions, only 30% (51) of the pairings gave a Kappa score greater than 0.4.

The surfaces for which less than half of the dentist pairings agreed, when rating 1 was taken as a decision to restore, and ratings 2-6 as a decision to leave the surface unrestored, are listed in Table 7.4(i). The radiographic appearance of these surfaces can be seen in Figure 4.6 and the visual and microscopic appearances of these disputed surfaces are shown in Tables 7.4(i-iii). Of the sixteen surfaces where less than half the dentists agreed on a **definite** need for treatment, (Table 7.4(i)) eight (50%) exhibited cavitation when examined visually, and seven (45%) were seen to have lesions into dentine when examined histologically. The remaining disputed sites were sound or had lesions in enamel only.

The surfaces in which less than half the dentist pairings agreed, when ratings 1-3 were considered as a decision to treat and ratings 4-6 a decisions to leave unrestored (see Figure 4.6 for radiographic appearance) are presented in Table 7.4(ii). Of the 21 surfaces which were contentious when using this method of dichotomising the data, four (19%) had cavitated lesions, and three (14%) had lesions which were shown to have penetrated dentine when examined microscopically.

The surfaces in which less than half the dentist pairings agreed on treatment, when ratings 1-5 were taken as a decision to treat and rating 6 as a decision to leave unrestored are listed in Table 7.4(iii). The treatment required for 65 surfaces was agreed upon by less

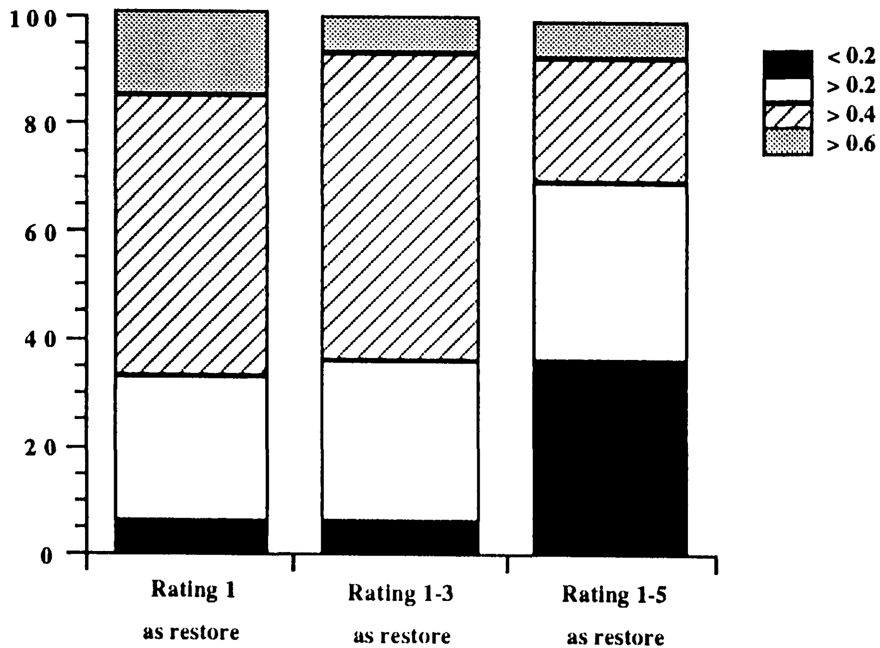


Figure 7.1 Proportion of dentist pairings (N=190) achieving Kappa scores <0.2, >0.2, >0.4, >0.6.

Table 7.4(i) The visual and microscopic appearance of tooth surfaces for which less than 50% of the dentist pairings agreed that the surface definitely required treatment

(Rating 1 as a decision to restore, 2-6 as decisions to leave unrestored)

Radiograph	Tooth	Visual appearance*	Microscopic appearance*
4	4d	0	0
4	6m	0	1
7	4d	2	3
7	5m	2	0
8	6m	3	3
8	6d	0	0
8	5m	3	9 (missing)
8	5d	4	9 (missing)
10	5m	4	9 (missing)
10	4d	0	0
10	7m	3	3
10	6m	7	1
11	6d	3	4
13	6d	0	3
14	5d	4	3
15	7m	3	3

* See Tables 5.1 and 5.2 for details of criteria for assigned scores.

Table 7.4(ii) The visual and microscopic appearance of the tooth surfaces for which less than 50% of the dentists agreed that treatment was, or could be considered necessary

(Ratings 1-3 considered as a decision to restore, ratings 4-6 as decisions to leave unrestored)

Radiograph	Tooth	Visual appearance*	Microscopic appearance*
1	7m	0	0
2	5d	5	0
3	5d	3	2
3	5d	0	0
4	6m	1	2
4	6m	0	0
4	7m	0	0
7	5m	5	5
7	4d	0	0
8	5d	0	0
8	7m	3	2
9	6d	0	0
9	5m	2	5
10	6m	0	0
11	7m	0	0
11	5d	0	0
12	4d	2	3
13	5d	2	0
13	4d	3	3
13	4d	4	4
14	6m	0	0

* See Tables 5.1 and 5.2 for details of criteria for assigned scores.

Table 7.4(iii) The visual and microscopic appearance of surfaces for which less than 50% of the dentist agreed that treatment was definitely not necessary

(Ratings 1-5 as a decision to restore, rating 6 as a decision to leave unrestored)

Radiograph	Tooth	Visual appearance*	Microscopic appearance*
1	6m	2	1
1	6d	3	1
1	7m	2	0
2	5m	3	3
2	6d	0	0
2	4d	3	3
2	4d	3	0
2	5m	0	0
3	5m	2	2
3	5d	3	2
3	4d	2	2
4	7m	0	0
5	5m	3	2
5	5d	0	0
5	5m	2	3
5	5d	3	3
5	6d	5	2
5	4d	9	9
5	6d	1	1
5	7m	0	0
6	5m	5	9
6	5d	2	9
6	7m	0	0
6	6d	1	2
7	5m	2	0
7	5d	3	2
7	5m	5	5
7	6d	0	0
8	4d	6	9

Table 7.4(iii) continued

Radiograph	Tooth	Visual appearance*	Microscopic appearance*
9	5d	2	0
9	6m	0	0
9	7m	2	1
9	6d	0	0
10	4d	0	0
10	5m	4	9
10	6d	3	3
10	5m	2	2
10	6d	3	3
10	7m	0	0
10	6d	0	0
11	5m	0	0
11	6m	2	3
11	7m	0	0
11	6d	2	1
11	5m	0	0
11	5d	0	0
11	6m	0	0
11	7m	0	0
12	4d	0	0
12	5d	0	0
12	6m	0	0
12	7m	3	9
13	5m	4	4
13	5d	2	0
13	6m	1	0
13	5d	2	3
13	6d	0	3
14	7m	0	0
14	5d	3	3
14	6m	0	0
14	7m	0	0
14	4d	3	3
15	5d	0	0
15	6m	3	3
15	6d	0	0

* See Tables 5.1 and 5.2 for details of criteria for assigned scores.

than half the subjects, when the loosest criteria for treatment was utilised as the "cut-off" level at which practitioners would, in actuality restore a tooth. Of these 65 surfaces 14 (22%) had cavitation, and 15 (23%) had lesions extending into dentine.

It can be seen from these Tables (7.4(i-iii)) that the number of dentist pairings agreeing on treatment decisions decreased as the criteria for a positive decision were lowered. There were more surfaces agreed upon by the majority of dentists when the strictest criterion only was used as a decision to restore.

7.4 DISCUSSION

The responses to the question concerning the appropriate point at which dentists felt a tooth **should** be restored is a potential source of variation amongst them. The dentists in this study did not unanimously agree about the stage of the carious process they were attempting to detect on the radiographs, prior to a decision to restore a tooth. This difference of opinion as to the most appropriate stage of lesion development at which a filling ought to be placed should lead to systematic variations between dentists. The Kappa scores generated when observers 6, 9, and 11 were compared to others, who held different treatment criteria, showed that differences in treatment criteria did not necessarily lead to very low levels of agreement. Such disagreement between dentists would be expected if it were only stated lesion depth criteria which caused variation between dentists treatment decisions. If lesion depth criteria **were** the source of variation, then pairs of dentists with similar views as to the correct point of intervention would consistently achieve higher Kappa scores, than those who held different opinions. This was not the case.

The fact that the Kappa scores of pairings between dentists with unlike lesion depth treatment criteria were not excessively low showed that their beliefs as to how lesions are represented on a radiograph, differ. Alternatively the variation may be because dentists value errors differently (see Chapter 6). If the above two explanations are incorrect, then agreements between dentists with unlike treatment criteria are "mistakes", - i.e. the practitioners are unable to correctly identify a radiograph which shows a lesion of the depth which they consider **should** be restored.

Dentists agree more readily on positive treatment decisions than on negative. The overall Kappa score of 0.502, for decisions to definitely restore teeth indicates they have moderate agreement amongst themselves concerning tooth surfaces which definitely require restoration. However, there is a poor level of agreement about definitely negative treatment decisions. The high probability of a second observer agreeing with a first for decisions to leave a tooth unrestored is merely a reflection of the high numbers of decisions made not to restore teeth. The Kappa score of 0.196 indicates that most of the agreement about definite negative decisions is due to chance, and dentists are not consistent about decisions to definitely leave a tooth unrestored. Therefore, although dentists judge positive decisions in a similar manner, dentists seem to differ in their views as to the radiographic criteria which dictate that a tooth should not be restored. The dentists were therefore differing widely in their setting of criteria for no-treatment whilst demonstrating similar views as to the radiographic appearance which indicated a need for treatment. However, the fact that no dentist pairing achieved substantial agreement (K>0.8) whichever rating scale cut-off level was used, was disappointing.

It is clear from the results that the sources of variation among dentists are systematic as one shifts the diagnostic criteria. The fact that only one of the 16 surfaces for which less than half the dentists decisions agreed in Table 7.4(i), appears in Table 7.4(ii), implies that it must have been rated, by the dentists not deciding to definitely restore it, as rating 2 or 3. Therefore when the dentists disagreed, their views to which category of the rating scale should be used varied only slightly i.e. they were disagreeing with "near-misses", rather than being in total disagreement. This suggests that dentists shift the certainty of their decisions at similar rates, but do not begin to consider restoration at the same point of radiographic evidence.

The increasing numbers of tooth surfaces causing disagreement, as the cut-off criteria are slackened suggests that it is the decision not to place a surface in category 6 (definitely leave unrestored) which causes greater numbers of differing views between pairs of dentists. Dentists seem to be clear as to when to decide to fill a tooth, and less clear as to when a tooth should be left unrestored. This finding is of great interest when viewed in the light of the findings of Chapter 5, which showed that dentists frequently made incorrect positive treatment decisions.

CHAPTER 8. MEASUREMENT OF UTILITIES OF DENTAL HEALTH STATES

8.1 INTRODUCTION

Before using decision trees to structure and analyse the clinical problem of a tooth's need for restorative treatment, an important step is to identify the value of all health outcomes which may arise from the decision whether or not to perform the intervention. The medical literature[241-244] has given some considerable attention to the preferences and values which should be incorporated into clinical decisions, but hitherto, population utilities for dental health states have not been assessed. Some studies have concentrated on **decision-makers'** views, and their preferences for various outcomes[79,152] but none have investigated community-based preferences for various states of tooth surfaces. Considering that, in the UK, patients **choose** whether or not to present to the dental profession for care, and therefore dental treatment is of an elective nature, it is surprising that there is little research investigating patients preferred outcomes of treatment/no treatment.

The dental health outcomes which patients truly desire have largely been ignored, probably because the assumption is usually made that the professional view of health is the 'right' view. Therefore, a filled tooth is considered to be superior to a decayed, filled tooth. However the differential value of diseased and untreated, and diseased and treated teeth, have not been assessed. It is a commonly made assumption that the views of the public are equivalent to those of the dental profession and health care planners, but this is not necessarily true. Even if it were, the value placed on different health states by dental professionals have not been explicitly measured.

If the value, or utility, of each dental health state is known, it becomes possible to rationalise treatment decision making. The availability of such utilities would also enable progress to be made in the future, towards assessments of the quality of teeth, in order that a system of Quality Adjusted Tooth Years (QATYs) be introduced into dental service planning, just as the QALY (Quality Adjusted Life Year) is utilised in medical service planning[245,246].

The issue of service planning is crucial to the philosophy underlying this thesis. Decisions about how much money should be spent on competing dental health programmes demands that the planner synthesises the biological, clinical, demographic, economic, sociologic and political pressures relevant to the choice of programmes and allocation of resources. Because the resources available for dental health care are always limited, the decision as to how they should be allocated - to preventive or restorative services - involves placing judgements as to the value of the health states achieved or avoided by these programmes. That is, the worth of the outcomes of a programme are inherent in the decision to implement such a programme. Thus, the current provision of resources for restorative, rather than preventive dental services implies an assumption that the F (filled) component of the DMF (decayed, missing, filled) index is of equal or more value than sound teeth component. Rationally, the higher the utility (or worth) of a dental health outcome achieved by a service or programme, (or the lower the worth of the dental health state it prevents), the more deserving it becomes for allocation of resources. Undertaking rational dental management decisions, therefore, requires that the utilities of various dental health states are known. The data presented in this Chapter attempt to quantify the benefit derived from dental services, as perceived by the population receiving them.

Given that the UK dental health service is currently largely funded by the public, an argument can be made that the evaluation of dental health states should not be made implicitly by government officials and health professionals, but by incorporating the explicitly measured dental health utilities of the general public, into the decision-making process.

8.2 AIMS

This chapter describes pilot studies which utilise two different methods of measuring the utilities of dental health states. The aim of the studies described were to examine the practicality of assessing dental health state utilities from uninformed members of the general population.

The objectives of the two studies were therefore:

(1) To determine whether visual analogue and standard gamble questionnaires designed to measure dental health state utilities are understood and accepted by the general public.

(2) To determine whether the results obtained by each method of assessing dental health state utilities were empirically acceptable.

(3) To examine the affect of dental attendance on dental health state utilities as this behaviour can be taken as a measure of the value people place on their dentitions.

(4) To examine the comparability of the two methods of assessing dental health state utilities.

(5) To provide dental health state utility values in order that decision analysis may be applied to dentists' restorative treatment strategies (Chapter 9).

8.3 METHODOLOGY

The states of dental health chosen were those whose occurrence, clinical course, and response to restorative therapy were well documented.

Each individual's perception of the value of a health state may depend on whether or not he/she has had experience of that particular disease or pathology. The dental health state utilities measured in these studies might therefore be related to the person's experience of both decay, and dental services. However the studies reported here are the first to try to quantify dental health state utilities. They are therefore pilot studies, which explore methodologies which have potential for the measurement of tooth state utilities, rather than being studies which seek to definitively measure the utility of dental health states.

8.4 UTILITIES OF DENTAL HEALTH STATES - STUDY 1 USING A VISUAL ANALOGUE SCALE TO ASSESS THE UTILITY OF DENTAL HEALTH STATES

8.4.1 Introduction

Several methods exist for eliciting value judgements. One of the simplest of these is the visual analogue scale. A visual analogue scale, for the purpose of measuring the value of health states requires two extremes of health to be written at either end of a ten centimetre line. The respondent is then asked to determine the value of an intermediate health state by marking the point on the line which he/she feels best represents the relative value of that state, in comparison to the two extremes.

Previous evidence has shown that people appear to value anterior and posterior teeth differently[39,40]. It has also been noted on several occasions that regular and irregular attenders view their teeth in different ways. Therefore, this investigation studies the effect of these variables on dental health state utilities as measured using a visual analogue scale. If the results are acceptable in the light of prior knowledge[39,40] then it would seem that such a technique may have some internal validity and predictive value for the measurement of dental health state utilities. Likewise, the visual analogue scale would have to be rejected as a means of quantifying the value of dental health, if it produced empirically unacceptable results.

8.4.2 Aims

The first study described in this Chapter was carried out (a) in an attempt to investigate whether visual analogue scales might be of use for measuring dental health state utilities in the population; (b) in order to assess whether the values for various dental health states, as measured using a visual analogue scale were empirically acceptable and, (c) to derive utility values or dental health states which could be used in a decision analysis (Chapter 9).

8.4.3 Selection of Dental Health States

Four states of a tooth were selected for this first study

(1) a completely healthy restored tooth

(2) a completely healthy, restored tooth which will require re-restoration within 5 years

(3) a decayed non-painful tooth

(4) a decayed painful tooth.

A questionnaire was designed, consisting of 10 questions (Appendix VII). The age, sex and occupation of the individual was ascertained, as was their regularity of dental attendance. Respondents were then asked to mark the value they placed on each dental health state on a visual analogue scale for both posterior and anterior teeth. A pilot study involving 14 individuals resulted in a slight rewording of the questionnaire.

8.4.4 The Sample

It was impractical to draw a random sample of the UK population for participation in this study. Therefore individuals in one geographical location (Glasgow) were involved. They were interviewed by a team of researchers who merely stopped pedestrians in Glasgow City Centre and asked them to participate.

Interviewees were eligible for participation if they were over 18 years of age, were prepared to be interviewed and were able to communicate in English. No target was set for the number of interviews, the final sample being determined by the number of interviews which could be completed by six interviewers in one day.

8.4.5 Results

The sample consisted of 215 individuals over 18 years of age, their age distribution being detailed in Figure 8.1. It can be seen that this distribution is positively skewed, in that the younger age groups are proportionally over-represented.

Of the sample 49% (105) were male and 51% (110) were female, and 69% (149) reported that they attended the dentist regularly (i.e. without pain or problems as a prompt to service utilisation). However, 29% (62) reported that they only visited a dentist when they considered that they had an oral health problem (Figure 8.2).

The mean utility placed on each dental health state examined for anterior and posterior teeth is shown in Table 8.1. It can be seen here that, for both anterior and posterior teeth, the utilities increase with increasing 'health' of the tooth, and Table 8.2 details this

Figure 8.1 Age distribution of the sample in Study 1

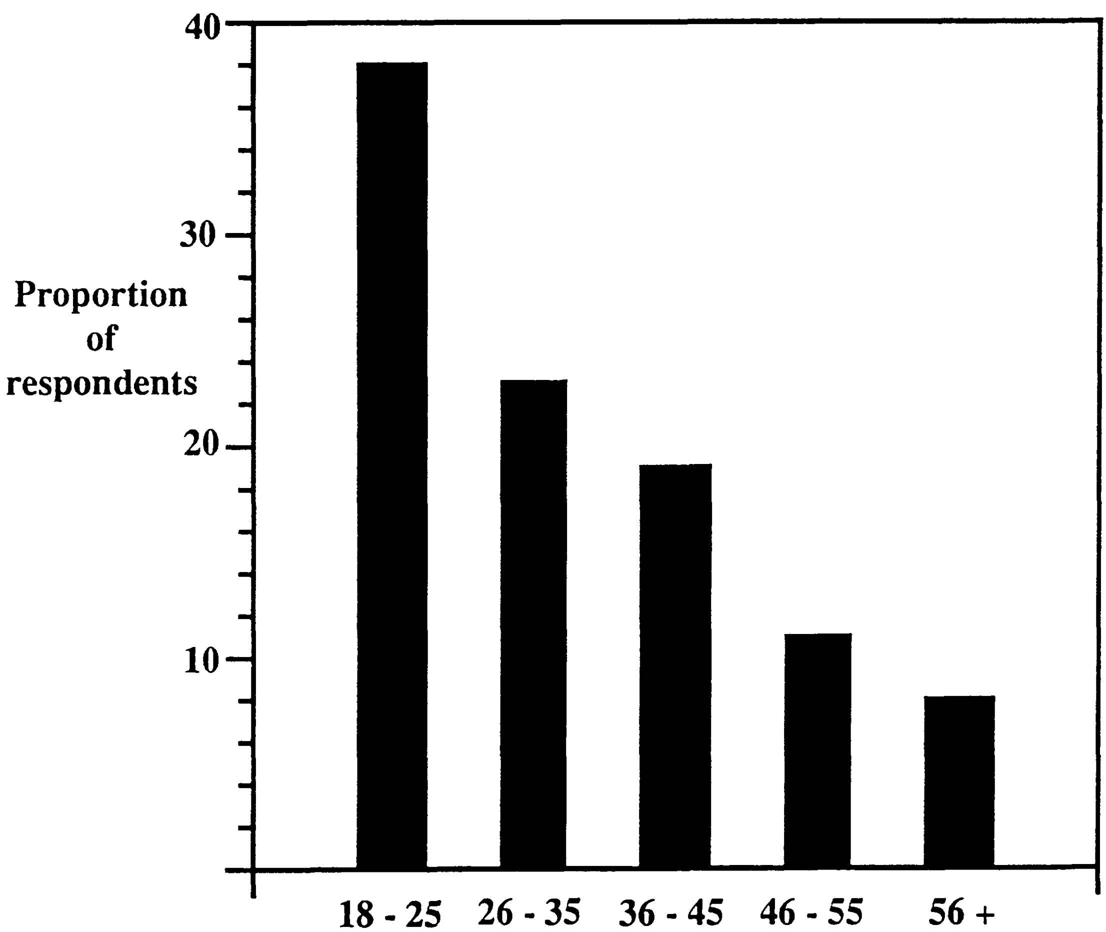


Figure 8.2 Attendance pattern of respondents in Study 1

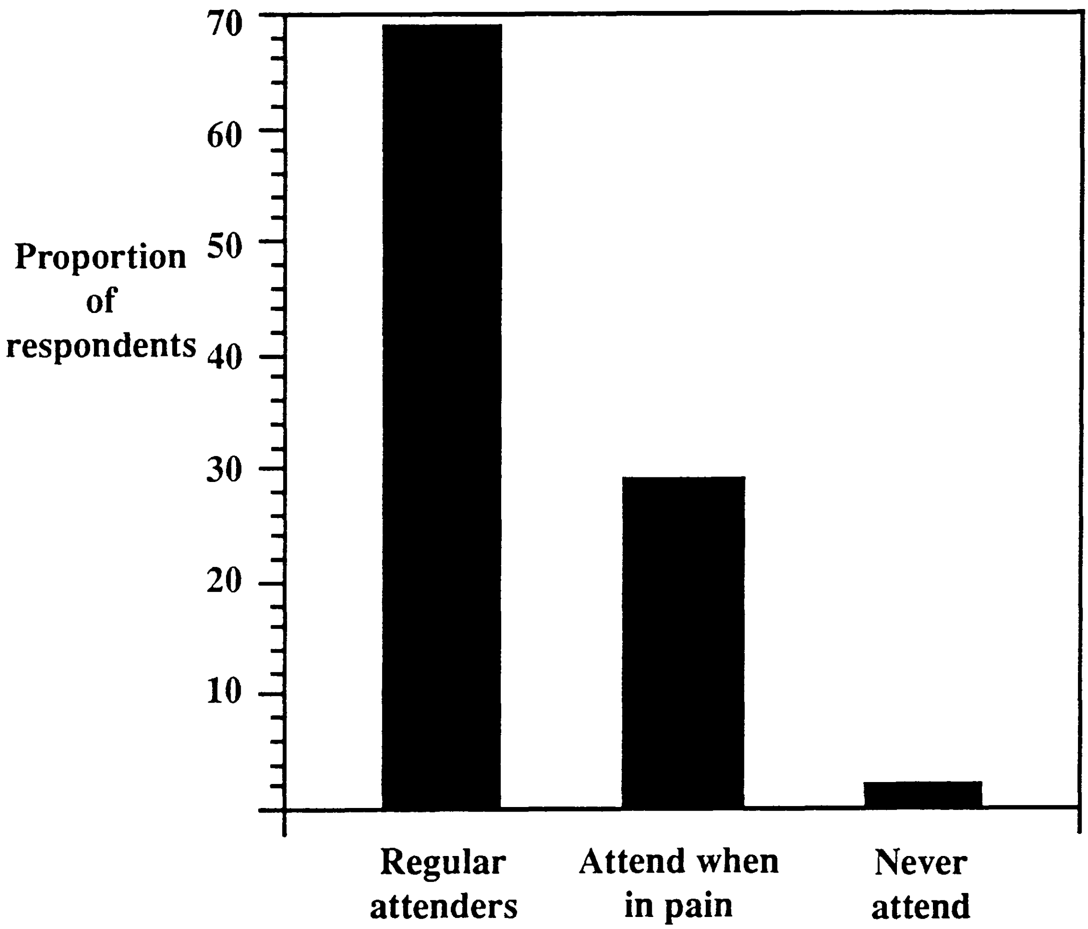


Table 8.1 Average utility (and range) for four dental health states for anterior and posterior teeth using a 10 cm visual analogue scale (Study 1)

Dental Health State Assessed	Average Utility (Range)			
	Anterior teeth		Posterior teeth	
Decayed, painful	0.49	(0.21-.052)	0.31	(0.11-0.45)
Decayed, non-painful	0.60	(0.33-0.65)	0.42	(0.36-0.46)
Filled, refilled	0.79	(0.61-0.82)	0.56	(0.52-0.61)
Healthy, filled	0.89	(0.65-0.94)	0.62	(0.55-0.67)

Table 8.2 Average utility placed on each dental health state using 10 cm visual analogue scale, by age group (Study 1)

State of Health	Average Utility in Age Group					
Age	18-25	26-35	36-45	46-55	56-65	65+
Healthy, filled anterior tooth	0.88	0.87	0.89	0.88	0.92	1.00
Healthy, filled posterior tooth	0.55	0.57	0.65	0.71	0.75	0.93
Filled, refilled anterior tooth	0.74	0.79	0.80	0.79	0.91	0.67
Filled, refilled posterior tooth	0.46	0.54	0.58	0.62	0.74	0.80
Decayed, non-painful anterior tooth	0.54	0.65	0.56	0.54	0.62	0.77
Decayed, non-painful posterior tooth	0.34	0.47	0.45	0.41	0.50	0.83
Decayed, painful anterior tooth	0.43	0.51	0.59	0.41	0.48	0.33
Decayed, painful posterior tooth	0.23	0.35	0.40	0.27	0.32	0.37

information by age group. The Table demonstrates that the utilities for filled teeth were greater in the older age groups, whilst older individuals placed lower value on decayed teeth than the middle aged group.

Non-regular dental attenders held consistently lower average values for each of the states of dental health examined. (Table 8.3).

8.4.6 Achievement of aims - Study 1

The first aim of the study was to determine whether or not the visual analogue scale was understood and accepted by the public. Only 27 people who were asked to participate refused to be the interviewed. This high acceptance of the interview (89%) shows this type of questionnaire to be generally acceptable to the public.

The second aim was to examine the empirical acceptability of the visual analogue scale as a measure of the value of various dental health states. It can be seen from Table 8.1 that the ranges of response were narrow, and that the values for anterior teeth, for each state of health, were higher than for posterior teeth. This result is in accordance with what might be expected from previous evidence[39,40]. The Table (8.1) also shows that the sequence of values .49, .60, .79, and .89 for decayed painful, decayed non-painful, temporarily filled and permanently filled, anterior teeth, respectively, are as might be expected, i.e. disease which would lead to unpleasant sequelae are assigned a lower utility by the public.

The utilities for posterior teeth showed a similar increase in values, with improvements in utility being in agreement with normative views of dental health, i.e. the states of health which a dentist would perceive as "better", are valued more highly by the responding population.

The differential values placed on teeth by regular and irregular dental attenders gives the most convincing evidence that this method of measuring dental health state utilities does indeed reflect the value systems which dictate individual's behaviour with regard to their dental health. Thus, the methodology appears to have some predictive value. Therefore the visual analogue scale seems to truly measure beliefs about the importance of dental health, and therefore might, with development, act as a useful tool for the measurement of tooth-state utilities.

The final aim of the study was achieved in that empirically acceptable utility values for dental health states, which could be used in a decision analysis were achieved. It is not, in fact crucial to the decision analysis to have achieved definitive measurements of dental health state utilities, as sensitivity analyses can examine the effect of variations in these

Table 8.3 Average utility placed on each dental health state using 10 cm visual analogue scale, for regular and irregular attenders (Study 1)

Dental Health State	Average utility for regular attenders	Average utility for non-regular attenders
Healthy, filled anterior tooth	0.90	0.85
Healthy, filled posterior tooth	0.64	0.53
Filled, refilled anterior tooth	0.79	0.75
Filled, refilled posterior tooth	0.57	0.48
Decayed, non-painful anterior tooth	0.64	0.50
Decayed, non-painful posterior tooth	0.44	0.35
Decayed, painful anterior tooth	0.52	0.41
Decayed, painful posterior tooth	0.31	0.27

values. It does however give the decision analysis greater meaning when utilities derived from population studies are available for interpretation of the decision analysis.

8.4.7 Comments on Study 1

In conclusion, although this study had major flaws due to the nature of the sample, the concurrence of the results with all previous evidence[39,40,217-219] suggests that the utilities derived from the study may be useful for quantifying the value of the general public assign to various dental health states.

8.5 UTILITIES OF DENTAL HEALTH STATES - STUDY 2 USING A STANDARD GAMBLE QUESTIONNAIRE TO ASSESS THE UTILITY OF DENTAL HEALTH STATES

8.5.1 Aims

The second study described in this Chapter was carried out (a) as an attempt to validate the study described previously and (b), as an example of a use in dentistry, of a technique previously employed in medical fields to assess health state utilities.

8.5.2 Background

The need to develop a valid index to distinguish between various dental health outcomes reflects the growing appreciation amongst public health dentists that there is a need to incorporate patient and public preferences into treatment, and management, decision-making. To achieve such a goal, the dentist or dental manager must provide the patient with data about the possible outcomes of available treatments, and the public must be able to comprehend and use these data. The public must be given alternative health outcomes which might arise from service planning and treatment decisions, in order to detect truly their preferences, in terms of health care policies.

Therefore the method employed in this second "utility" study was to present each respondent with a large series of dual alternatives. For each response, the subject was given a choice between a certainty of having an intermediate health outcome (e.g. a decayed tooth), or a given probability of a completely healthy tooth with a complementary probability of the worst possible dental health outcome (extraction). (See questionnaire in Appendix VIII).

8.5.3 Method

As in study 1, four states of a tooth were selected for assessment of their utilities

(1) a completely healthy restored tooth

(2) a completely healthy restored tooth which will require re-restoration within 5 years

(3) a decayed non-painful tooth

(4) a decayed and painful tooth

A questionnaire was designed (Appendix VIII) consisting of four standard gambles, regarding the tooth states listed above. The respondents undertook the gambles simultaneously for both anterior and posterior teeth, in order that the relative values could be examined as in Study 1.

Each respondent indicated the probability at which they would live in the intermediate outcome (see above), in preference to risking full dental health or dental extraction (see Appendix VIII).

As in the previous study, it was not deemed practical to administer the lengthy questionnaire (Appendix VIII) to a random sample of the population. Therefore in May of 1989 a separate street sample of 110 individuals in Glasgow participated. This was deemed to be an acceptable sampling method, due to the problem-definition nature of the study.

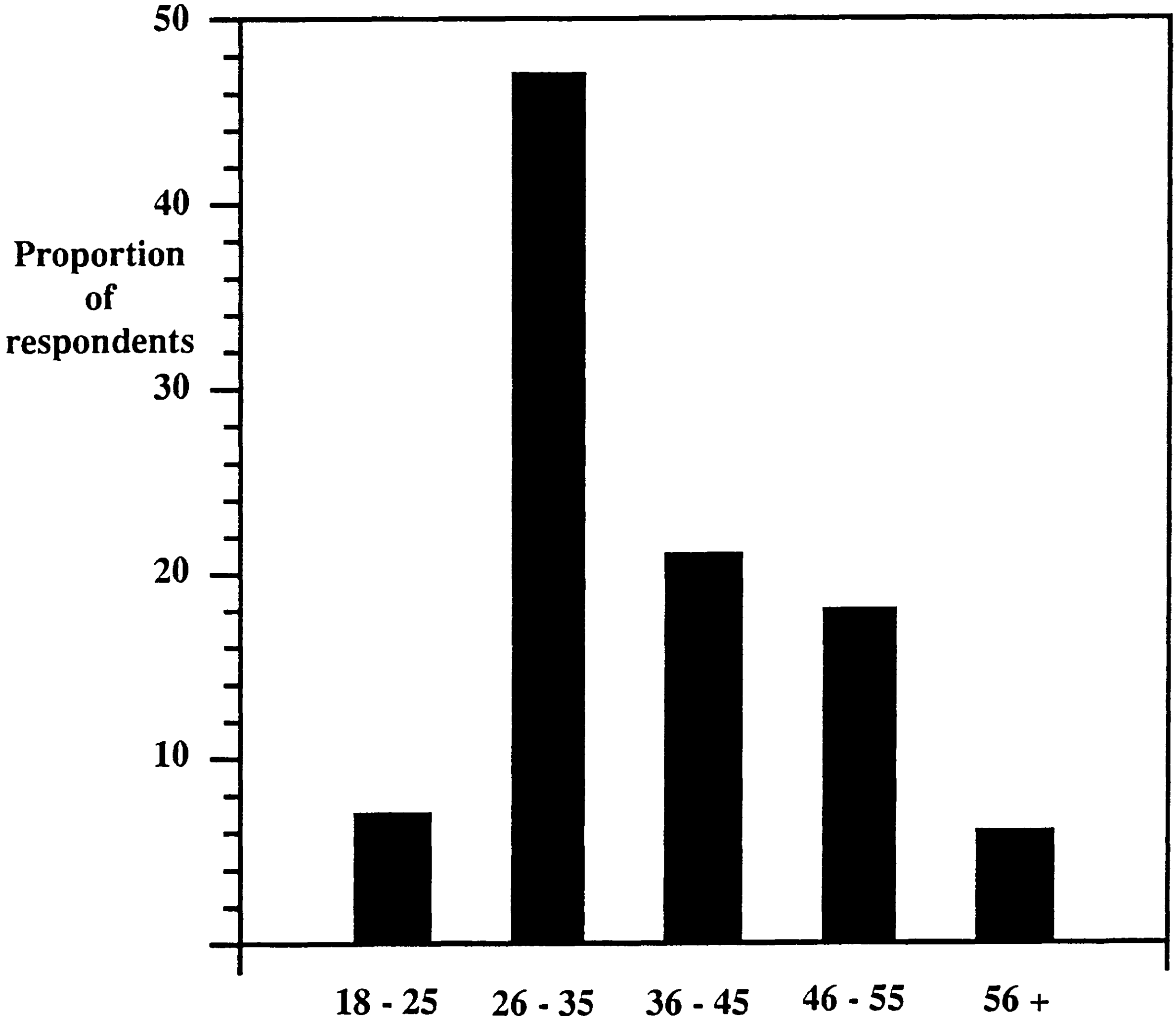
The inclusion criteria were the same as in Study 1.

8.5.4 Results

The sample consisted of 110 individuals with an age distribution as demonstrated in Figure 8.3. Here it can be seen that the modal group was the 26-35 year olds, with all other age groups comprising less than 20% of the sample. Of these 52% (57) were male and 48% (53) were female. Of the 110 participants 63% (69) reported that they had visited the dentist regularly whilst 28% (31) were irregular attenders.

The mean utility for each dental health state examined, for anterior and posterior teeth is detailed in Table 8.4. It can be seen from this Table that, as in Table 8.1 (Study 1) the

Figure 8.3 Age distribution of respondents in Study 2



tooth utilities increased with increasing 'health' of the tooth. When choosing between a sound tooth for life (with varying probability of extraction) and a certainty of a decayed tooth, the average utility for a decayed painful posterior tooth was .46 (Table 8.4). When choosing between a sound tooth for life (with varying probabilities of the tooth extraction) and a certainty of a filling which must be replaced within five years, the average utility for a short-term-filled tooth was 0.69 (i.e. a posterior filled tooth with the filling lasting for 5 years is worth 69% of a sound tooth).

If a posterior restoration was deemed to have an indefinite lifespan, a filled posterior tooth was considered to have 76% of the value of a sound tooth.

The differences between the mean utility scores among regular and irregular attenders are given in Table 8.5. Here it can be seen that regular attenders consistently gave each tooth state a higher mean utility value.

8.6 COMPARISON OF STUDIES 1 AND 2

Data in Table 8.6 indicate the differences between the samples and results of the two studies.

It shows that the age distribution of the Study 1 sample differed from that in Study 2. However, by age standardizing the results of Study 1 to Study 2 (Table 8.7) it can be shown that the two methodologies are comparable.

In these studies two measuring instruments were used. The visual analogue scale is included because of its simple structure, its potential low cost, ease of administration and its widespread use in the field of attitude measurement. The standard gamble questionnaire was used because of its wide acceptance as a method of measuring health state utilities.

Comparison of the studies give rise to an important question. Is it preferable to accept the results of a somewhat simplistic methodology which has never previously been tested, (Study 1) despite the fact that it shows some interesting and empirically acceptable results, or is it preferable to accept the results of a study which uses a methodology which has been tried, tested and validated in other contexts, but never in the dental field? (Study 2).

8.7 Comments

To quantify the feasibility of these instruments three measures could be used: firstly their acceptability to the public; secondly their ease of use by the interviewer, and thirdly, their cost. Both techniques appeared to be acceptable in that very few subjects were unwilling

Table 8.4 Average utility (and range) placed on each dental health states as assessed by the standard gamble questionnaire (Study 2)

Dental Health State Assessed	Average Utility			
	Anterior teeth		Posterior teeth	
Decayed, painful	0.50	(0.3-0.8)	0.46	(0.2-0.90)
Decayed, non-painful	0.56	(0.3-0.95)	0.51	(0.2-0.95)
Filled, refilled	0.72	(0.4-1.0)	0.69	(0.3-1.0)
Healthy, filled	0.76	(0.4-1.0)	0.72	(0.5-1.0)

Table 8.5 Average utility placed on each dental health state using standard gamble questionnaire for regular and irregular attenders (Study 2)

Dental Health	Average Utility for regular attenders	Average Utility for irregular attenders
Decayed, painful anterior tooth	0.60	0.50
Decayed, painful posterior tooth	0.51	0.43
Decayed, non-painful anterior tooth	0.69	0.40
Decayed, non-painful posterior tooth	0.67	0.45
Filled, refilled anterior tooth	0.74	0.56
Filled, refilled posterior tooth	0.74	0.51
Healthy, filled anterior tooth	0.81	0.72
Healthy, filled posterior tooth	0.79	0.69

Table 8.6 Comparison of sample characteristics in utility Studies 1 and 2

		Study 1 (Visual Analogue)	Study 2 (Standard Gamble)
N		215	110
Sex	Male	49%	51%
	Female	50%	49%
Age Distribution			
	18-25	38%	7%
	26-35	23%	47%
	36-45	19%	21%
	46-55	11%	18%
	56+	8%	6%
Attendance			
	Regular	69%	64%
	Irregular	29%	33%
	Never	2%	1%
Utilities			
	Completely healthy (anterior)	1.0	1.0 (assumed)
	Completely healthy (posterior)	1.0	1.0 (assumed)
	Filled, healthy (anterior)	0.89	0.76
	Filled, healthy (posterior)	0.62	0.72
	Filled, refilled (anterior)	0.79	0.72
	Filled, refilled (posterior)	0.56	0.69
	Decayed, non-painful (anterior)	0.60	0.56
	Decayed, non-painful (posterior)	0.42	0.51
	Decayed, painful (anterior)	0.49	0.50
	Decayed, painful (posterior)	0.31	0.46

Table 8.7 Average utility for four dental health states, for Studies 1 and 2 with results age-adjusted to population in Study 2. (Omitting 65+ age group, N=4)

	Average Utility (age-adjusted to Study 2)	
	Study 1	Study 2
Decayed, painful anterior tooth	0.50	0.49
Decayed, painful posterior tooth	0.34	0.31
Decayed, non-painful anterior tooth	0.62	0.60
Decayed, non-painful posterior tooth	0.67	0.45
Filled, refilled anterior tooth	0.80	0.79
Filled, refilled posterior tooth	0.57	0.56
Healthy, filled anterior tooth	0.88	0.89
Healthy, filled posterior tooth	0.62	0.62

to go through with the interview. If an instrument was unacceptable, one would expect many interviews to be broken off (this only happened in two cases with the standard gamble technique, and in no case with the visual analogue technique). Surprisingly, the subjects seemed to find the visual analogue technique slightly more difficult to understand than the standard gamble. The interviewers in both studies, after a brief introductory discussion lasting approximately one hour, found both interview schedules manageable, although the additional length and more complex structure of the standard gamble questionnaire required greater explanation by the investigator. The costs of the standard gamble interview schedule were greater due to increase printing requirements.

With regard to the validity of these measurements, it is widely accepted that the von Neumann Morgenstern standard gamble[202] can be taken as the standard technique for use in medical fields. However, since this is the first use of this technique in a dental context, it would be unwise to accept the result unquestioningly.

Therefore, at present, there is no explicit quantitative estimation of the effects of dental health services and dental interventions on dental health outcomes, and making quantitative analyses of dental health outcomes is a difficult problem. However, two studies are described here which attempt to assess population utilities for filled, decayed, decayed and painful, and decayed and pain-free teeth.

The groups participating in the studies can clearly not be assumed to be representative of the British population as a whole. However, these studies are the first to try to place a quantitative value on dental health outcomes, and in the absence of more extensive investigations, the values derived from this information will be used in the decision analyses presented in the next Chapter.

These studies contribute to a search for a good instrument by which to measure social preferences for dental health states. The standard gamble technique is somewhat complex and costly to administer, each interview taking approximately 20 minutes, and it is probably only useful for application by professional interviewers with educated subjects. It remains for a dentally-oriented standard gamble technique to be widely tested on a random sample of the general public.

The visual analogue technique is less time-consuming, and therefore would be less costly. However, the respondents sometimes found this procedure difficult to understand.

In summary, these studies suggest that the standard gamble technique has potential as an instrument for measuring social preferences for dental health states, but great care must be taken to keep the procedure simple if it is to be used successfully on the general public.

As this is a new and difficult field of research, the results must be considered tentative, and call for further population studies and an expansion of the number of health states examined. Nonetheless, the findings are presented as a contribution towards the development of what, in the author's opinion is a badly needed instrument.

CHAPTER 9. DECISION ANALYSIS

9.1 INTRODUCTION

Chapter 5 investigated the relationship of dentists' decisions to certain "gold standards" - i.e. the depth of the carious lesion as seen under a microscope. The values of dental practitioners, which are implicit in treatment decisions, and those which the dentist explicitly states, have also been examined. However, from a rigorous scientific standpoint, the "correctness" of a treatment decision rests upon the probability of, and the value of, the various outcomes which may arise as a result of the decision. In other words, decisions about the most appropriate treatment for a tooth surface should depend upon the degree of likelihood of progression towards outcomes which may be favourable, or unfavourable, in patients' eyes. It is therefore invalid to base measurement of the value of treatment decisions on the conventional wisdom relating to depth of lesion, or on simple evaluation using sensitivity and specificity. For example, if a patient values an extraction more highly than a restoration then it is clearly unnecessary to hold strict lesion-depth criteria as a basis for the decision to restore a tooth. While it is probable that practitioners do implicitly base clinical treatment decisions on their estimates of patients' preferences, and on estimation of the likelihood of lesion progression, the role of these factors have not hitherto been explored in relation to treatment decisions.

This Chapter seeks to clarify what level of intervention is appropriate, given the utilities a population of potential patients assign to certain outcomes. In order to understand the concept of this analysis it is necessary to leave aside traditional assumptions about 'correct' treatments. Decision trees can be constructed, firstly to clarify the probabilities which are relevant to a decision to restore a tooth, and secondly to examine the effects of altering various parameters; namely, the predictive values of dentists' decisions, population utilities, increasing activity and incidence of disease, and rates of restoration failure.

In order to construct an appropriate decision tree concerning restorative dentistry, it is necessary to list, and evaluate, ALL the possible outcomes of a dentist's action or inaction, when he is faced with a treatment decision. Enumerating the outcomes is a simple procedure, whilst evaluating them is more difficult. In Chapter 8, an attempt was made to derive population utilities for the outcomes of dentists' decisions. There is one outcome, however, which was not included as it was considered to be both unwise and unethical to attempt to assess the value of an **unnecessarily** filled tooth. Such a question might have raised uncertainty in the public's mind as to the effectiveness or integrity of the dental profession. In the decision analysis which is to be described it was therefore necessary to make assumptions about these values.

The other information which was required for the construction of a decision tree was the probability of occurrence of the various outcomes. Probabilities of lesion progression, and probabilities for restoration failure are derived from the literature reviewed in Chapter 2, Section 2.2.2 and 2.2.4. The probabilities of true positive, true negative, false positive and false negative treatment decisions when dentists make decisions about carious lesions of various depths, are derived from the results described in Chapter 5.

In order to assess whether treatment brings about outcomes which the population view as favourable, under any given set of circumstances, all of the values, probabilities and utilities can be subjected to sensitivity analysis. This allows one to vary them in a systematic fashion and to measure the effect of such changes on the expected utility of a decision.

In essence the decision tree takes the place of what would otherwise be a somewhat unethical clinical trial. For example, one can assess the expected utility of a decision to treat enamel-only lesions, or a decision to leave dentine lesions untreated. One can also examine the utility of treating lesions even when a dentist is uncertain that treatment is required.

9.2 AIMS

The aims of this chapter are therefore:

1) To examine the expected value of positive and negative treatment decisions, and to investigate whether alteration of treatment criteria affects the expected utility of these decisions.

2) To ascertain how the likelihood of pain would affect the expected utility of treatment decisions.

3) To ascertain the effect of restoration failure rates on the expected utility of positive and negative treatment decisions.

4) To ascertain the effect of the population utility of unnecessary treatment on the expected utilities derived from the decisions dentists make.

5) To ascertain the effect of varying the likelihoods of lesion progression on treatment decision-making.

6) To ascertain the effect of increased caries incidence on treatment decision-making.

9.3 METHOD

Decision trees can be constructed to examine the utility or 'usefulness' of decisions made at any treatment threshold. When a decision is made to treat, or leave a tooth unrestored, there is a given probability (the *a posteriori* probabilities derived from chapter 5) of the surface being sound; of it having a lesion into dentine; of it having a lesion into the inner enamel; or of a lesion into the outer enamel being present. The binary decision to be assessed concerns the decision about whether a tooth surface should be treated or not. In this analysis three trees were analysed. Tree number one (Fig. 9.1) concerned the posterior probabilities brought about when only ranking 1 in the dentists' questionnaire (Appendix III) was classed as a decision to treat. Tree number 2 (Fig. 9.2) concerns analysis of treatment decisions when rankings 1-3 were considered as a decision to treat. Tree number 3 (Fig. 9.3) is derived from the probabilities of caries of various depth which arise when only a treatment ranking of 6 was considered to be a decision to leave a tooth unrestored and all other decisions were deemed to be positive. The posterior probabilities, i.e. the probabilities of a given lesion depth apropos a decision to treat or not treat, are shown in Table 9.1. These data are derived from the results of the rating scale experiment and subsequent sectioning of the teeth which is described in Chapter 5.

If a treatment decision is made about a tooth surface there is then a given probability that the lesion would have progressed had the tooth not been restored. These probabilities are shown in Table 9.2 and are derived from the literature reviewed in Section 2.2.2. Once restored, the filling in a tooth may be permanent or may fail. Thus the end outcomes of a decision to restore a tooth are either a "necessary filling", which may or may not require to be replaced, or an unnecessary filling: which again may be replaced or may last a lifetime. The probabilities of restoration failure were derived from the literature (see Table 2.2, Chapter 2), and the population utilities of the outcomes used in the decision tree were derived from the results of the studies described in Chapter 8.

Following a negative treatment decision, any lesion may or may not progress, a sound surface may develop a lesion, and lesions may or may not give rise to pain. It is also possible that the lesion may regress. The utilities for these possible outcomes are shown in Table 9.3.

In order to assess whether treating lesions using different decision criteria, affects the expected utility of treatment decisions, the maximum expected utility of positive and negative decisions was assessed. This was derived by calculating the cumulative probability of each outcome, multiplying them by the utility of the outcome, and summing these factors for positive and negative decisions. Thus, by adding together all the potential benefits of treatment decisions weighted by their probability of occurrence, it is possible to examine the maximum expected utility of positive and negative decisions.

FIG 9.1
DECISION TREE FROM DECISIONS MADE USING TREATMENT THRESHOLD 1 AS POSITIVE DECISIONS

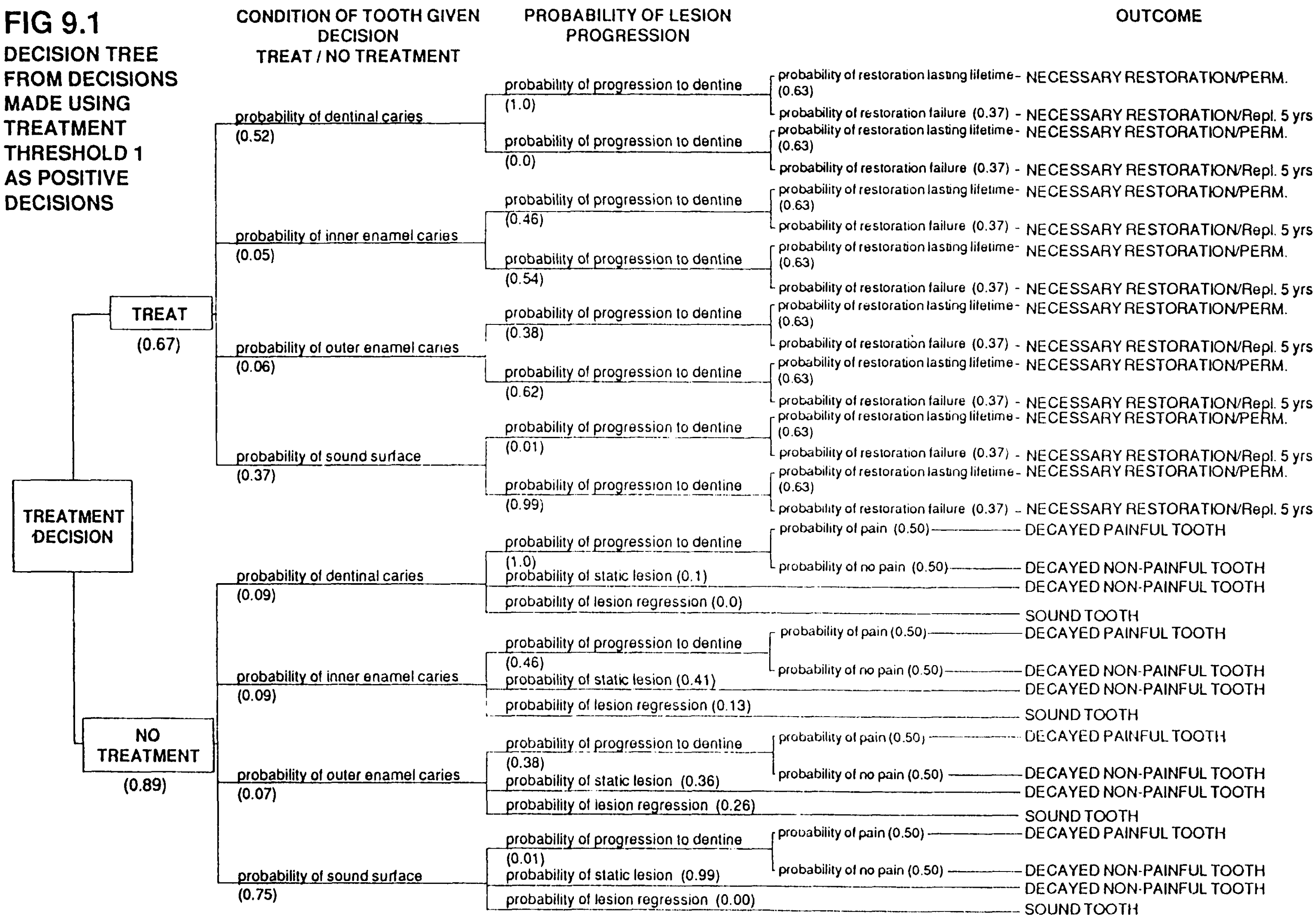


FIG 9.2
DECISION TREE FROM DECISIONS MADE USING TREATMENT THRESHOLD 1,2 & 3 AS POSITIVE DECISIONS

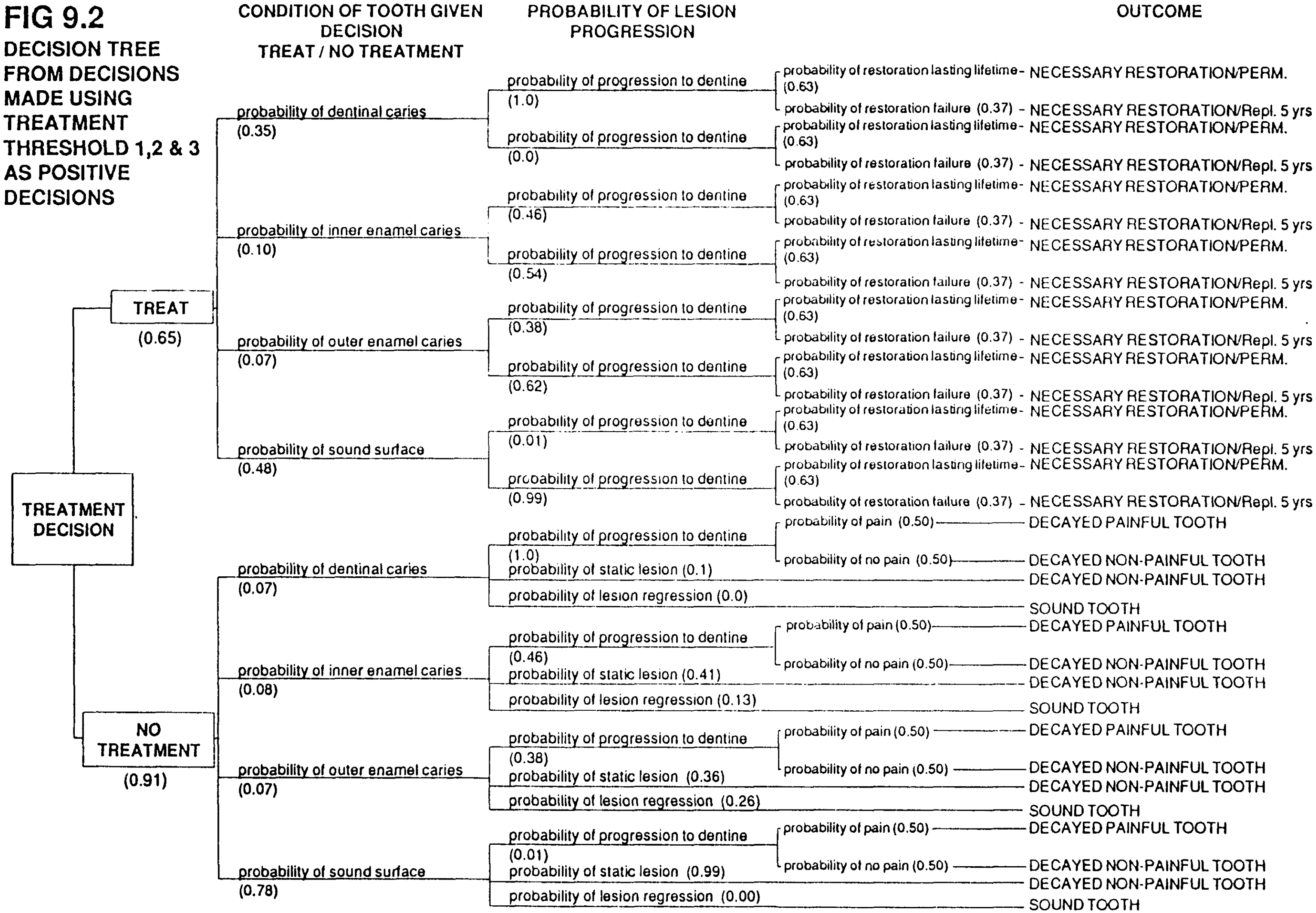


FIG 9.3
DECISION TREE FROM DECISIONS MADE USING TREATMENT THRESHOLD 1-5 AS POSITIVE DECISIONS

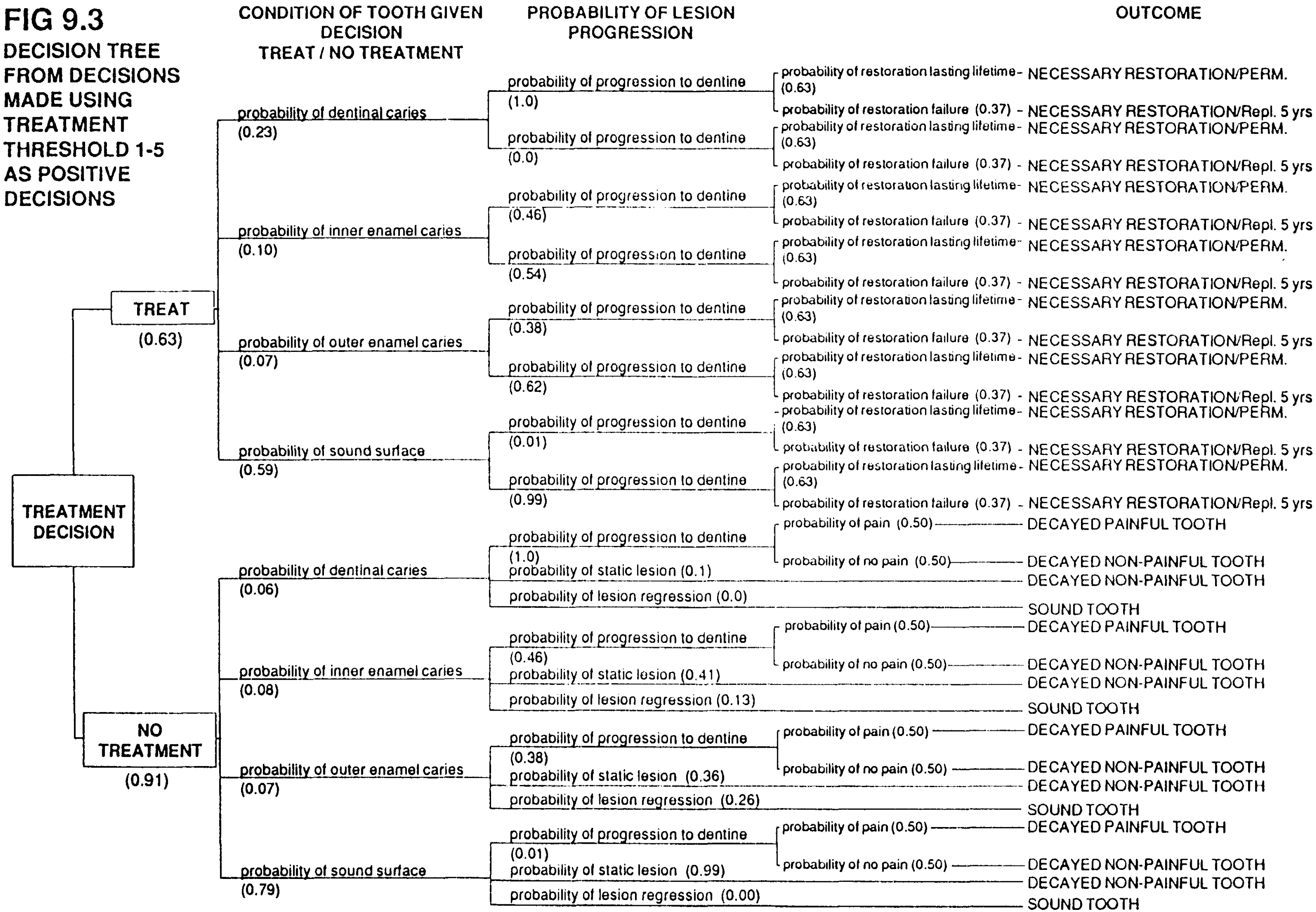


Table 9.1 Probabilities of lesion depths given dentists' decisions, as used in the decision analysis

	Lesion into dentine	Lesion in inner enamel	Lesion in outer enamel	Sound tooth
Dentists score considered as positive decision	Probability	Probability	Probability	Probability
1	.52	.05	.06	.37
1-3	.35	.10	.07	.48
1-5	.23	.10	.07	.60
Dentists score considered as negative decision				
2-6	.09	.09	.07	.75
4-6	.07	.08	.07	.78
6	.06	.08	.07	.79

All values in this table are derived from results in Table 5.3.5 (Chapter 5).

Table 9.2 Probabilities utilized at nodes of decision trees and their sources

	Probability of progress to dentine within 3 years (source)	Probability of regression to sound within 3 years (source)
Lesion in dentine	1.00 (assumed)	0.0 (assumed)
Lesion in inner enamel	.46 (average of extreme values shown in Table 2.4)	.13 (Dummer et al (see Section 2.2.2)
Lesion in outer enamel	.38 (average of extreme values given in Table 2.1)	.26 (Dummer et al (see Section 2.2.2)
Surface sound	0.00 (assumed*)	1.00 (assumed*)

*) The probability of a restoration being permanent is the average of the extreme values noted in Table 2.2 (Chapter 2).

The probability of development of pain in a tooth which has caries into dentine is assumed.

Table 9.3 Outcome utilities utilised in decision trees

Outcome	Utility*
a completely sound tooth	1.00
a necessary** restoration which will last for life	0.76
a necess. ary restoration which lasts for less than 5 years	0.68
an unnecessary** restoration which will last for life	*0.60*
an unnecessary** restoration which lasts for less than 5 years	*0.57*
a decayed painful tooth	0.46
a decayed painless tooth	0.56

* The derivation of these utilities is described in Chapter 8.

** This assumes that a filling is 'necessary' if the tooth has, or will have, within 3 years, a carious lesion within the dentine.

The utilities in italics are assumed, for reasons explained earlier.

Therefore, in order to interpret a decision tree, one requires to follow a set of circumstances down the tree. e.g. given a decision to treat a tooth surface there is a given probability of (i) the tooth being sound, (ii) the tooth surface having a lesion into the outer enamel, (iii) the tooth surface having a lesion into inner enamel, and (iv) the tooth surface having a lesion into dentine. Following each of these branches then leads to the chance of progression of each lesion type, and the final 'chance' which is considered in the tree is whether or not a restoration would be permanent. The probabilities in each route through the tree are multiplied by each other and this cumulative probability used to weight the utility at the end of the branch.

Similarly, for a decision not to treat a surface, there is a given probability (see Table 9.1) that there will be a lesion of a particular depth. For each of these types of lesions, there is a given probability that the lesion will (i) progress, (ii) remain static, or (iii) regress. In the case of (i), there is then a given probability that pain will occur. Thus the utility of each end node (i.e. outcome) is again weighted by the cumulative probability with which it will occur.

The following one-way sensitivity analyses were performed (i) varying the probability of pain; (ii) varying the probability of restoration failure; and (iii) varying the value of an unnecessarily filled tooth; (iv) varying the rates of lesion progression, and (v) varying the incidence of caries in microscopically sound teeth.

9.4 RESULTS

The expected utility of a negative treatment decision was always greater than a decision to treat a tooth, even using the strictest treatment criteria. At treatment threshold one (only a ranking of 1 considered to be a decision to restore) the expected utility of a negative treatment decision was 0.89, and a positive treatment decision was 0.67.

At the other treatment thresholds the differences between the expected utilities of decisions increased: at treatment threshold two (ranking 1-3 considered to be a decision to restore) the expected utility of a negative treatment decision was 0.91 and that of a positive decision 0.65. At the third diagnostic level analysed, the utilities of negative and positive decisions were 0.91 and 0.63 respectively. The decision trees may be seen in Figures 9.1, 9.2 and 9.3. In the following text all figures denoted a) represent findings from the first decision tree (where rating 1 is a decision to restore), figures marked b) derive from the analysis where ratings 1-3 are decisions to restore, and those labelled c) are from the data where ratings 1-5 are all considered positive. The expected utility of positive treatment decisions are therefore maximised by holding high treatment thresholds. i.e. only restoring teeth when absolutely certain that that is what is required.

Shown in Figure 9.4(a-c) are the effects of increased pain probability on the utility of treatment decisions for the three treatment thresholds. Increased likelihood of pain from a decayed tooth minimally decreases the expected utility of a negative treatment decision, and marginally improves the expected utility of a positive treatment decision, but not to the extent where a positive decision brings about a higher expected utility than a negative one. Even when pain is a certainty in a lesion which has progressed, the decision to not treat still has greater expected utility.

The effect of changing the probability of restoration failure on the expected utility of treatment decisions is illustrated in Figures 9.5 (a-c). These Figures demonstrate that decreasing the likelihood of restoration failure affects the expected utility of positive treatment decisions more than the likelihood of pain. (Figure 9.4(a-c)). The utility of a positive decision is highest when dentists are treating only when using treatment threshold 1 (Figure 9.4(a)) and when there is a 100% chance of the restoration being permanent. However, the utility of a negative treatment decision is always greater than that of a positive one.

In Figure 9.6 (a-c) the effect of changing the population utility of unnecessary restoration on the expected utility of treatment decisions is shown. From these Figures, it is seen that if the utility of unnecessary fillings increases, the expected utility of a positive treatment decision rises. However, the expected utility of a positive decision still does not exceed that of a negative one. This implies that even if the population value for an unnecessarily filled tooth is at an equivalent value to that of a sound tooth (i.e. 1), the decision to fill a tooth always attracts a lower expected value than a decision to leave it unrestored.

Illustrated in Figure 9.7 (a-c) is the effect of increasing the probability of progression of outer enamel lesions on the expected utility of treatment decisions when the probability of pain after progression = 0.995, and when the likelihood of lesion regression is considered to be zero. The figure demonstrates that, even when lesion progression is a certainty, there is little increase in the utility of a decision to restore. The Figure also shows that the disparity between the expected utility of a decision to treat, and the expected utility of a decision to leave unrestored, increases as the treatment threshold is lowered.

The effect of increased likelihood of inner enamel lesion progression on the expected utility of treatment decisions when pain is virtually a certainty if the lesion progresses (probability of pain = 0.995) is examined in Figure 9.8(a-c). Again, lesion progression likelihood has little effect on the expected utility of positive and negative treatment

Figure 9.4 (a-c) The effect of increased probability of pain on the expected utility of treatment decisions, at three treatment thresholds
a = Rating 1 as positive
b = Rating 1-3 as positive
c = Rating 1-5 as positive

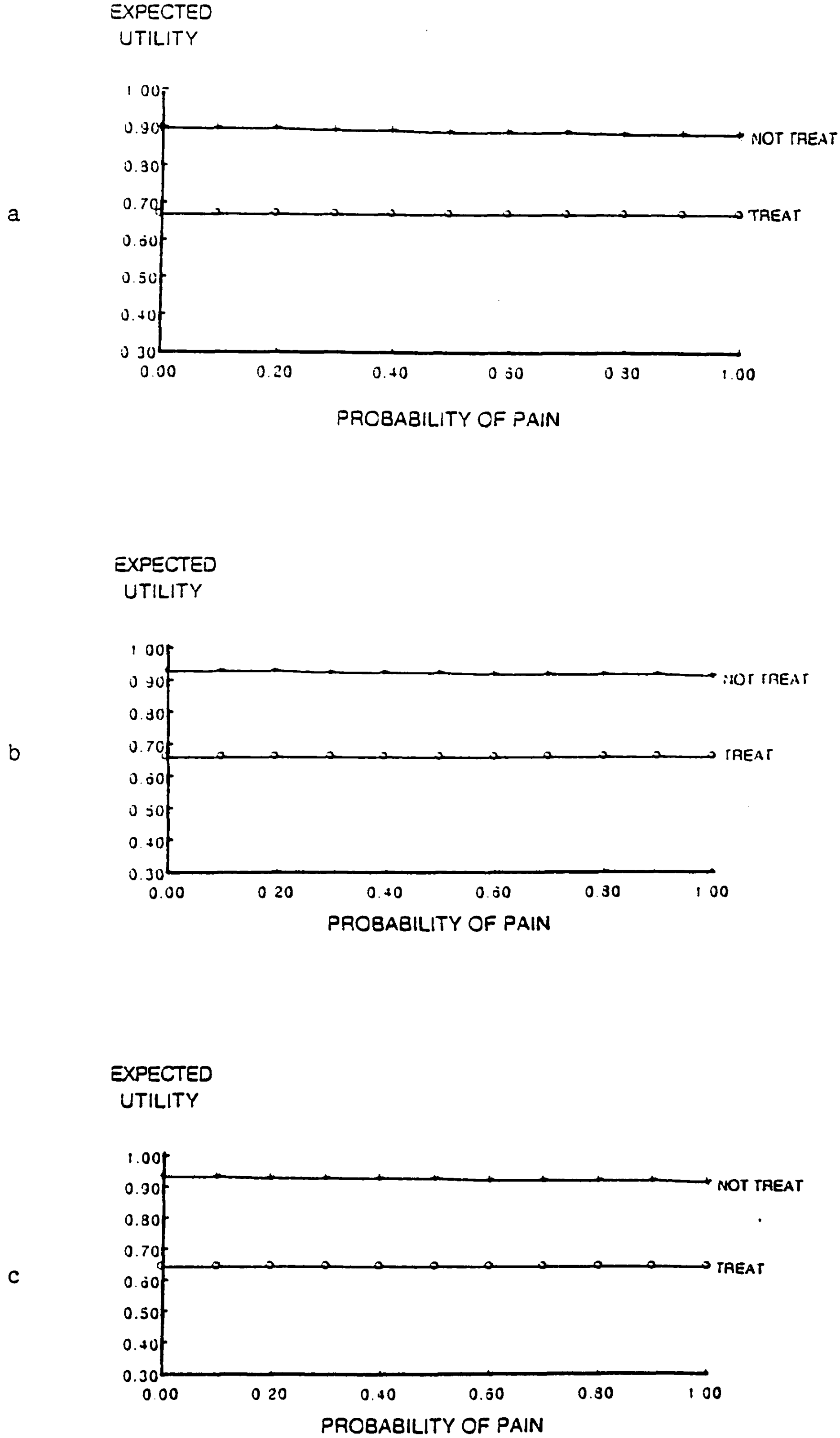
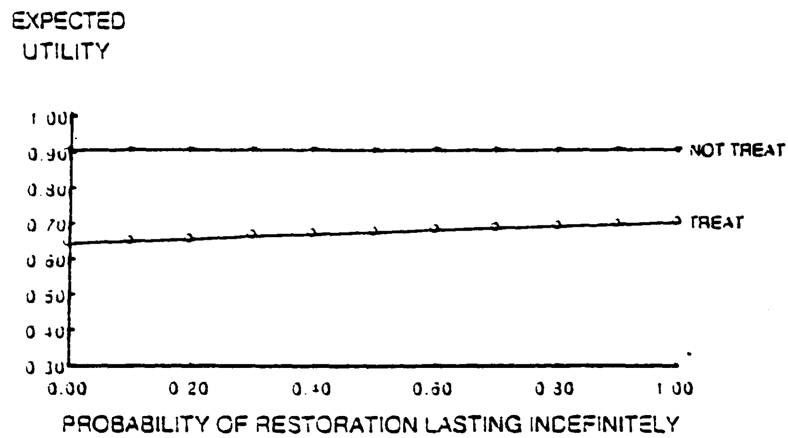
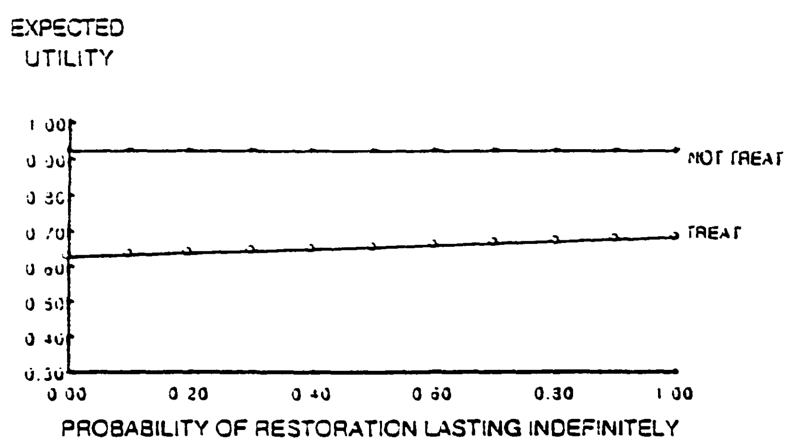


Figure 9.5 (a-c) The effect of increasing probability of restoration permanence on the expected utility of treatment decisions, at three treatment thresholds
a = Rating 1 as positive
b = Rating 1-3 as positive
c = Rating 1-5 as positive

a



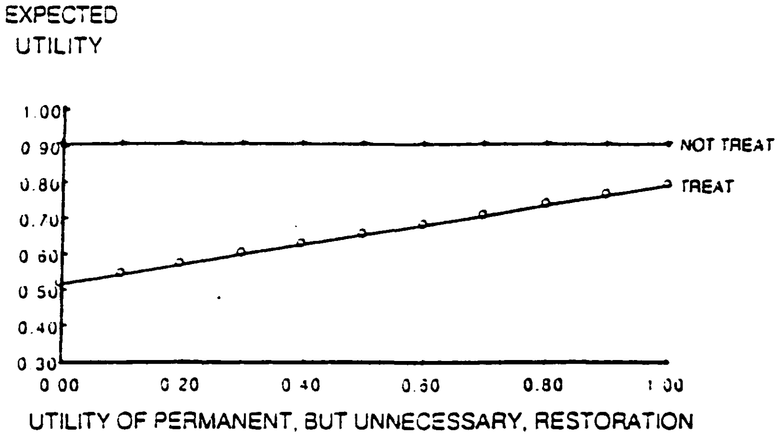
b



c

Figure 9.6 (a-c) The effect of increasing the population utility for unnecessary restorations on the expected utility of treatment decisions, at thre treatment thresholds
a = Rating 1 as positive
b = Rating 1-3 as positive
c = Rating 1-5 as positive

a



b

c

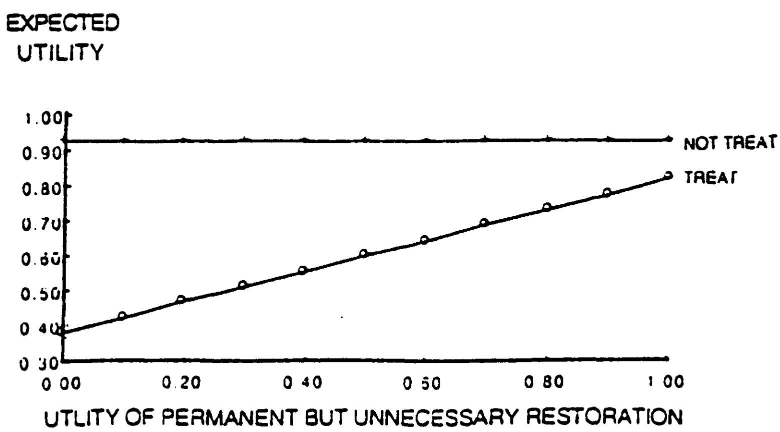


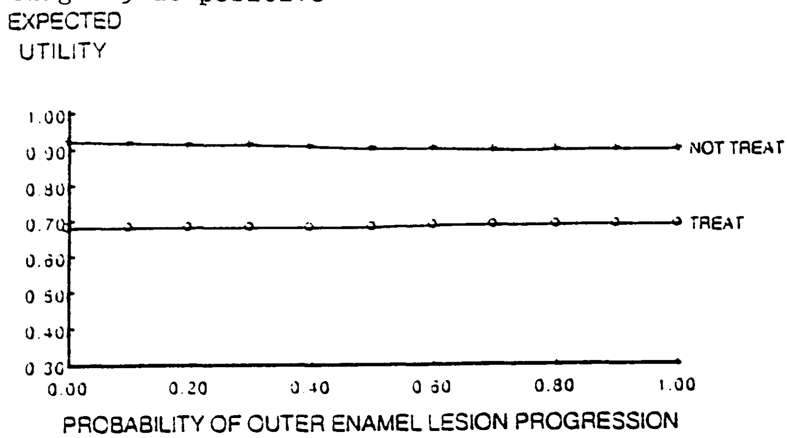
Figure 9.7 (a-c) The effect of increased probability of progression of outer enamel lesions on the expected utility of treatment decisions when the probability of pain after progression = 0.995.

a = Rating 1 as positive

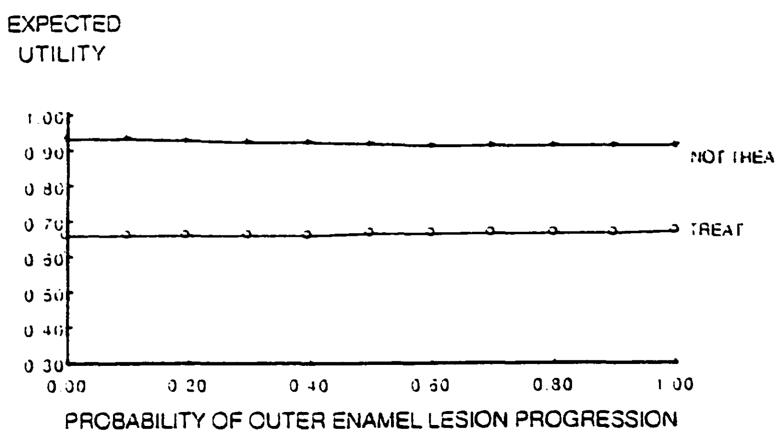
b = Rating 1-3 as positive

c = Rating 1-5 as positive

a



b



c

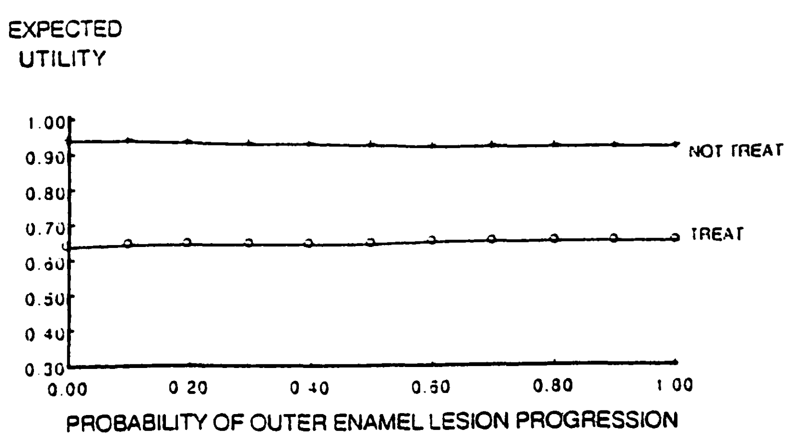
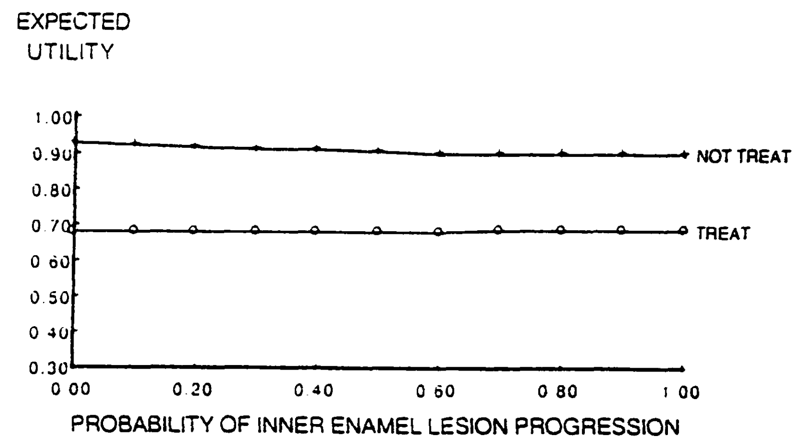
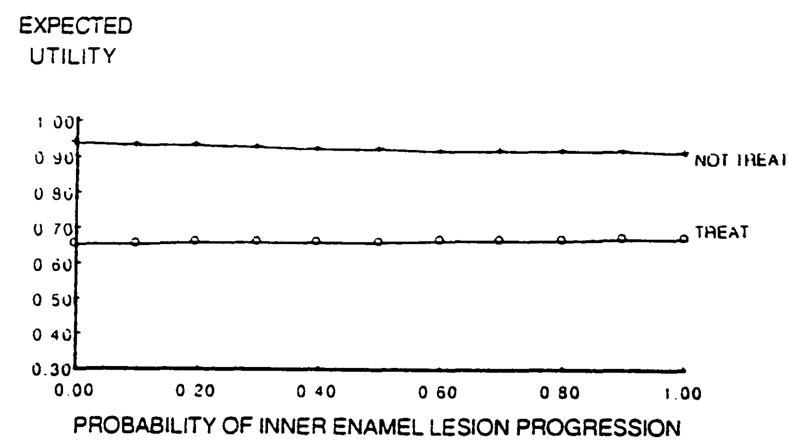


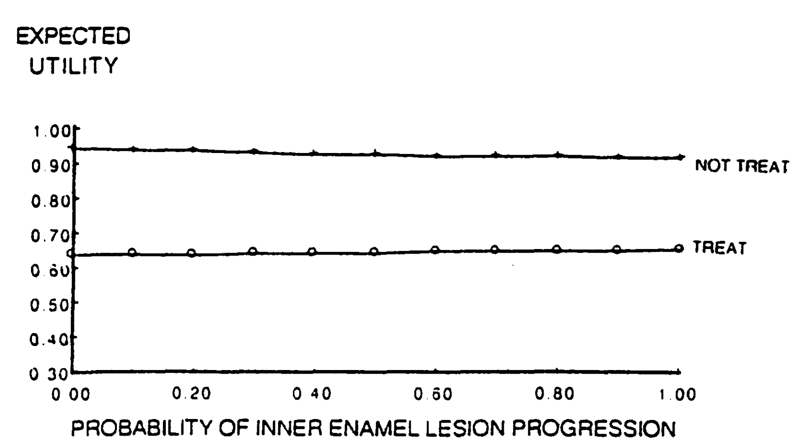
Figure 9.8 (a-c) The effect of increased probability of inner enamel lesions on the expected utility of treatment decisions at three treatment thresholds
a = Rating 1 as positive
b = Rating 1-3 as positive
c = Rating 1-5 as positive



a



b



c

decisions and, as in Figure 9.7(a-c) the strictest treatment threshold gives the highest expected utility for decisions to treat.

Figure 9.9 demonstrates using treatment threshold 1 alone, how treatment decision making is affected by the likelihood of decay in microscopically sound tooth surfaces. The Figure shows that the expected utility of a decision not to treat a tooth surface declines as the likelihood of decay in a sound tooth surface increases. If the probability of decay in a previously sound tooth exceeds 0.56 - the threshold - then the expected utility of a decision to treat exceeds that of a decision not to treat. This threshold is reached at a probability of 0.59 at when ratings 1-3 constitute a positive decision and at a probability of 0.61 when decisions 1-5 are deemed to be positive.

This concept is pursued via the data in Figure 9.10 by examining the optimal treatment decision when the probability of pain, and the probability of caries in sound teeth, are varied simultaneously. This Figure shows that as the probability of disease incidence in decay free teeth increases, a lower probability of pain from decayed teeth is required before treatment becomes the optimal decision.

The final analysis was to determine at what treatment threshold, the expected utility of a decision to treat becomes greater than that of a decision not to fill a tooth. If the predictive power of a negative decision is set at 70%, the expected utility of treatment rises to 0.72, with the expected utility of no treatment equalling 0.85. If the predictive power negative equals 60%, the expected utilities become even closer, at 0.72 and 0.79 respectively, for treatment and no treatment.

If the predictive power of negative decisions is equal to 0.50, then the population utility of unnecessary fillings becomes relevant to the expected utility of decisions. If dentists' treatment thresholds were held at a level where the predictive power of a negative decision was less than 0.5, as long as the utility of an unnecessary permanent filling was greater than 0.93, then decisions to treat have the highest expected utility. No other parameters i.e. increased probability of pain, restoration failure, or lesion progression, affected the highest expected utility of decisions being associated with a negative decision.

If the predictive power of a negative test falls to 40% the threshold value is equal to 0.76, i.e. if the population utility of an unnecessary filling is greater than this value, at this predictive power, the expected utility is maximized by decisions to treat.

However, if the predictive power of a positive test is raised to 95% the "no treatment" decision still attracts the greatest expected utility, with only the utility of a permanently filled tooth affecting the optimal decision (predictive power negative set at 70%). If the

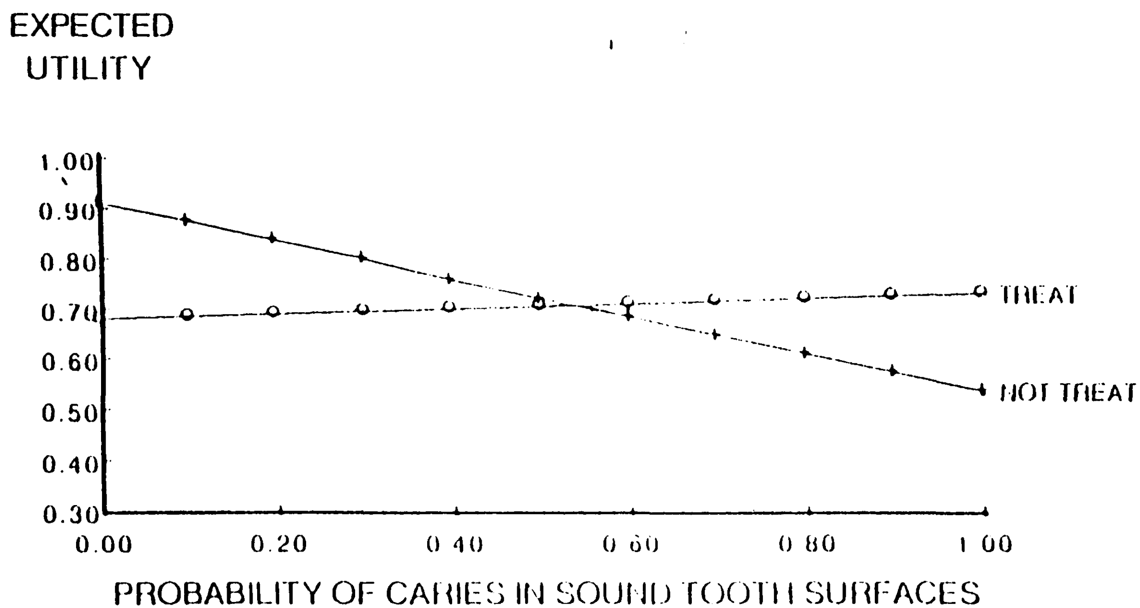


Figure 9.9 The effect of increased probability of disease in sound teeth on the expected utilities of treatment decisions (threshold 1 only)

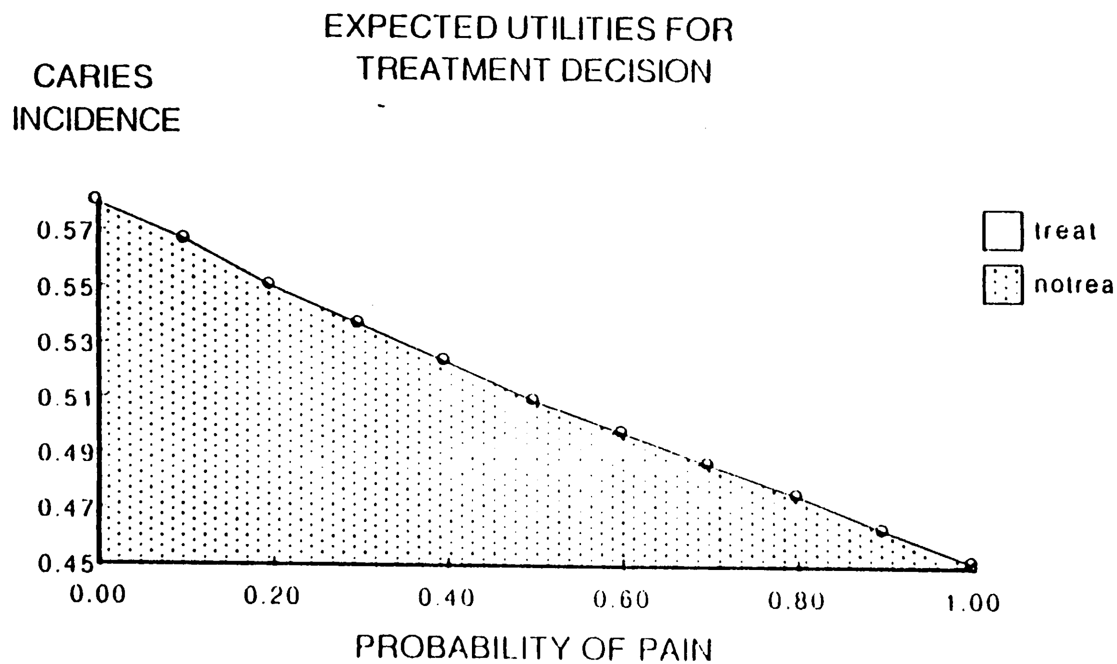


Figure 9.10 The effect of simultaneously increasing the probability of pain, and caries incidence on the expected utility of treatment decisions (threshold 1 only)

population utility of a permanently filled tooth is greater than 0.96, then treatment becomes the preferred option. At a predictive power negative of 60% (and predictive power positive = 95%) the threshold value for the utility of a filled tooth becomes 0.87, i.e. at this treatment threshold, if a permanently filled tooth were valued by the population as having 87% of the value of a sound tooth, then the maximum utility is achieved by opting to treat the tooth surface.

Once again the model is most sensitive to disease incidence. At a predictive power positive of 95% and predictive power negative equal to 70%, the decision to treat is optimal when the probability of disease incidence in untreated teeth exceeds 0.30, i.e. if 30% of sound tooth surfaces are expected to decay within 3 years. Similarly, if the predictive power positive of dentists' decisions were 95% and predictive power negative were 60% then, if the incidence of disease in sound teeth is greater than 18%, the decision to treat is optimal.

No other parameters present in the decision tree other than increased disease incidence and increased predictive powers positive affect the conclusion that maximum expected utility is obtained by negative treatment decisions.

Altering the positive predictive power alone, while leaving the predictive power negative set at the value achieved by the dentists in the study (75%) has little effect on the maximum expected utility of decisions. If the positive predictive power is raised to 95%, then treatment becomes the preferred option, but only under the circumstances that the utility of a permanently filled tooth was equivalent to that of a sound tooth.

9.5 DISCUSSION

The results in Chapter 5 showed that the dentists in this study were achieving the highest specificity values when the highest thresholds for restoration were used. However the sensitivity of the decisions increased as the treatment thresholds were lowered. The decision analysis shown attempts to determine which of these properties is of greater importance.

The analysis presented shows that, no matter which treatment threshold were used, the decision not to treat always gave the maximum expected utility. This result is due to two factors - firstly, the population utilities derived from the results of the study in Chapter 8 indicate that the utilities for decayed, pain-free and decayed, painful teeth were quite high, i.e. people do not see decay as a major problem, even when pain becomes an intervening factor. Secondly the dentists in the study almost invariably achieved higher specificity values than sensitivity. Therefore, because untreated caries is not viewed with great concern by the public, and because dentists are more proficient at identifying sound teeth than at recognizing decayed teeth, a decision **not** to treat a tooth maximises the expected utility in almost all cases.

Results indicate that, if all dentists use and continue to use the treatment thresholds employed by the dentists in this study, a positive treatment decision will, under any circumstances, lead to a "poorer" outcome than that achieved by leaving the tooth unrestored. However, if the utilities of positive treatment decisions are compared, it is clear that they achieve the greatest value when the highest treatment thresholds are held. i.e. when a dentist only fills a tooth when he is absolutely certain that it is decayed.

If practitioners were to set even higher treatment thresholds, they would lower the rate at which sound teeth are deemed to require treatment (37% at the highest thresholds in this study). By so doing, they would also be increasing the specificity of their positive treatment decisions. However, lowering of the threshold at which treatments were instituted would increase the sensitivity of the decisions, and therefore the expected utility of positive treatment decisions would be raised.

At the start of the study it had seemed likely that if pain was a consequence of an unrestored lesion, this would influence the expected utility of positive treatment decisions. However, although increased likelihood of pain enhanced the expected utility of positive treatment decisions, it never increased the value of a positive decision beyond that of a negative one. This is mainly a consequence of the fact that the population utility of a decayed pain-free, both, and a decayed painful tooth, were not found to be markedly different (0.56 and 0.46 respectively). Had these utilities been more widely separated, the influence of pain on treatment threshold setting would have been greater. This result leads to speculation about the population utilities. Clearly, someone who had previously experienced agonising pain from a tooth might place a much lower utility on a painful tooth than someone whose experience suggested that dental pain was always mild. Therefore the past experience of an individual must be considered when making implicit or explicit judgements on the values they apply to dental health states.

The results also show that, although a high probability of a restoration never requiring replacement enhances the expected utility of a decision to restore, it still does not raise the utility of a positive decision above that of a negative. This finding implies that, if dentists continue to treat disease using their current criteria, even if a new and revolutionary restorative material was developed which was guaranteed to last for a patient's lifetime, decisions not to treat still attract the greatest benefits, in the view of the population. However, increased permanence of restoration increases the utility of a positive decision to a noticeable degree.

It has been stated previously that it was deemed unethical to assess population utilities for unnecessarily filled teeth. The analysis described in Figure 9.6 shows that, if the population is unconcerned by unnecessary treatment, (there would be no chance of caries in dentine within three years if a sound tooth had been left unrestored), then the expected utility of positive decisions is dramatically increased. However, even if the population utility of an unnecessarily restored tooth is equal to that of a sound tooth, treatment does not become the preferred option. A necessarily filled tooth's value is only 76% that of a sound tooth. Therefore it is impossible for the utility of an unnecessarily filled tooth to be greater than this value (unless, of course, the population prefer unnecessary to necessary fillings).

The analysis shown in Figure 9.7 shows the expected utility when pain from a progressing lesion is almost certain. Even when there is a high chance of a lesion progressing, the virtual certainty of pain does not alter the optimal decision - namely a decision not to treat. Therefore even if every "missed" lesion progresses inexorably to pain, because of the relatively high population utility (0.46) put on a decayed and painful tooth, the decision to leave a tooth unrestored always gives the maximum utility. This is because mistakes are commonly made when dentists treat according to their thresholds shown in Chapter 5. This implies that dentists should employ a much more circumspect attitude to restorative treatment.

The main parameter which is found to affect the expected utility of treatment decisions is the probability of progressive decay in previously microscopically sound teeth. Data in Figure 9.8 indicate that positive treatment decisions become the most valued option, only if over half the sound teeth left untreated progress to decay. It is interesting to speculate whether dentists make an assumption that new disease is highly probable, which makes them set their treatment thresholds at the levels seen. Such speculation raises the hypothesis that dentists are perhaps still holding treatment philosophies more appropriate to the era when caries had an extremely high incidence.

It is shown in Figure 9.9 that, if the incidence of decay is greater than 0.58, even if pain is unlikely, treatment has the greater expected utility. If pain is always associated with decay then the incidence of disease can be as low as 0.45 before decisions to treat give the maximum expected utility. In reality, it is only if the incidence of caries rises dramatically that this information is useful. However, such a rise is not impossible, and amongst some groups of the population, a caries incidence of this order may be found. For example if a particular group of patients, of a particular age, sex and social class had this level of disease incidence, then the treatment thresholds used by the dentists in this study would be appropriate.

It can be seen that, even if dentists were to change their thresholds and false negative rates were to rise to 50%, it is still only the utility of unnecessary fillings which would alter the maximum expected utility being derived from 'not treating' to it being achieved by 'treating'.

It is only when the false negative decision rates are increased to 60%, that the population value of an unnecessary filling, affects the expected utility of treatment decisions. This finding implies that when false negative rates rise to 60%, positive treatment decisions are preferred, only if patients are unaware of any difference in value between a tooth which has been filled necessarily, and one which is unnecessarily restored.

Finally, the analyses presented consider the outcomes when the sensitivity of dentists' decisions is improved to 95%. In these circumstances it is still only disease incidence which affects the maximum expected utility of treatments when compared to non-treatment. This implies that, even if all fillings lasted a lifetime, and progression of all lesions is unremitting, dentists should not opt to fill teeth, even when their decisions are extremely accurate unless a high incidence of disease is expected. This is a consequence of the population not regarding a decayed tooth, even when painful, as a particularly poor health state. Therefore, unless caries incidence increases rapidly, restorative dental treatment is wasteful of resource in terms of the utilities held by the population sampled in Chapter 8. It can be assumed that the reason why people seek treatment is either that they underestimate the probability of untoward consequences of restorative treatment or that their individual utilities for the outcomes are different from the population averages shown in these analyses.

CHAPTER 10. DISCUSSION

10.1 INTRODUCTION

The work described in this thesis has attempted to evaluate critically the treatment decisions which are made by dentists on the basis of a much utilized diagnostic test - the bitewing radiograph. Research in the field of validating and evaluating dental diagnostic tests is limited at the present time.

A significant amount of public funds are expended on the bitewing procedure[126-128], yet there is little research supporting its validity as a means of diagnosing disease which is in need of treatment because, historically, it has been readily accepted as a valid procedure by the profession.

Diagnostic tests, such as bitewing radiographs, are frequently, although not invariably, initiated by practitioners seeking evidence of disease, rather than by the patient who is seeking a cure for a specific complaint. Therefore it is vital that those who undergo such an examination will benefit from the procedure. An assumption seems to be made that the benefit which a patient accrues by having their dentition radiographed as part of their "check-up", is that disease is detected and treated at a stage earlier than that at which it would become noticeable without a radiograph. A second assumption is that this early detection benefits the patient, in that the outcome derived from early detection with consequent restorative treatment, reduces the chance of sequelae of the disease which the patient would view as unfavourable. A third assumption is that a diagnostic test has an inherent validity, which is stable despite variable interpretations of the test results.

This work questions these assumptions, firstly by examining whether or not a patient can rely on the detection of disease at a stage when it is appropriate to restore the lesion, and secondly by investigating whether or not the patient can expect a better or poorer outcome contingent on the disease being detected and treated restoratively. Finally this thesis examines the factors which influence the validity of the above two assumptions.

10.2 THE REVIEW OF THE LITERATURE

10.2.1. Caries prevalence

The review of the literature confirms the conventional wisdom that the prevalence of dental caries is declining, at least in industrialized countries[27-33]. The review has also established beyond doubt that the caries process is slow and that progress of lesions in permanent teeth is by no means inevitable, at least in the short term[5]. However, the variability in lesion initiation and progression rates highlighted in the studies reviewed has important implications. The final analysis in this thesis demonstrates the increased importance and validity of the bitewing radiograph in individuals who are at high risk for caries. This thesis has therefore demonstrated that diagnostic efforts and services should be directed towards identifying such individuals, as it is for them that diagnostic procedures have the greatest impact on the expected utility of treatment decisions. One recommendation that can be made as a result of this thesis, is that the search for predictors of caries activity should continue unabated, so that the applications of tests and treatments are able to reach their maximum efficiency and effectiveness.

10.2.2 Signal detection theory

Signal detection is an extremely important means by which diagnostic tests can be evaluated. Its importance lies in the acceptance that the observers who use a diagnostic test are an integral part of the test, and will bring their own attitudes, beliefs and discriminatory powers to the test. The sensory process of lesion detection is a separate, although important, factor in the decision process of treatment planning[72,73].

The review of the literature concerning signal detection has highlighted the fact that once a practitioner is aware of the decision attitude he holds, he may be able to control his error rates in order that his treatment decisions conform to his beliefs about the costs of errors[76,77]. This brings the discussion back to the issue of withholding diagnostic tests unless they are justifiable. Thus, if a practitioner is aware of his error rates, and can recognise patients who have a high probability of disease, he may be able to adjust his decision attitude to conform with the patients' views of the values of his decisions, and will thus only apply the tests when they are likely to be of benefit.

Another important finding from the search of the literature concerning signal detection, is that the personal attributes of individual observers influence the accuracy of their treatment decisions to a greater extent than the training they receive[80,83]. The studies described show that dental practitioners who all receive similar training, show little conformity in their treatment decision-making. This suggests that the method of selection of the individuals who become dentists will have greater impact on the care delivered to the public, than alterations in the methods by which they are trained.

10.2.3 Restoration failure

The literature review has detailed the extent and probability of restoration failure[86-101]. Too often, clinical treatment decisions are made on the assumption that the treatment option will lead to a 'good' outcome, or at least a "better" outcome than leaving disease untreated. It is of great importance that such assumptions are made in the light of the probability of the real long-term outcomes. All of the potential outcomes of a decision to treat a tooth should be known before a treatment decision can be made rationally. One of the great advantages of using a decision tree for analysis of treatment options is that it forces one to enumerate and quantify the probability of all the possible outcomes - not only those which are desirable. The increasing popularity of clinical audit in dental fields, is likely to generate a wealth of information about the middle and long-term outcomes of treatment. This information **MUST** be made known to clinicians in order they can utilise this important data in their day-to-day clinical decision-making. The literature reviewed in this context has demonstrated that restoration failure is of considerable importance. The final analysis, presented in Chapter 9, has demonstrated that dentists must take the possible failure of their treatments into account when they decide to restore an approximal lesion. Dentists should therefore be encouraged to subject their own work to careful scrutiny, in order to ensure that the outcomes which they **feel** they achieve, do in reality occur. They would then provide themselves with feed-back about the outcomes of actions. Likewise, new materials and methods for treating patients must be subjected to research which examines their long-term success, as no treatment decision can be made rationally unless the true picture about restoration longevity is known.

10.2.4 Dental radiography

The review of the literature supports the hypothesis that dental radiography may bring to the observer's attention many more lesions than those which require restorative treatment[103,104], and that the bitewing radiograph must be seen only as a source of information about the probability of a lesion, rather than as a absolute diagnostic test. In essence, a practitioner should not be seeking to identify radiolucencies but should aim to use the radiograph, in conjunction with their experience, and any other information available, which may shed light on the probability of caries, in order to determine how lesions in need of treatment are predicted by a radiograph.

The review also indicates that, despite the increasing volume of evidence which questions the clinical importance of a positive radiographic finding, the number of films being used is increasing linearly[126,127]. The health costs of exposure of the population to radiation from x-rays have been shown to be small[128]. However, the fact that there are any costs, in terms of human life, contingent upon the use of radiographs, suggests that research aimed at clearly delineating the usefulness of bitewing radiographs as routine clinical tests, is long overdue.

10.2.5 The relationship of radiographic appearance to caries prevalence and lesion depth

It is clear that the routine use of bitewing radiographs in dental practice was preceded by a considerable amount of literature which strongly advocated its frequent application even before its usefulness for preventive procedures had been demonstrated[132-134]. Close scrutiny of this early literature demonstrates, merely that more lesions were shown on radiographs than could be detected clinically. Such a finding does **not** necessarily imply that the extra lesions shown on radiographs are clinically important. It can be assumed that this early literature encouraged the use of the technique, despite the fact that the terms "radiolucency" and "lesions in need of treatment" were assumed, rather than proven, to be equivalent. Later work demonstrates that the two items are not one and the same. Indeed, the literature reviewed highlights the importance of applying Bayesian principles to radiographic results.

10.2.6 Decision analysis

It is clear from the review of the previous uses to which decision analysis has been put, that one of the prerequisites for the construction of a useful decision tree is the availability of reliable information about both the beneficial and poor outcomes of decisions[144-148]. Thus, information about the probabilities of the outcomes of diagnostic tests and treatment decisions, must be made available to clinicians **and** patients before rational decisions about test and treatment options can be achieved. The review also highlights that optimum decisions cannot be made without first investigating the population's views about the values of health outcomes[194].

The review of the potential applications of decision analysis has demonstrated that the concept that treatment decisions are either 'right' or 'wrong' is incorrect[147], and that the simple dichotomy of disease/no disease is not applicable to any chronic ailment, particularly caries. Treatment decisions can achieve a high or low value outcome, each with a given probability, and it is the scientific evaluation of decisions, under the numerous sets of conditions which may prevail in an individual setting, which must be taken into account.

10.2.7 Variations amongst clinicians in treatment planning

The literature review has delineated the many studies which have shown that clinicians both diagnose caries, and arrive at treatment decisions, in an idiosyncratic manner[173-177]. It is clear that inconsistencies may be due not to errors but to differing attitudes, and one of the key components of this thesis is the search to explain the routes through which the discrepancies between clinicians' treatment thresholds arise.

The studies concerning observer variation show that, although systematic variation between clinicians is comparatively easy to quantify and explain, the random variation in decision-making is more difficult to interpret.

The sources of variation which might be supposed to arise through differing views of the radiographic signs which indicate the presence of disease, can be counteracted by 'errors' made in determining an appropriate interventive threshold. Therefore, the idiosyncrasies in dentists' treatment decisions highlighted in the review of the literature, can be explained by variations in clinicians' knowledge and attitudes, and their subsequent behaviours. For example, a practitioner may have appropriate knowledge about the radiographic appearance of a particular depth of lesion, and about the 'correct' depth of lesion at which to intervene. However, he may then hold a treatment threshold which will lead to under- or over-treatment. Alternatively, he may have inappropriate information about the clinical appearance of lesions requiring treatment, but his attitudes may compensate for this lack of knowledge. Finally, a clinician's behaviour may not adhere to either his knowledge or attitudes, due to misinterpretation of the true outcomes of his treatment decisions.

10.2.8 Evaluation of health outcomes

Defining how the public view the seriousness and severity of illness of any kind is a complex and time-consuming task, as demonstrated by the plethora of uncoordinated studies which have taken place in an attempt to quantify "health"[210-216].

It might be assumed, given the complexity and difficulty of valuing health outcomes, that the implicit patient values elicited by the clinicians treating them, might be sufficient. However, for as long as the Westernised medical model prevails, and doctors and dentists are expected to dictate the best treatment options for their patients, it is unlikely that clinicians will pay enough attention to patients' values. It is therefore preferable to measure values in an explicit way.

The concept which underlies the search for population utilities for health outcomes, is that it is not the professions who should choose patients' preferences, but the patients themselves. Such a premise of public sovereignty must be supported with respect to medical policy-making for publicly-funded health services. It is no longer sufficient for professions to adopt the paternalistic view that they "know best". It is a salient lesson that the public appear not to care particularly about the health state of their teeth. This being so, then perhaps the idea that people 'should' attend dental services for prevention and palliation of disease is untenable, and a case can be made for demand-led, rather than need-led services.

However, before such issues can be resolved, it is essential that population values about dental health are ascertained through further research. A more realistic view of how the population views various states of health and dental health, is also required before dental cost-benefit studies can have any real value.

The participation of the population in determining how resources are used for health care is an obvious, although frequently ignored, ethical issue.

10.2.9 Conclusions drawn from current knowledge

The conclusions which may be drawn in the light of the literature review are as follows.

1. Continuing monitoring of dental disease prevalence and severity is vital if anachronistic, and therefore inappropriate, clinical practices are to be revised in the light of changes in disease patterns.

2. Longitudinal studies of the progression and regression of lesions are comparatively rare - presumably because of the cost and difficulty of conducting such studies. A further reason for the paucity of such studies may be that ethical questions can be raised about investigations which detect, but do not treat disease, when the long-term outcomes of such inactivity are unknown. Finally, in populations where disease is treated by local practitioners at an unknown rate, it is impossible to definitively determine the rate of carious lesion progress.

3. Signal detection is a highly useful but much under-utilised theory in dental research. Further application of this technique to aspects of dental clinical decision making is timely, in order that dentists' decision attitudes may be more thoroughly understood.

4. There is a priority for research concerning the long-term fate of dental restorations. It is vital that rigorous clinical audit and monitoring is undertaken in order that practitioners can have information about restoration longevity made available to them. It is also essential that all new materials used for restoring teeth should have their relative permanence examined before they are adopted for use.

5. Dental practitioners should be made aware that all decisions (and particularly these involving diagnostic radiography) deal with probabilities, and not with certainties. To introduce the concept that mistakes are inevitable may be unpalatable to a profession which has traditionally accepted a somewhat simplistic view of disease and its signs and symptoms. If such a concept were accepted, the use of radiographs on a routine basis might become more circumspect.

6. Subjecting dental clinical decisions to decision analysis, would have a wide and important impact on both policy planning, and treatment of individual patients. The technique should be brought to the attention of the dental profession, if only to encourage them to seek the information they require in order to make rational, rather than pragmatic, treatment decisions.

7. Further research should explore the sources of random variation among dental practitioners when they make treatment decisions. Such research should aim to define the role of knowledge and attitudes as predictors of clinical behaviour.

8. It is a sad indictment of dental profession members that they have hitherto made almost no attempt to ascertain the true relationship of the normative view of dental 'health' in relationship to the population's perceived views of dental 'health'.

10.3 THE AIMS, OBJECTIVES AND METHODOLOGY OF THE STUDY

The overall aim of this thesis was to identify appropriate thresholds for interventive treatment of approximal carious lesions. However to achieve this it is necessary to take into account not only the outcomes of treatment decisions, but also the value of those outcomes. Likewise, it is necessary to investigate all the factors which influence the decision to restore teeth, be they perceptual, attitudinal, or informational. Therefore several objectives had to be achieved in order to attempt to answer the research question. Sufficient information was available from the literature to indicate, (although not unequivocally ascertain) the rate of carious lesion progression in permanent teeth, and the probability of restoration failure within a given time-period. Likewise, the results revealed in the literature concerning the relationship of a dentist's interpretation of a radiographic lesion and its actual visual and microscopic state, are equivocal.

Previous studies have tended to assume that lesions penetrating radiographically into dentine were cavitated, or have assumed that dentists' treatment decisions were based on the expectation of cavitation. They have therefore tended to fail to appreciate that treatment decisions, and treatment thresholds are based on perception, information and attitudes, as well as on the visual detection of lesions.

In order to achieve the objective of determining the relationship between treatment decisions, radiographic appearance and actual state of the tooth surface, it was necessary to (i) examine the depth of lesions which dentists intended to restore; (ii) examine the depth of lesions at which they did restore teeth, and (iii) relate these lesion depths to the visual and microscopic appearance of teeth.

By comparing what a dentist intends to do, and his subsequent actions, the accuracy of his radiographic interpretation can be ascertained. By determining the influences on his decision criteria, outwith the relationship between the actual lesion depth and the radiographic appearance, one can examine how a dentist's attitudes affects his accuracy. The two factors can act in conjunction, or in opposition, to determine which teeth a dentist decides to restore.

Finally, by comparing the actions of a number of dentists, the variation between them could be noted, and the analyses used in the study could be utilised to determine some of the sources of this variation. Meeting these objectives required an investigation which had several components. Hence the study population of dentists gave information as to their intentions, their actions and their attitudes, and validation of decisions used both visual and microscopic criteria.

Although the prime project assessed the dentists' attitudes to the setting of treatment thresholds, in order to achieve the main objective of the study, it was essential that the decisions made by participants were related to the population's view of the outcomes. A simplistic measure of these values can be made using visual analogue scales. However, this methodology is flawed for the following reason. When a decision to restore a tooth is taken, it is always made under conditions of uncertainty (although this may be of a greater or lesser degree). Therefore one is dealing with probabilities, and it is thus inappropriate to assess patients' preferences purely in terms of rating various health states against each other. Most people will accept risks in their lives in order to achieve something which they desire. This concept has been shown to be true of medical health states, and therefore is almost certainly true of dental health states. The importance of the use of the standard gamble technique for eliciting health state utilities, is that its use implies an acceptance that there is always a chance of false negative and false positive treatment decisions. In essence, the standard gamble technique takes into account all four cells of the traditional decision matrix, whilst the visual analogue scale can only deal with true positive and true negative decision outcomes.

It can therefore be seen that the methodologies used throughout this thesis have a single unifying concept, which accepts that all decisions relate only to probabilities, and that all decisions therefore have risks associated with them. The aim was to incorporate both dentists' and the public's risk attitudes in order that a coherent and more rational view of the outcomes of dentists' decisions may be taken.

The final part of this thesis applied decision analysis to the findings of the literature review and the two studies (the dentist's rating scale experiment, and the assessment of dental health state utilities). Therefore the elements of uncertainty and preference, mentioned earlier, are specifically combined in order to determine the benefits of 'treatment' or 'no treatment' decisions.

The decision trees constructed have also highlighted where more information is required, and the improvements in diagnostic and restorative techniques which would make a decision to restore a tooth more attractive in terms of expected utility. The meaning of the analysis could even be extrapolated to be assumed to indicate where health promotion efforts might be directed, in order to encourage patients' use of services. Health education might influence the utilities people hold for various health outcomes and therefore, service utilisation would be improved.

One problem may arise if treatment thresholds based on decision trees using population utilities are applied to individual patients. Each individual may hold different utilities and risk attitudes. Therefore, the construction of a tree may be a useful tool for clarifying the issues a dentist should raise with each individual patient, before deciding on a treatment option, rather than for indicating the precise treatment an individual should receive. Decision analysis of the type demonstrated in this thesis therefore has an important role. More widespread use of the technique would force the dental profession to consider all the possible outcomes of its actions, and would lay open in a stark fashion, all the assumptions they make. Moreover it would also force the profession to consider how patients really feel about treatment and finally, it would help to describe precisely, why a treatment offered to one patient might be withheld from another.

10.4 THE SIMULATED BITEWING RADIOGRAPHS

Considerable time and effort went into the construction of models which would allow realistic bitewing radiographs to be produced. Nonetheless it is unnecessary to point out that the resultant radiographs did not imitate exactly those taken *in vivo*. However, the aim of the work reported in Chapter 4 was not to produce bitewing films indistinguishable from those taken of humans, it was to simulate the radiographic properties of the human tissues normally seen on such radiographs.

The teeth used were natural ones and the radiographs were taken employing exactly the same tube current (mA), exposure time (sec) and tube voltages as would be used when radiographing a patient. The tube-object distance was also the same as that pertaining in the normal situation, and the perspex screen used to separate the models and x-ray tube had a similar mass coefficient of attenuation to that of cheek tissues. For these reasons, the images produced could be expected to replicate the density and contrast of enamel and dentine as seen in clinical films. The simulation of the soft and hard tissues surrounding the teeth was less successful. The bone-meal/acrylic mixture used to mimic bone, did not, nor could be expected to, reproduce the anatomical features or trabecular pattern of the human tissue. However, such was not the aim - it was the radiodensity of the material which was of greatest importance. Since the models in which the teeth were set were produced from impressions of the actual anatomical features of a maxilla and mandible, the bone thickness surrounding the teeth was similar to that of the real-life situation. Also, given that the radiodensity of the mixture was made as near as possible to that of a 'live' mandible, the aim of simulating the radiodensity of the tissues surrounding the dentition was achieved. Furthermore, few previous studies have attempted to produce such similarities between phantom models and the real-life situation.

Since receiver operating characteristic (ROC) curves were to be derived from the data of the dentist-observer study, it was also considered important that structures resembling the radio-density of the periodontal membrane and lamina dura were introduced into the models. Signal detection theory accepts that detection of 'signals' or, in this case, carious lesions, is affected by the amount of 'noise' generated from the observed object. Therefore, to detect lesions from radiographs where there are no distracting structures of varying radiodensity, would ignore the radiographic 'noise' generated on (real) bitewings by the anatomical structures normally surrounding the teeth.

Considerable trial and error in the construction of models resulted in radiographs which were deemed to be acceptable for the purpose of the study. The subject of treatment decision validation made on the basis of bitewing radiographs would be much easier if a simple, quick and inexpensive radiographic technique for simulating the human dentition could be devised.

The practitioners involved in the study were asked to express their opinions of the films, having not been informed that the films were taken *ex vivo*. Only three participants recognised that the films were not "real". However, seven commented on the somewhat strange appearance of the "lamina dura" and others mentioned the density variability of the films.

It would be unwise to consider this lack of comment by the involved practitioners as conclusive evidence that the bitewings were accepted as "authentic". The complex nature of the task which they performed seemed to have concentrated their minds on the areas of the films with which their treatment decisions were concerned, and therefore they may have noticed other features less than might have been expected.

A very important factor in this study was the selection of teeth which were used in the phantom models. It was necessary to achieve a disease prevalence in the test radiographs which did not exceed dramatically that which would be possible in an actual population of individuals. On the other hand, it was also necessary to select a sufficient number of teeth which would generate radiographic 'signals' of disease, in order that the number of trials needed to construct the ROC curves was kept within reasonable limits.

The aim was to select 240 anatomically suitable teeth for construction of 15 "dentitions". Of these, it was deemed sensible (for the reasons cited above), for 20% of the surfaces to have lesions which would have a possibility of appearing as radiolucencies on the resultant radiographs. Other requirements were that some restored teeth should be included, again in order to attempt to avoid the observers' feeling that they were making decisions under artificial conditions. The extent to which this objective was met is uncertain.

Obtaining 240 teeth, 80% of which were sound, and of which at least 15% were required to have only minimal lesions was, in itself, a considerable task. Teeth with early caries, and sound teeth, are only rarely extracted, and approximately 3000 had to be examined, before 15 'dentitions' could be assembled.

In conclusion, the radiographs produced fulfilled the function for which they were designed - namely, radiographs which contained an appropriate number of sound and carious teeth, in which the radiodensity of anatomical structures was simulated.

10.5 THE ACCURACY OF DENTISTS' TREATMENT DECISIONS

This study suggested that the correlation between cavitation and the extension of a lesion into dentine is by no means perfect, and the figures reported in this thesis suggest that cavitation is less often seen in conjunction with dentine caries than previously reported. Earlier studies[110] suggest that a lesion into dentine is combined with cavitation in 88% of cases, whereas the comparable figure found here was 77%. Such information indicates that if cavitation is taken as the point at which restorative therapy should be instituted, it cannot be predicted reliably from a radiograph even **IF** the film gave an exact indication as to whether the lesion had penetrated dentine.

Therefore, once again, it is clear that treatment decisions made on the basis of bitewing radiography are arrived at in relation to a series of probabilities, each contingent on the other. Assuming that a dentist would wish to restore lesions with cavities, and recognising that he can only detect lesions into dentine with a given probability, then, even if such a lesion is seen, there is only a 77% chance that a cavity would exist. Hence it is evident that even the most accurate treatment decision-maker is only likely to be correct in a proportion of occasions. Therefore, dentists who were wary of using the rating scale values of 1 and 6 were acting in a rational manner, and these who tended to adhere to the extremes of the scale were not accepting there was doubt in their judgements, unless of course they felt that the term 'definitely requires restoration' was not equivalent to '100% probability that the lesion requires restoration'.

The data generated in the study does, however, indicate that there is a trend for increasing likelihood of positive treatment decisions as the visual size and microscopic depth of a lesion increased. Therefore, the probability of a positive treatment decision increased, as the probability of cavitation increased.

Validation of treatment decisions is not straightforward, for two reasons. Firstly, there is no universally accepted point at which a decision to restore is "correct". and secondly, dentists would rarely make a simple judgement as to whether or not they would restore a tooth only on the basis of a radiograph. It was for these reasons that the results presented in Chapter 5 and the analyses in subsequent Chapters used three depths of lesions as validating criteria. It might be held that few people would accept that enamel lesions should be restored but, in fact, results of the study indicate that some practitioners do consider enamel lesions as evidence of restorative treatment need. However the methodology employed here ensured that, whatever one's opinions as to the 'correct' criterion, the results can be interpreted adequately .

The variation in dentists' reported treatment thresholds was expected, as previous authors have reported similar results[171,173,176,177]. Interestingly, although the sample was small, the proportions of dentists with each treatment threshold were broadly similar to those found in 1127 general and community dental practitioners in Scotland[175]. Whether this variation arises from training, or prior experience, has been examined in Chapters 6 and 7, and will be discussed later in this Chapter.

One of the most intriguing findings, which is highly relevant to this thesis, is that the depth of lesion which a dentist planned to restore, had little effect on how often he chose a 'correct' treatment option, nor on how often he made both type I and type II errors. It might have been expected that dentists with a "restore early" philosophy might have achieved higher sensitivity values, at the expense of low specificity. i.e. a dentist who wishes to restore **all** enamel and dentine lesions might be expected to be 'over-treating', when the validating criteria was caries into dentine. However, throughout this study, the dentists' intentions appeared to have little or no effect on treatment decisions. This suggests that concerns about over-treatment based on the treatment criteria **reported** to be used by dentists, may be unfounded.

Whether the accuracy of treatment decisions in this study is considered to be 'good' or 'bad', depends on one's views as to the importance placed on the outcomes of the dentists' actions. Basically, it depends on the intrinsic value of a sound, decayed, filled, and unnecessarily filled tooth. If a filled tooth is almost equivalent to a sound tooth then dentists should strive never to miss the opportunity to restore a lesion. Alternatively, if a decayed tooth is almost equivalent to a sound tooth in its value and longevity, then it is more acceptable to leave lesions unrestored, as the gain from treating them is small. Therefore, although on initial perusal of the sensitivity and specificity values achieved by the dentists in this study, one might feel that many 'mistakes' are being made, careful evaluation of the outcomes of the 'mistakes' indicated that they may be due to a rational manipulation of treatment thresholds, by the dentists involved.

More relevant to this thesis than sensitivity and specificity is the predictive power of positive treatment decisions. This value gives the proportion of tooth surfaces designated for treatment which are, in fact, carious. These posterior probabilities are the most clinically relevant measure. It could be argued, that if dentists were informed of the simple fact that over one-third of the teeth they decide to fill, do not have dentinal caries, those who feel that dentine caries is an indicator of treatment need, might quickly adjust the criteria they use when making a positive decision. However, it could also be argued that a dentist is in constant receipt of data about the predictive power of his positive treatment decisions. As a dentist opens a cavity, presumably he is able to detect whether or not a lesion is truly present. Thus, feedback from the action taken should dictate future treatment criteria. This statement implies that dentists should implicitly utilise Bayesian theory to influence their treatment planning i.e. they should be able to predict posterior probabilities from their past experience. Whether or not this is the case is open to speculation. Dentists who do not recognize when a tooth has been restored unnecessarily may perhaps be placing more faith in radiographic, than in clinical appearance.

Some of the dentists in this study apparently prescribed treatment for tooth surfaces which were not (in terms of conventional wisdom) in need of restorative care. Explanations for this observation are offered later in this Chapter, but one possibility which must be considered is that the techniques used for validation of the decisions (both visual inspection and microscopy) are not sufficiently accurate to detect disease which is, in fact, present. However, the visual and microscopic evaluations agreed with each other to an acceptable degree. Likewise the fact that the probability of a positive decision increased as the depth of lesion increased, suggests that the validating techniques were accurate even though it is generally accepted that all 'gold-standards' used to validate decisions may be flawed. Nonetheless it is usual to accept, for the purposes of such studies, that the validating criteria are as near to the 'recording angel' version of the truth as can be achieved.

The generation of ROC curves for each participant at each validating criterion, gives a revealing insight into the dentists discriminatory powers. These appear to have very little to do with what the practitioner stated he intended to detect, as his discriminatory ability was not affected by the depth of lesion he said he wished to restore. Thus, it can be concluded from Chapter 5 that dentists hold criteria which cause them frequently to treat tooth surfaces without dentinal caries, and that these criteria do not appear to be influenced by the dentists' views of the depth of lesion which should be restored.

10.6 THE EFFECT OF DENTISTS' VALUES ON THEIR TREATMENT DECISION-MAKING

The area under an ROC curve represents the probability that a random pair of carious/non-carious radiographic images will be correctly ranked, according to their disease status. However this probability only conveys the dentists' ability to discriminate between carious/non-carious lesions when sensitivity (the ability to detect disease when it is present) is weighted equally to specifity (the ability to correctly detect tooth surfaces which are non carious).

Therefore, although discrimination is tested, the curves do not explain why dentists operate in the way they do. For example, one dentist might hold beliefs which lead to a high rate of false positive diagnoses, in order to achieve a high rate of true positives. i.e. his treatment threshold would generate a point situated in the upper right hand corner of an ROC curve; whilst another dentist may operate at the opposite end of the curve. i.e. he accepts a low true positive rate in order to be sure to keep his false positive rate within a range which he feels to be acceptable.

It is clear, therefore, that sensitivity and specificity are neither equally weighted, nor uniformly weighted in the same way by individual practitioners. External factors affect the decision process, most specifically the relative costs which the dentist feels are contingent upon the two types of diagnostic errors (false positives and false negatives). The evidence presented in this work has attempted to separate these intrinsic discriminatory qualities (or detection of lesions) as much as possible, from the decision attitudes of the practitioners involved, by utilising ROC analysis and exploratory attitudinal questionnaires.

The initial, qualitative study detailed in Chapter 6, sought to identify issues which might influence the weight practitioners gave to diagnostic errors. No claim is made that the ten issues listed in the Chapter are exhaustive nor that the group of participants are representative of practitioners as a whole. This thesis merely utilised the views and practices of a small group of dentists in an attempt to investigate in depth (rather than breadth), the influences pertaining to dentists' behaviours. Research which seeks to explain why people behave in the way they do is often fraught with difficulties and limitations, and it is clearly unrealistic to attempt to quantify every influence which may pertain to a practitioner's treatment decisions. However, variation between dentists' treatment policies have been given considerable attention by the U.K. media, and yet little research has been directed towards pin-pointing any, let alone, all sources of error. Also, although the responses to the questionnaire are used to derive numerical expressions of the practitioners' decision attitudes, this is merely an empirical exercise which allows one to go some way to determining whether the practitioners weight specificity and sensitivity equally.

Comparing the results of Chapters 5 and 6, it must be noted that of the dentists who, overall, felt that always filling carious teeth was most important, only one dentist's treatment decisions ever achieved greater sensitivity than specificity, and this was only when all tooth surfaces rated 1-5 in the rating scale were considered as decisions to restore. Thus, either the questionnaire methodology for assessing decision attitudes was not valid, or else the dentists were unable to detect lesions which they would, in reality, wish to restore.

Such a hypothesis can be explored by examining Table 6.4. In this Table any dentist whose operating ratio for $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ is lower than his reported value, is overtreating in relation to his reported attitudes. However, as described above, it is more realistic merely to view this Table in terms of values which are greater or lesser than 1. Therefore, if a dentist's reported value for $\frac{V_{TN} + C_{FP}}{V_{TP} + C_{FN}}$ is greater than 1.0, he wishes to avoid false positive treatments more than he wishes to avoid false negatives. It is thus possible to indicate to a dentist, the interventive threshold at which he comes closest to behaving in a way which is concordant with his expressed views.

The majority of the dentists set treatment thresholds that brought about treatment decisions which concurred with their expressed attitudes, even when they were considered to restore teeth which they felt should "probably" be left unrestored. However, a minority were overtreating in relationship to their stated attitudes.

Some dentists wished to pursue an aggressively interventionist philosophy. Such a practitioner appears to hold a view that dentistry operates as a discipline which seeks, above all, to thwart the natural progress of disease. Inherent in such a philosophy is a high tolerance of any undesirable consequences of restorative treatment. Dentists with such an attitude might treat teeth, even when lesions might regress spontaneously, as restoration is seen as a means of promoting a definite and rapid return to health. The dentists who had a non-interventionist approach, would not wish to restore teeth unless such an action had a very high probability of long-term success, and when they could be sure that caries would not regress, despite any preventive efforts made. Clearly the most prudent approach lies somewhere between these two philosophies. However in this study, the results of both the reported and calculated decision attitudes appear to be skewed towards non-interventionist behaviour.

The data presented suggest two sources of variation between practitioners. Firstly, if the areas beneath individual dentist's ROC curves are dissimilar, variation stems from innate differences in discriminatory powers. Secondly, variations can be seen to arise from the treatment thresholds chosen by the dentists, which were, in turn influenced by the weighting placed on the outcomes of their decisions (TN,TP,FN,FP). The importance assigned to each outcome varied markedly between dentists (Table 6.2) and this therefore may explain many of the differences noted between observers, in this and other studies[179-182].

Finally, however, variation may be introduced by perceptual differences between dentists. Given that participants who intend to treat in a non-interventionist manner, may actually overtreat in relation to their attitudes, another source of variation is true "errors" of judgement, in that a dentist's treatment thresholds do not allow him to achieve his stated objectives. These may be regarded as mistakes, in relation to both the dentists' knowledge and his attitudes.

10.7 VARIATION BETWEEN DENTISTS IN TREATMENT DECISION-MAKING

One source of variation between dentists, was their reported treatment criteria. There was no uniform view as to the stage at which a carious lesion **should** be treated. It has been a recurrent theme in this thesis that it is impossible to determine one 'depth of lesion' at which a tooth should always be restored. The point of intervention should vary according to the dentists' views of patients' values, and according to the probability of favourable and unfavourable outcomes. Therefore, the source of the observed variation in treatment criteria may arise from each dentist's experience, in terms of his patients' values, and the success rates of his treatments. This is, of course, an optimistic view. The alternative explanation of the disparity amongst the twenty members of the profession sampled in this study, is that information concerning regression of lesions, and the slowness of caries progression, has not reached the practitioners. They may therefore be practising without the benefit of research evidence to guide their treatment decisions. However, further analysis of the agreements between dentists suggests that the variation in stated criteria may not be as important as it might at first seem. Since many dentist pairs with differing treatment thresholds still achieved reasonable levels of agreement, it would appear that the practitioners' judgments of the 'correct' level of intervention, are relatively unimportant when compared to the costs and benefits they assign to errors. Therefore, a dentist who intends to treat early in lesion progression (when caries extends only through enamel), may find himself making very similar treatment decisions to those made by a dentist who intends to treat only deep lesions. This can arise if the "early treater" holds a very high treatment threshold, and places a high value on true **negative** decisions, and a very high 'cost' on false **positive** decisions. Alternatively the "late-treater" may hold a low treatment threshold, and may place a high value on true **positive** decisions, and a high 'cost' on false **negatives**. Such a pair of dentists may intervene at similar levels, due solely to their value systems, rather than their stated criterion. This has implications for those training dentists, particularly for clinical trial calibration. It is insufficient to concentrate only on teaching what does, and does not constitute a 'carious' lesion. Each examiner in a clinical trial or screening exercise, must also place the same weight on each of the four components of the decision matrix if they are to 'diagnose' in similar ways. If dentists in practice are to treat patients in a similar manner to their colleagues, it will be necessary to obtain views as to the values of TN, TP, FN, FP, and attempt to make all practitioners accept them. This may be difficult and will depend on the origins of the values dentists assign to correct decisions and decision errors. For example, do they absorb their value systems from their tutors at dental school, or do they develop their values from their own experience? Alternatively, perhaps the values are derived in some way, from the values held by the patients each dentist has treated in the past. The results presented in Chapter 7 give no indication as to the sources influencing participants decisions. Clearly, to investigate these phenomena requires further research.

Another source of variation amongst dentists, is that they may actually see different images when they examine a bitewing radiograph. This may be due to differing lighting conditions and, possibly variation in visual acuity. The intention in the study carried out was to assess variation among dentists, as they practised normally. There was, therefore, no attempt to standardise the conditions under which the radiographs were viewed. In fact, the dentists involved were encouraged to view the test radiographs as they would their patients' films. Most utilised a standard view box, although some raised the films to the daylight in order to examine them more closely. Only one dentist used magnification. Therefore, one explanation of the results shown may simply be that the dentists saw different images due to their various viewing methods. The evidence that this was not the case, is that the results are largely in agreement with other studies in which practitioners have made treatment decisions under standardised conditions[182]. It also seems to be a much more empirically sound tenet, to take into account the effect of individuals' usual attitudes and norms when making treatment-decisions, as the research then has more relevance to the real-life practice of dentists. Personal attributes explain the majority of individuals' everyday behaviours[247]. It thus seem likely that such parameters also affect treatment decision behaviours.

A further possible explanation of the wide variations between practitioners is that they interpret radiographic images differently. Hence if two dentists agreed on the depth of lesion which warranted treatment, and also held similar views as to the values and costs associated with treatment decisions and diagnostic errors, they may still make different decisions if they had different beliefs as to how a lesion of a particular depth is represented on a radiograph. For example, a dentist may believe that the radiographic image cast by a carious lesion consistently underestimates its depth, whilst another dentist may believe that a radiographic shadow consistently appears to be deeper than the lesion itself. Such beliefs may lead to over-treatment in the first instance, and under-treatment in the latter.

Variation in treatment decision-making is a complex issue. The studies presented in this thesis have clarified some of the potential sources of variation, but dentists' attitudes and values appeared to be of the greatest importance, and little research has addressed this issue hitherto. Others, such as the effect of varying knowledge criteria, and viewing conditions, are also relevant, but have been investigated previously[79] and have not been found to explain consistently the variations seen between practitioners.

10.8 DENTAL HEALTH STATE UTILITIES

There is clearly a need to develop an instrument to quantify the benefits of dental treatment and preventive programmes. Data in Chapter 8 show that further consideration of; the items selected for measurement, the methods of scaling the values, and research determining the reliability, validity and responsiveness of the instruments described here, are needed.

Proper "weights" or utilities for use in structuring dental health-care decisions should be non-arbitrary, community-based and should truly reflect the public's view of the relative desirability of the various dental health states. These criteria imply that the measurement instrument must be reliable, valid and, importantly, able to be used with the general public. The studies described in Chapter 8 are offered as a small contribution towards the development of such an instrument.

Torrance[189] has pointed out, that although measures of Standard Gamble Technique reliability show there is inherent imprecision when assessing individual attitudes, the applications for which this instrument are intended require population means, rather than individual values. Thus precision could be enhanced by increasing the sample size and improving the sampling technique.

With regard to the validity of these measurements, it is widely accepted that the standard gamble can be taken as a criterion, when used in medicine. However, the study presented is an initial use of this technique in a dental context.

Currently, there is no explicit quantitative estimation of the effects of dental health services and dental interventions, on dental health outcomes. Clinical, management and planning decisions are based typically on qualitative subjective judgments that treatment benefits are likely to outweigh any harm which might arise from dental services. However, the rising cost of dental health-care[248], increasing competition between dentists[249], concern over wide variations in practice patterns[26], and greater involvement by patients and consumer groups in dental health-care policy-making, imply it has now become imperative that quantitative methods of assessing health outcomes are developed.

Given the results presented in Chapters 5-7, techniques which measure dental health outcomes without examining patients' attitudes to risk, are unlikely to be useful. The Chapters clearly show that treatment decisions are made under conditions of uncertainty. The decision criteria utilised by dentists incorporate the values implicit in the trade-off between true and false decisions. Therefore, any instrument which seeks to quantify the values patients place on the outcome of treatment decisions, must include an element which allows them to trade worst and best scenarios. for intermediate outcomes. Therefore, it is the view of the author that the well-respected Standard Gamble Technique (reported in this thesis), or the time trade-off technique, (described elsewhere[189]) for assessing dental health utilities would be the most useful, in terms of giving valid measurements of utilities.

A "unit of currency" such as a QATY (Quality Adjusted Tooth Year) would allow direct comparison of the benefits accruing to a population, in terms of years of tooth life, weighted according to their value, or 'utility'. If the population utilities of various tooth states were known, along with the longevity of teeth in various states, it would be possible to compare, for example, restorative and preventive dental programmes, in terms of the QATY's they provide. Such a unit is clearly needed as, at present, it is virtually impossible to compare the value of a service which provides filled teeth, of limited life-span, to a service providing sound teeth via preventive programmes. This is because the current dental health unit of measurement, the DMFT index, does not recognise that decay compromises the life-expectancy of a tooth more than a filled surface, and that a sound tooth is possibly of greater functional value than a filled one. For example, a DMF of 12 might represent a dentition with four decayed, four missing and four filled teeth: it might equally represent a dentition with six decayed and six filled teeth. Therefore, two functionally dissimilar dentitions give a similar 'score' when the DMF index is used. However by weighting the D, M and F components by their utility value, one composite score, which describes the overall "health" of the dentition could be derived.

For example if D=4, M=4, F=4, S=16, composite health score =

$$\frac{(0.46 \times 4) + (0 \times 4) + (0.72 \times 4) + (1 + 1\ 6)}{28} = 0.74$$

while if D=6, M=0, F=6, S=16, composite health score =

$$\frac{(0.46 \times 6) + (0 \times 0) + (0.72 \times 6) + (1 + 16)}{28} = 0.82.$$

The current lack of appropriate dental health measures perhaps explains why dental health health economics is a virtually non-existent field. An overall measure of the 'health' of a dentition would also undoubtedly be an extremely useful tool for those involved in dental health policy planning. In a world where demands for health-care are increasing rapidly, and health-care becomes increasingly expensive and highly technological, there is an increasing tendency for funding to be awarded on an economic basis. Therefore, while it is easy to attempt to justify dental services in terms of reducing toothache and days-off-work, unless clear and quantitative economic reasons are given for providing restorative care, dentistry may find itself underfunded, as it is likely, in the future, to increasingly have to compete with other services for funding.

In the context of this thesis, utility values have been sought for dental health states, specifically for the purpose of utilising these numerical values in a decision analysis. The assumption is, therefore, that the numerical values reflect the benefits and detriments associated in the populations' minds with the various states of health. There are problems with such an assumption. Firstly, there is an implicit problem in assigning a numerical value to an abstract entity, which is intangible, such as a 'value' or 'usefulness' of a state of health. Many of the outcomes of dental treatment are intangible, and therefore not easily measurable. What is the 'utility' to an individual of a planned, non-painful loss of teeth? What is the utility of avoiding restorative care? The problem of giving numerical magnitudes to intangible values is a profound one, and it is difficult to propose a solution, as similar difficulties have been found in medical health-care, and hitherto, no methodological procedure exists in order to solve these problems.

The second difficulty which arises when assigning utilities to dental health states, is that each of the outcomes has multiple attributes, which the work described has attempted to condense into one numerical value. For example, in the case of "tooth filled, and requiring re-restoration within five years", one should examine time lost from work to attend the dentist (measured in days), the economic cost of two restorative procedures (measured in

£'s), pain (measured in an appropriate unit), and the benefit of retaining a complete dentition (measured in other appropriate units). The 'utilities' assigned in the study described in Chapter 8 express all these attributes in one common unit. The validity of doing so is questionable, as the weights applied to each attribute by an individual will be dependent on his experience of that attribute.

These difficulties are inherent to decision analysis, and they are not readily managed. However, in defence of the approach taken, it would seem more sensible to attempt to place some sort of value on outcomes, even if they are only comparatively, and empirically acceptable. The claim is not being made that the utilities assigned in this thesis are in any way a meaningful numerical expression of the values of various health states. They merely indicate the comparative values of different outcomes, the probability of occurrence of which, are affected by the decisions dentists make.

Therefore, the studies presented in Chapter 8 have shown that it is possible to go some way towards determining the dental health state values of the general public. Such research is vital if the costs and benefits of dental care are to be investigated, and if rational policy-planning is to be based on decision analysis models.

10.9 THE DECISION ANALYSIS

The decision analysis presented in Chapter 9 of this thesis provides an insight into the real-world problem of restorative treatment decision-making. Although the models presented are rather complex, the insight and conclusions which can be drawn from the analyses are only helpful if the models really represent the problem.

The trees presented offer their principal benefits by having the capacity to allow one to ask "What if restorations fail more often than we imagine"; "What if decay **always** gives rise to pain"; "What if bitewing radiography becomes a more accurate diagnostic tool?" etc.

Thus, although the studies described (Chapters 2, 5 and 8) allowed the derivation of lesion progression rates, restoration failures, true positive and negative decisions, and the values populations might hold for certain states of dental health, such real-life data are not required in order to construct the tree. Whether or not such data is collected prior to construction of a decision tree is, in fact, immaterial. What is important is that the tree clarifies which clinically relevant data needs to be made available to a clinician when he is making a treatment decision. Thus, although it is unlikely that a dentist would seek to evaluate, in depth, the utilities for states of dental health held by an individual patient, the analyses presented showed the important influences on the expected utilities of a decision are: the strictness of the treatment criteria held, the likelihood of restoration permanence, the patient's valuation of unnecessary treatment, and the incidence of new disease in that patient. These findings do not concur with conventional thinking which tends to imply that the important parameters when making a treatment decision are the depth of the lesion, the likelihood of pain from the tooth, and the rate of lesion progression. These factors have been shown, in this study, to have minimal effect on the expected utilities of treatment decisions, and should therefore occupy the dentist's mind to a lesser degree than they might at present.

It may be that dentists actually do make clinical decisions about their patients in partnership with them. The information listed above as that which is most relevant to the decision, may in fact be sought from patients, during a dentist's clinical examination. This then may go some way to explaining the variations in treatment planning which occur when several dentists examine one patient. The relationship between a dentist and a patient, and the interactions and quality of communication between them may have a sizeable influence on the information derived. For example, when a new and obviously nervous patient presents to a dentist, what judgement will the dental surgeon make about the patient's attitude to treatment? Will he presume that the patient will place a low value on positive treatment decisions, or would he conclude that the patients' values must be high, since he has attended for treatment despite his obvious fear?

To some extent Chapter 6 explored how a dentist's attitudes influenced the decisions he made, but further research is required in order to examine the importance of, and the route by which dentists glean information about their patients' attitudes to treatment. It is, however, inherently obvious that this type of data-gathering on the dentist's behalf will be profoundly affected by the dentist-patient relationship.

It therefore seems that previous research examining dentists' treatment decisions, and indeed the way dentists are trained, have over-concentrated on the mechanistic view of the carious process, and have under-emphasised the importance of the salient factors highlighted in the decision analysis presented in this thesis.

The evaluation of the decisions made by the dentists in this study presented in the decision analysis, utilised utilities derived from a non-randomly selected population. The decisions the dentists' made related to an unknown person, of the age of 16, who was presented in a vignette which described the patient as having average caries experience. The decisions taken by the dentists were therefore made without them being able to make any judgements as to the tooth utilities which the hypothetical patient might hold. Thus, the findings of the decision analysis have clarified the types of information which should be gained from patients before dentists can make rational decisions. Although the practitioners from whom the data have been derived may have guessed at the patient's caries-risk-status from the radiographs, it must be assumed that the utilities held by the patient in question were derived from the practitioners' experience with his own pool of patients. Clearly, their patients may well hold completely different sets of values from those held by the populations involved in the utility studies (Chapter 8). Therefore, the results presented must not be interpreted as implying that the positive treatment decisions made by the practitioners were 'wrong' - they were merely inappropriate with regard to the populations used in the utility studies. Further exploration of how practitioners' decisions are influenced by their patients' attitudes and values are overdue. Likewise, the discrepancies between the utilities held by patients who attend, and those who do not attend dentists, would provide information which would indicate where some of the barriers to dental attendance lie.

CHAPTER 11. CONCLUSIONS AND RECOMMENDATIONS

1. All dental diagnostic tests should be evaluated and validated in terms of their sensitivity, specificity and predictive power, in order that rational judgements of the costs and benefits of such techniques can be made.

2. Research into factors which will predict likely caries activity is of vital importance. If individuals who are likely to be at high risk could be identified, the validity of any caries diagnostic tests be raised. Also valuable resources, currently wasted on tests which produce unreliable and unimportant information, and subsequent treatment which is frequently unnecessary, could be redirected towards those who would benefit most from them.

3. Longitudinal studies of both lesion progression and treatment outcomes are required, in order that reliable information is available to both practitioners and researchers. Sensible treatment decision-making is not possible, unless the long-term outcomes of treatment policies are known.

4. The innate abilities and attitudes of dentists performing diagnostic tests have a greater impact on their performance than their levels of knowledge, or the training they receive. This finding has enormous implications for those involved in the selection and education of future members of the dental profession.

5. Enumeration and realistic evaluation of all the outcomes of a positive treatment decision are required before the decision can be based on anything approaching rationality. Untoward outcomes should be brought to dentists' attentions and they should take these possible negative outcomes into account when planning treatment for their patients.

6. Research concerning valuation of the utility placed on dental health outcomes is in its infancy. Further definitive studies are urgently required in order that appropriate education, cost-benefit studies, and decision analysis can be applied to contentious issues in dentistry.

7. An effective and cheap method of simulating bitewing radiographs using *ex vivo* teeth with real carious lesions would greatly simplify efforts to evaluate bitewing radiography.

8. The relationship between dentine caries and cavitation is not a direct one. The dental profession needs to determine which of the two states justifies restorative treatment, and then delineate more clearly the radiographic appearance which most accurately indicates when such a need for treatment is present or absent.

9. The dental profession should appreciate more clearly that treatment decisions are made on the basis of their judgements of disease probabilities, rather than their ability to determine whether disease is absent or present. They should be made aware that caries is a process, rather than a diagnosis, and that rules about the exact radiographic appearance of a lesion in need of treatment cannot be drawn up even by "experts". Each dentist's judgements about treatment decisions must be made on the basis of the probability of disease, the probability of various outcomes, and the patient's views about the relative values of these outcomes.

10. Dentists should be made aware that the predictive power of their positive treatment decisions is low, even in places where caries is more prevalent than in many United Kingdom populations.

11. Dentists' intentions concerning the depth of lesion at which they wish to institute restorative care have little effect on the accuracy, sensitivity, specificity and predictive power of their decisions.

12. The values which dentists assign to each cell of the decision matrix affect the treatment decisions they make to a greater extent than their beliefs about the depth of lesions which warrant clinical intervention.

13. Further research is required in order to determine the route by which dentists come to assign different weightings to correct and incorrect treatment decisions. Training of dentists should be directed towards encouraging appropriate weightings to TN, TP, FN, and FP decisions, rather than in teaching them definite radiographic criteria which warrant restorative treatment.

14. If the population for whom treatment decisions were being made, were the same as that from whom the utilities in this thesis were derived, dentists would achieve maximum expected utilities by opting for a non-interventionist approach.

15. The probability of new disease in a patient is the parameter most relevant to decisions to treat lesions restoratively. Dentists should therefore employ radiography and opt for a more restorative attitude with patients in whom disease

rates are high. i.e. the *a priori* disease probability is high. Such a 'high-risk only' intervention philosophy is to be encouraged.

16. Information transfer between research work, dentist and patients, is much more relevant to clinical decisions than has previously been recognised.

Final Comments

This thesis has explored the treatment decisions made about approximal carious lesions in posterior teeth by a small group of dental practitioners.

The importance of this work lies in the validation of these decisions, not only in terms of 'gold standards' derived from the visual and microscopic appearances of the tooth surfaces, but also in terms of their utility to a population. Although it is accepted that the decisions made may not be representative of those made by all practitioners, this thesis has explored the decision attitudes of the practitioners, and the weightings they place on correct and incorrect decisions. This analysis has made a contribution to the search for explanations of the variations seen between dentists. Likewise, the study has highlighted the relevance of both populations', and individuals' values to a dentists decisions.

The findings presented have relevance to the considerations and training which dentists must undertake before their decisions can be made on a rational basis.

The studies have introduced three new methodologies by which dentists' decisions may be evaluated - namely ROC analysis, populations' perceived values, and decision analysis. These three techniques are intimately related, and further work remains to be undertaken before the methodologies can be developed in order to make them generally applicable to dental practice.

REFERENCES

1 Elderton RJ, Nutall NM. Variations among dentists in planning treatment. Br Dent J 1983; 154: 201-6.

2 Merrett MCW, Elderton RJ. An in vitro study of restorative dental treatment decisions and dental caries. Br Dent J 1984; 157: 128-33.

3 Elderton RJ. Treatment variation in restorative dentistry. Rest Dent 1984; 1: 3-8.

4 Watts A, Kay EJ, Paterson RC, Blinkhorn AS. Further studies on treatment decisions related to fissure caries. J Dent Res 1989;68:592.

5 Dirks OB. Post-eruptive changes in dental enamel. J Dent Res 1966;45:503-9.

6 Swets JA, Pickett RM. Evaluation of diagnostic systems. Methods from signal detection theory. New York 1982. Academic Press.

7 Swets JA, Tauner WP Jr, Birdsall TG. The evidence for a decision making theory of visual detection. (Tech Rep 40). Ann Arbor University of Michigan 1955.

8 Egan JP. Signal detection theory and ROC analysis. New York. Academic Press 1975.

9 Beck JR, Shultz EK. The use of the relative operating characteristic (ROC) curves in test performance evaluation. Arch Path Lab Med 1986;110:13-20.

10 Arrow KJ. Social choice and individual values, 2nd ed. Yale University Press 1963.

11 Kidd EAM, Pitts NB. A reappraisal of the bitewing radiograph in the diagnosis of posterior approximal caries. Br Dent J 1990;169:195-200.

12 Groenveld A. Longitudinal study of prevalence of enamel lesions in a fluoridated and non-fluoridated area. Community Dent Oral Epidemiol 1985;13:153-63.

13 Pitts NB, Renson CE. Monitoring the behaviour of approximal caries in the permanent dentition of posterior approximal lesions by image analysis of bitewing radiographs. Br Dent J 1987; 162: 15-21.

14 Maryniuk GA. In search of treatment longevity - a 30 year perspective. J Amer Dent Assoc 1984;109:739-44.

15 Department of Health Publication. Radiation protection in dental practice. London GPS (Print) Ltd 1989.

16 Department of Health. Dietary sugars and human disease. HMSO 1989.

17 Ziff S. The toxic timebomb: how the mercury in your fillings can affect your health. Aurora Press. New York 1984.

18 Dorfman DD, Alf E. Maximum likelihood estimate of parameters of signal detection theory and determination of confidence intervals - rating-method data. J Math Psych 1969;6:487-96.

19 Howat AP. Comparison of the sensitivity of caries diagnostic criteria. Caries Res 1981; 15: 331-7.

20 Sewerin I, Stolze K. Variabiliteten ved radiologisk caries-diagnostik blandt 845 danske tandllœger. Tandlœgebladet 1979; 83: 343-51.

21 Poulsen S, Bille J, Rugg-Gunn AJ. Evaluation of calibration trial to increase interexaminer reliability of radiographic diagnosis of approximal carious lesion. Community Dent Oral Epidemiol. 1980; 8: 135-8.

22 Pliskin JS, S hwartz M, Grondahl H-G, Boffa J. Reliability of coding depth of approximal carious lesions from non independent interpretation of serial bitewing radiographs. Community Dent Oral Epidemiol 1984; 12: 366-70.

23 Bulman JS, Osborn JF. Measuring diagnostic consistency. Br Dent J 1989;166:377-81.

24 Dice LR. Measures of the amount of ecological association between species. Ecology 1945;26:297-302.

25 Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measurement 1960;20:37-46.

26 The dentists who disagreed. Lead Article. Sunday Telegraph Magazine. 1988 Jan 31st.

27 Anderson RJ. The changes in dental caries experience of 12-year-old school children in two Somerset schools. Br Dent J 1984; 157: 20-2.

28 Downer MC. Secular changes in caries experience in Scotland. J Dent Res 1982; 61: 1336-9.

29 Fe jerskov O, Antoft P, Gadegaard E. Decrease in caries experience in Danish children and young adults in the 1970's. J Dent Res 1982; 61: 1305-10.

30 Kalsbeck H. Evidence of decrease in prevalence of dental caries in the Netherlands. J Dent Res 1982; 61: 1321-6.

31 Brunelle JA. Carlos JP. Changes in the prevalence of dental caries in US school children, 1961-1980. J Dent Res 1982; 61: 1346-51.

32 Glass RL. The declining prevalence of dental caries. The evidence and impact on dental education, dental research and dental practice. J Dent Res 1982; 61: 1301-81.

33 Brown RH. Evidence of decrease in the prevalence of dental caries in New Zealand. J Dent Res 1982; 61: 1327-30.

34 Clerehugh V, Blinkhorn AS, Downer MC, Hodge HC, Rugg-Gunn AJ, Mitropoulous CM, Worthington HV. Changes in the caries prevalence of 11-12 year old school children in the North West of England from 1968-1981. Community Dent Oral Epidemiol 1983; 11: 357-70.

35 Blinkhorn AS, Holloway PJ, Davies TGH. Continued effects of a fluoride dentifrice and mouthrinse on the incidence of dental caries. Community Dent Oral Epidemiol 1983; 11: 7-11.

36 Downer MC, Holloway PJ, Davies TGH. Clinical testing of a topical fluoride caries preventive programme. Br Dent J 1976; 141: 242-7.

37 Hodge HC, Holloway PJ, Davies TGH, Worthington HV. Caries prevention by dentifrices containing a combination of sodium monofluorophosphate and sodium fluoride. Report of a 3 year clinical trial. Br Dent J 1980; 149: 201-4.

38 Rugg-Gunn AJ, Holloway PJ, Davies TGH. Caries prevention by daily fluoride mouthrinsing. Report of a three-year clinical trial. Br Dent J 1973; 135: 353-60.

39 Todd JE, Walker AM. Adult Dental Health Vol I. England and Wales. 1968-78, 1980 HMSO, London.

40 Todd JE, Walker AM, Dodd P. Adult Dental Health, Vol II. UK 1982 HMSO, London.

41 Dirks OB, Amerongen J van, Winkler KC. A reproducible method for caries evaluation. J Dent Res 1951; 30: 346-69.

42 Stephen KW, Creanor SL, Russell JI, Burchell CK, Huntington E, Downie CFA. A 3-year oral health dose-response study of sodium monofluorophosphate dentifrices with and without zinc-citrate: anti-caries results. Community Dent Oral Epidemiol 1988;16:321-5.

43 Boyd JD, Cannon JJ, Leighton RE. Epidemiologic studies in dental caries. V. Placement of fillings as a source of statistical error of estimate as to dental caries progression rates. J Dent Res 1952; 31: 354-8.

44 Boyd JD, Wessels KE. Epidemiologic studies in dental caries III. The interpretation of clinical data relating to caries advance. Am J Pub Hlth 1951;41: 976-9.

45 Boyd JD, Wessels KE. Leighton RE. Epidemiologic studies in dental caries IV. Variability of progression rates of dental cavities in the occlusal surfaces of second molar teeth. J Dent Res 1952; 31:124-8.

46 Boyd JD, Wessels KE, Cheyne VD. Is the lactobacillus count a valid index of activity of dental caries? Proc Soc Exp Biol 1949;71:533-6.

47 Parfitt GJ. The speed of development of the carious cavity. Br Dent J 1956;100:204-6.

48 Berman DS, Slack GL. Caries progression and activity in approximal tooth surfaces. Br Dent J 1973; 134: 51-57.

49 Haugejorden O, Slack GL. Progression of approximal caries in relation to radiographic caries scoring codes. Acta Odont Scand 1975; 33: 211-7.

50 Rugg-Gunn AJ. Approximal carious lesions, a comparison of radiological and clinical appearances. Br Dent J 1972; 133: 481-4.

51 Zamir T, Fischer D, Fishel D, Sharav Y. A longitudinal radiographic study of the rate of spread of human approximal dental caries. Arch Oral Biol 1976; 21: 523-6.

52 Kolemainen L, Rytomaa I. Increment of dental caries among Finnish dental students during a period of 2 years. Community Dent Oral Epidemiol 1977; 5: 140-4.

53 Murray JJ, Majid ZA. The prevalence and progression of approximal caries in the deciduous dentition in British school children. Br Dent J 1978; 145: 161-4.

54 Gröndahl H-G, Hollender L. Dental caries and restorations. A six year longitudinal study of the caries increment of proximal surfaces. Swed Dent J 1979; 3: 47-55.

55 Gröndahl H-G, Hollender L, Malmcrona E, Sundquist B. Dental caries and restorations in teenagers. Swed Dent J 1977; 1: 51-7.

56 Granath L, Kahlmeter A. Matsson L, Schroeder U. Progression of proximal enamel caries in early teens, related to caries activity. Acta Odont Scand 1980; 38: 247-51.

57 Hyde EJ. Caries-inhibiting action of three different topically applied agents on incipient lesions in newly erupted teeth. J Can Dent Assoc 1973; 39: 189-93.

58 Hollender L, Koch G. Influences of topical application of fluoride on rate of progress of carious lesions in children. Odontol Revy 1969; 20: 37-41.

59 Muhler JC, Spear LB, Bixter D, Stookey GK. The arrestment of incipient dental caries in adults after the use of three different forms of SnF_2 therapy results after 30 months. J Am Dent Assoc 1967; 75: 1402-6.

60 Powell KR, Barnard PD, Craig GG. Effect of Stannous fluoride treatments on the progress of initial lesions in approximal surfaces of permanent posterior teeth. J Dent Res 1981; 60: 1648-54.

61 Schwartz M, Grondahl HG, Pliskin JS, Boffa J. A longitudinal analysis from bitewing radiographs of the rate of progression of approximal carious lesions. Arch Oral Biol 1984; 29: 529-36.

62 Cook SR. A longitudinal radiographic study of caries progression in dental students. Aust Dent J 1984; 29: 315-20.

63 Darvell BW, Pitts NB. The progression of approximal caries in enamel: a mathematical model. Aust Dent J 1984; 29: 111-5.

64 Pitts NB. Monitoring of caries progression in permanent and primary posterior approximal enamel by bitewing radiography. A review. Community Dent Oral Epidemiol 1983; 11: 228-35.

65 Shaw L, Murray JJ. The progression of approximal caries in the permanent dentition of British children. Community Dent Hlth 1986; 3: 199-205.

66 Ekanayake LS, Sheiham A. Reducing rates of progression of dental caries in British school children. A study using bitewing radiographs. Br Dent J 1987; 163: 265-9.

67 Dummer PMH, Oliver S, Shaw L. Changes in the distribution of decayed and filled tooth surfaces and the progression of approximal lesions. Br Dent J 1988; 164: 277-81.

68 Good IJ. The estimation of probabilities. An essay on Modern Bayesian methods. Cambridge Mass: MIT Press 1965.

69 Raiffa H. Decision analysis. Introductory lectures on choices under uncertainty. Reading Mass:Addison-Westley 1968.

70 Green DM, Swets JA. Signal detection theory and psychophysics. New York:John Wiley and Sons 1966.

71 Lusted LB. Decision making studies in patient management. N Engl J Med 1971; 284: 416-24.

72 Lusted LB. Perception of the roentgen image: applications of signal detection theory. Radiol Clin North Am 1969; VII: 435-45.

73 Yerushalmy J. The statistical assessment of the variability in observer perception and description of roentgenographic pulmonary shadows. Radiol Clin North Am 1969; 7: 381-92.

74 Garland LH. On the scientific evaluation of diagnostic procedures. Radiology 1949; 52: 309-27.

75 Edwards WM, Cox RS. Garland LH. The solitary nodule of the lung. An analysis of 52 cases treated by thoractomy and a study of preoperative diagnostic accuracy. Am J Roentgenol 1962; 88: 1020-30.

76 Alcorn FS, O'Donell E. Mammogram screeners: Modified program learning for non-radiologic personnel. Radiology 1968; 90: 336-41.

77 Alcorn FS, O'Donell E. The training of non-physician personnel for use in a mamography program. Cancer 1969; 23: 879-84.

78 Goldstein IL, Mobley WH. Chellenni SJ. The observer process in the visual interpretation of radiographs. J Dent Educ 1971; 35: 485-91.

79 Mileman PA. Radiographic caries diagnosis and restorative treatment decision making. Thesis. University of Groningen. 1985.

80 Herman PG, Hessel SJ. Accuracy and its relationship to experience in the interpretation of chest radiographs. Invest Radiol 1975; 10: 62-67.

81 Sheft DJ, Jones MD, Brown RF. Screening of chest roentgenograms by advanced roentgen technicians. Radiol 1970; 94: 427-9.

82 Thomas EL, Lansdown EL. Visual search patterns of radiologists in training. Radiol 1963; 81: 288-92.

83 Kundel HL, Lafollette PS. Visual search patterns and experience with radiological images. Radiol 1972; 103: 523-8.

84 Tuddenham WJ, Calvert WP. Visual search patterns in roentgen diagnosis. Radiol 1961; 76: 255-6.

85 Swensson RG, Hessel SJ, Herman PG. Omissions in radiology: Faulty search or stringent reporting criteria. Radiol 1977; 123: 563-7.

86 Ottolengui R. Failures with amalgam as commonly used. Dent Cosmos 1925; 67: 998-1008.

87 Tingley HE. The amalgam problem. New York J Dent 1940;10:410-7.

88 Easton GS. Causes and prevention of amalgam failures. J Am Dent Assoc 1941; 28: 392-5.

89 Roper LH. Restorations with amalgam in the army: an evaluation and analysis. J Amer Dent Assoc 1947; 34: 443-50.

90 Healey HJ, Phillips RW. Clinical study of amalgam failures. J Dent Res 1949; 28: 439-42.

91 Harvey OD. Caries resistance of filling materials. Dent Prog 1962; 2: 197-8.

92 Moore DL, Stewart JL. Prevalence of defective dental restorations. J Pros Dent 1967; 17: 372-8.

93 Attalla MN, Gibb GH. Failures in Class II amalgam restorations. J Can Dent Assoc 1968; 34: 11-5.

94 Black GV. Amalgam and amalgam fillings. Br J Dent Sci 1896: 39; 370-2.

95 Allan DN. The durability of conservative restorations. Br Dent J 1969; 126: 172-6.

96 Robinson AD. The life of a filling. Br Dent J 1971; 130: 206-8.

97 Richardson AS. Boyd MA. Replacement of silver amalgam restorations by 50 dentists during 246 working days. J Can Dent Assoc 1973; 8: 556-9.

98 Elderton RJ. An in vivo morphological assessment of occlusal amalgam restorations. PhD Thesis. University of London 1975.

99 Allan DN. A longitudinal study of dental restorations. Br Dent J 1977: 143: 87-9.

100 Crabb HSM. The survival of dental restorations in a teaching hospital. Br Dent J 1981; 150: 315-8.

101 Paterson N. The longevity of restorations. Br Dent J 1981; 157: 23-5.

102 Arnett JH, Ennis LM. Dental infection and systemic disease. Am J Med Sci 1933; 185: 777-81.

103 Burket LW. The accuracy of clinical and roentgenologic diagnosis of dental caries as determined by microscopic studies. J Dent Res 1941; 20: 70-5.

104 Trithart AH, Donelly CJ. A comparative study of proximal cavities found by clinical and roentgenographic examinations. J Am Dent Assoc 1950; 40: 33-7.

105 Webber RL, Benton PA, Ryge G. Diagnostic variations in radiographs. Oral Surg 1968; 26: 800-9.

106 Hefferen JJ. A review of approaches to the detection of dental caries. J Am Dent Assoc 1973; 86: 1359-64.

107 Grondahl H-G. Radiographic caries diagnoses and treatment decisions. Swed Dent J 1979; 3: 109-17.

108 Jung T. Die Kontakt Aachenkaries, Rontgenbefund und histologischer Befund. Dtsch Zahnartzl Z 1965; 20: 630-5.

109 Leijon G. Initial approximal primakaries i rontgenbilden. Swed Dent J 1969; 6: 227-37.

110 Marthaler TM, Germann M. Radiographic and visual appearance of small smooth surface carious lesions studied on extracted teeth. Caries Res 1970; 4: 224-42.

111 Purdell-Lewis DJ, Groenveld A, Pot TJ, Kwant W. Proximal carious lesions. A comparison of visual, radiographical and microradiographical appearance. Neth Dent J 1974; 81: 6-15.

112 Katz MA. A probability graph describing the predictive value of a highly sensitive diagnostic test. N Engl J Med 1974; 291: 1115-6.

113 Bille J, Thylstrup A. Radiographic diagnostic and clinical tissue changes in relation to treatment of approximal carious lesions. Caries Res 1982; 16: 1-6.

114 Aken J van. Limitations in clinical diagnosis of dental caries of approximal surfaces. Proc 12th ORCA Congr Utrecht 1965. Adv Fluoride Res Dent Caries Prev 1966; 4: 89-92.

115 Bucholz RE. Histologic-radiographic relation of proximal surface carious lesions. J Prev Dent 1977;4:23-4.

116 Downer MC. Concurrent validity of an epidemiological diagnostic system for caries with histological appearance of extracted teeth as the validating criterion. Caries Res 1975; 9: 231-46.

117 Gwinnett AJ. Comparison of proximal carious lesions as seen by clinical radiography, contact microscopy and light microscopy. J Am Dent Assoc 1971; 88: 1079-80.

118 Leijon G, Marken KE. Roentgenological diagnosis of proximal caries. Acta Odont Scand 1970; 12: 35-61.

119 Mileman PA, Purdell-Lewis DJ, Weele LT van der. Effect of variation in caries diagnosis and degree of caries on treatment decisions by dental teachers using bitewing radiographs. Community Dent Oral Epidemiol 1983; 11: 356-62.

120 Howat AP, Holloway PJ. The effect of diagnostic criteria on the efficiency of experimental clinical trials. J Dent Res 1977; 56: 116-22.

121 Howat AP, Holloway PJ, Brandt RS. The effect of diagnostic criteria on the sensitivity of dental epidemiological data. Caries Res 1981; 15: 117-21.

122 Nowak AJ, Creedon RD, Musselman RJ, Troutman KC. Summary of the conference on radiation exposure in pediatric dentistry. J Am Dent Assoc 1981; 103: 426-8.

123 Kendall GM, Darby SC, Rae S. The level of dental radiography in Great Britain. Br Dent J 1981; 151: 377-9.

124 Elderton RJ, Eddie S. Uptake of general dental service treatment in Great Britain 1965-1983, with special reference to radiography, prosthetics and the care of children. Br Dent J 1986;161:93-8.

125 United Nations Scientific Committee on the Effects of Atomic Radiation. 1977 United Nations. New York.

126 Wall BF, Kendall GM. Collective doses and risks from dental radiology in Great Britain. Br J Radiol 1983; 56: 511-6.

127 Taylor FE, Webb GAM. Radiation exposure of the UK population. NRDB-R77. National Radiological Protection Board 1978. Harwell.

128 Wall BF, Rae S, Darby SC, Kendall GH. A reappraisal of the genetic consequences of diagnostic radiology in Great Britain. Br J Radiol 1981; 54: 719-30.

129 Hirschmann PN. Newsletter of the British Society of Dental and Maxillofacial Radiology 1980 No 4.

130 Fixott HC. Use of roentgen rays in the practice of children's dentistry. J Am Dent Assoc 1937; 24: 91-5.

131 De Labarre FA. Preschool age dentistry. J Am Dent Assoc 1933; 20: 124-9.

132 White DP. Preliminary bitewing roentgenographic examination of naval aviation cadets. Naval Med Bull 1944; 43: 901-10.

133 Smith RK. The X-ray - an essential tool in diagnosis and prognosis of the child patients. J Am Dent Assoc 1942; 29: 796-9.

134 Barr JH. The diagnostic value at radiographic examination for proximal caries in deciduous posterior teeth. NZ Dent J 1945; 41: 89-92.

135 Gwinnett AJ. A comparison of proximal carious lesions as seen by clinical radiography and contact microradiography. J Am Dent Assoc 1971; 83: 1078-80.

136 Moller IJ. Clinical criteria for the diagnosis of the incipient carious lesion. In: James PMC, König KE, Held HR eds. Proceedings of the 12th Congress.

Advances in Fluoride Research and Dental Caries Prevention 4. Oxford: Pergamon, 1966; 67-72.

137 Moller IJ, Poulsen S. A standardized system for diagnosing, recording and analysing dental caries data. Scand J Dent Res 1973; 81: 1-10.

138 Newbrun E, Brudevold F, Mergamen HA. A microroentgenographic investigation of demineralized enamel: comparing natural and artificial lesions. Oral Surg 1959; 12: 576-84.

139 Soni NN, Brudevold F. Microradiographic and polarized-light studies of natural carious lesions. J Dent Res 1959; 38: 1187-94.

140 Soni NN, Silberkweit M, Parrish BA. A microradiographic-microphotometric and polarized light study of the demeralization pattern in carious lesions. Oral Surg 1965; 20: 321-9.

141 Crabb HSM, Mortimer KV. Two dimensional microdensitometry: a preliminary report. Br Dent J 1967; 122: 337-43.

142 Pitts NB. Detection and measurement of approximal radiolucencies by computer-aided image analysis. Oral Surg Oral Med Oral Path 1984; 58: 358-66.

143 Pitts NB. Regression of approximal carious lesions diagnosed from serial standardized bitewing radiographs. Caries Res 1986; 20: 85-90.

144 Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis. Science 1959; 130: 9-21.

145 Schwartz WB, Gorry GA, Kassirer JP, Essig A. Decision analysis and clinical judgement. Am J Med 1973; 55: 459-71.

146 Eraker SA, Politser P. How decisions are reached: Physician and patient. Annals Int Med 1982; 97: 262-8.

147 Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med 1980; 302: 1109-17.

148 Kassirer JP. The principles of clinical decision making. An introduction to decision analysis. Yale J Biol Med 1976; 49: 149-64.

149 McNeil BJ, Keeler E, Adelstein SJ. Primer on certain elements of medical decision making. N Eng J Med 1975; 293: 211-5.

150 Ingelfinger FJ. Decisions in medicine. N Eng J Med 1975; 293: 254-5.

151 Douglass CW, McNeil BJ. Clinical decision analysis methods applied to diagnostic tests in dentistry. J Dent Educ 1983; 47: 708-12.

152 Tulloch JC, Antczak-Boukoms AA. Decision analysis in the evaluation of clinical strategies for the management of impacted third molars. J Dent Educ 1987; 51: 652-60.

153 Antczak-Boukoms AA. Cost-effectiveness analysis of periodontal disease control. Abstract. Ninth Annual Meeting. Society for Medical Decision Making. Philadelphia 1987.

154 Marynuik GA, Tulloch JC, McNutt RA. A Clinical decision analysis of "faulty" amalgam replacement: when to reevaluate or replace? Abstract. Tenth Annual Meeting. Society for Medical Decision Making. Richmond 1988.

155 Antczak-Boukoms AA, Weinstein HC. Cost effectiveness analysis of peridontal disease control. J Dent Res 1987; 66: 1630-5.

156 Tulloch JC, Antczak-Boukoms AA, Berkey CS, Douglass CW. Selecting the optimal threshold for radiographic diagnosis of interproximal caries. J Dent Educ 1988; 52: 630-5.

157 Surmont P, Martens L, D'Hauwers R. A Decision tree for the treatment of caries in posterior teeth. Quint Int 1990; 21: 239-46.

158 Barza J, Pauker SG. The decision to biopsy, treat or wait in suspected herpes encephalitis. Ann Intern Med 1980; 92: 641-9.

159 Knill-Jones RP. Diagnostic systems as an aid to clinical decision making. Br Med J 1987; 295: 1392-6.

160 Malchow-Møller A, Thomsen C, Hilden J, Matzen P, Mindeholm E. A decision tree for early differentiation between obstructive and non-obstructive jaundice. Scand J Gastroenterol 1988; 23: 391-401.

161 Larvin M. McMahon MJ. Apache-II scores for assessment and monitoring of acute pancreatitis. Lancet 1980; 201-4.

162 Goldman L, Cook EF, Braid A. A Computer protocol to predict myocardial infarction in emergency department patients with chest pain. N Eng J Med 1988; 318: 797-804.

163 Bonsel GJ, Kompmaker IJ, 'T Veer van, Habbema JDF, Slooff MJH. Use of prognostic models for assessment of value of liver transplantation in primary biliary cirrhosis. Lancet 1990; 335: 493-6.

164 Curley RK, Cook MG, Fallowfield ME, Marsden RA. Accuracy in clinically evaluating pigmented lesions. Br Med J 1989; 299: 16-8.

165 Difford F. Defining essential data for audit in general practice. Br Med J 1990; 300: 92-4.

166 Hilden J. Mooney G. Medial Decision making. The challenges involved in the changing face of health care. Theor Surg 1988;2: 29-31.

167 Settle MB. White RP. Discussion summary. J Dent Educ 1983; 47: 713-4.

168 Mileman PA, Purdell-Lewis DJ, Dummer P, Weele LT van der. Variation in caries diagnosis and restorative treatment decision making when using bitewing radiographs. A comparison between dental schools. J Dent 1985; 13: 140-51.

169 Mileman PA, Vissers T, Purdell-Lewis DJ. The application of decision making analysis to the diagnosis of approximal caries. Community Dent Hlth 1986;3:65-81.

170 Bergren H, Welander E. The unreliability of caries recording methods. Acta Odont Scand 1960; 18: 409-20.

171 De Paola PF, Alman J. Assessment of the reliability of radiographic diagnosis in a caries clinical trial. J Dent Res 1972; 51: 1431-7.

172 Haugejorden O, Slack GL. A study of intra-examiner error associated with recording of radiographic caries at different diagnostic levels. Acta Odont Scand 1975; 33: 169-81.

173 Murray JJ, Shaw L. Errors in diagnosis of approximal caries on bitewing radiographs. Community Dent Oral Epidemiol 1975; 3: 276-82.

174 Richardson AS, Hale L, Williams JF. The role of radiographs in clinical studies of caries inhibiting agents. J Pub Hlth Dent 1971; 31: 158-65.

175 Nutall NHN, Pitts NB. Restorative treatment thresholds reported to be used by dentists in Scotland. Br Dent J 1990;169:119-26.

176 Rytomaa I, Jarvinen V, Jarvinen J. Variation in caries recording and restorative treatment plans among university teachers. Community Dent Oral Epidemiol 1979; 7: 335-9.

177 Nuttall NM. A quantitative and qualitative assessment of variation between treatment decisions made by dentists. - A health service perspective. PhD Thesis. University of Dundee 1987.

178 Hazelkorn HM. A comparison of dental treatment plans under different reimbursement systems. J Pub Hlth Pol 1985; 6: 223-35.

179 Nuttall NM, Davies JA. A comparison of epidemiological dental assessments with treatment subsequently undertaken. J Dent Res 1986; 65: 793.

180 Kay EJ, Patterson RC, Watts A, Blinkhorn AS. A preliminary investigation into the validity of dentists decisions to restore the occlusal surfaces of permanent teeth. Community Dent Oral Epidemiol 1988; 16: 91-4.

181 Espelid I. Radiographic diagnoses and treatment decisions on approximal caries. Community Dent Oral Epidemiol 1986; 44: 31-7.

182 Mileman PA, Purdell-Lewis DJ, Weele L Th van der. Variation in radiographic caries diagnosis and treatment decisions among university teachers. Community Dent Oral Epidemiol 1982; 10: 329-34.

183 Hinkle LE, Redmont R, Plummer N, Wolff HG. An examination of the relation between symptoms, disability and serious illness, in two homogeneous groups of men and women. Am J Pub Hlth 1960; 50: 1328.

184 Wyler AR, Masuda M, Holmes TH. Seriousness of illness rating. J Psychosom Res 1968; 11: 363-74.

185 Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being for a health status index. Hlth Serv Res 1973; 8: 228-45.

186 Williamson JWM, Alexander, Millar GE. Priorities in patient-care research and continuing education. J Am Med Assoc 1962; 204: 303-10.

187 Packer AH. Applying cost-effectiveness concepts to the community health system. Op Res 1968; 16: 227-35.

188 Goldsmith SB. The status of health status indicators. Hlth Serv Res 1972; 87: 212-20.

189 Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. Socio-econ Plan Sc 1976; 10: 128-41.

190 Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chron Dis 1978; 31: 697-704.

191 Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Hlth Serv Res 1972; 7: 118-33.

192 Sackett DL. Data requirement of users of hospital discharge abstracts for epidemiologic research. Med Care 1970; 8: 205-9.

193 Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. Oper Res 1979; 28: 206-24.

194 McNeil BJ, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. N Eng J Med 1982; 306: 1259-62.

195 Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981; 211: 453-8.

196 Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979; 47: 263-91.

197 Fishburn PC, Kochenberger GA. Two-piece Von Neumann Morgenstern utility functions. Dec Sci 1979; 10: 503-18.

198 Eraker SA, Sox HC. Assessment of patients preferences for therapeutic outcomes. Med Dec Making 1981; 1: 29-39.

199 Beyeth-Marom R. How probable is probable? A numerical translation of verbal probability expressions. J Forecasting 1982; 1: 257-69.

200 Lichtenstein S, Newman JR. Empirical scaling of common verbal phrases associated with numerical probabilities. Psychonomic Science 1967; 9: 563-4.

201 Llwellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The measurement of patients' values in medicine. Med Dec Making 1982; 2: 449-62.

202 Von Neumann J, Morgenstern D. The theory of games and economic behaviour (3rd ed). New York: John Wiley, 1953.

203 Culyer AJ. Measuring Health. Lessons for Ontario. Toronto: University of Toronto Press, 1978.

204 Weinstein MC, Feineberg HV. Clinical decision analysis. Philadelphia: WB Saunders, 1980.

205 Anderson NH. How functional measurement can yield validated interval scales of mental qualities. J Appl Psychol 1976; 61: 671-5.

206 Fischer GW. Utility models for multiple objective decisions: do they accurately represent human preferences. Dec Sci 1979; 10: 451-60.

207 Fauschel S, Bush JW. A health status index and its application to health services research. Operations Res 1970; 18: 1021-31.

208 Boyle MH, Torrance GW. Developing multiattibute health indexes. Med Care 1984; 22: 1045-57.

209 O'Connor AMC, Boyd NF, Trichler DL, Kriukov Y, Sutherland H, Till JE. Eliciting preferences for alternative cancer drug treatments. The influence of framing, medium and rater variables. Med Dec Making 1985; 5: 453-63.

210 Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chron Dis 1985; 38: 27-36.

211 Torrance GW. Measurement of health status utilities for economic appraisal. J Hlth Econ 1986; 5: 1-30.

212 Nikias MK, Sollecito WA, Fink R. A empirical approach to developing multidimensional oral health status profiles. J Pub Hlth Dent 1978; 38: 148-57.

213 Nikias MK, Sollecito WA, Fink R. An oral health index based on ranking of oral status profiles by panels of dental professionals. J Pub Hlth Dent 1979; 39: 16-26.

214 Kaplan RM, Bush JW, Berry CC. Health status: types of validity and the index of well-being. Hlth Serv Res 1976; 11: 478-507.

215 Marcus M, Koch AL, Gershen JA. A empirically derived measure of oral health status for adult populations. J Pub Hlth Dent 1980; 40: 334-45.

216 Marcus M, Koch AL, Gershen JA. Construction of a population index of adult oral health status derived from dentists' preferences. J Pub Hlth Dent 1983; 43: 284-94.

217 Cushing AM, Sheiham A, Maizels J. Developing socio-dental indicators - the social impact of dental disease. Community Dent Hlth 1986; 3: 3-17.

218 Westert GP, Oosterhaven SP, Beouma J, Schaub RHM. Significance of dental problems to the public and their comparability with general health problems. Community Dent Oral Epidemiol 1988; 16: 360-3.

219 Reisine ST. The effects of pain and oral health on the quality of life. Community Dent Hlth 1988; 5: 63-8.

220 Pitts NB. A Study of the progression and activity of approximal carious lesions in human posterior enamel. PhD Thesis. University of London 1984.

221 Oishi TT, Parfitt GJ. Effects of varying peak kilo voltage and filtration on diagnostic dental radiographs. Can Dent Assoc J 1976; 42: 449-52.

222 Hodge HC, Bale WF, Warren SL, Van Huysen G. Factors influencing the quantitative measurement of the roentgen-ray absorption of tooth enamel. Am J Roentgenol Theor Nucl Med 1935; 34: 817-38.

223 Wuehrmann AH, Curby WA. Radiopacity of oral structures as a basis for selecting optimum kilovoltage for intra-oral roentgenograms. J Dent Res 1952; 31: 27-32.

224 Van der Linden LWT, Van Aken J. Het ligamentum paradontale op de röntgenfoto. Ned Tijdschr Tandheelkd 1971; 78: 376-82.

225 Arnold LV. The radiographic detection of initial carious lesions on the proximal surfaces of teeth. Thesis. University of Groningen, 1983.

226 Webber RL, Benton PA, Cvar JF, Ryge G. Diagnostic significance of intratissue contrast in bitewing radiographs. Oral Surg 1969; 28: 352-8.

227 Hedin M, Halse A. Does narrow tubular collimator produce better dental radiographs. Dentomaxillofac Radiol 1981; 10: 43-8.

228 Jones DEA, Raine HC. Letter to the editor. Br J Radiol 1949; 22: 549-50.

229 Wainwright WW. Filtration for lowest patient dosage in dental radiography. Oral Surg 1963; 16: 561-571.

230 Seemann HE, Cleare HM. Reduction of X-radiation dosage by filtration. Dent Radiogr Photogr 1955; 28: 61-3.

231 Manson-Hing LR. Kilovolt peak and the visibility of lamina dura breaks. Oral Surg 1971; 31: 268-73.

232 Webber RL, Benton PA, Ryge G. Diagnostic variations in radiographs. Oral Surg 1968; 26: 800-9.

233 Richards AG, Webber RL. Constructing phantom heads for radiation research. Oral Surg 1963;16:683-90.

234 Downer MC, O'Mullane DH. Comparison of the concurment validity of two epidemiologic diagnostic systems of caries evaluation. Community Dent Oral Epidemiol 1975; 3: 20-4.

235 Levine RS. The scientific basis of dental health education. A policy document. Health Education Council. London: HMSO, 1985.

236 Kennedy DB. Paediatric dentistry - dental practitioners handbook No 21. Bristol: John Wright and Sons, 1976.

237 Guilford JP. Psychometric methods. New York: McGraw-Hill, 1936.

238 Conklin ES. The scale of values method for studies in genetic psychology. Univ Ore Pub 2 (Cited by Guilford, 1936).

239 Symonds PM. On the loss of reliability in ratings due to coarseness of the scale. J Exp Psych; 7: 456-61.

240 McNicol D. The power of signal detection theory. George Allen and Unwin: London, 1972.

241 Lusted LB. Decision making studies in patient management. New Engl J Med 1971;28:416-24.

242 McNeil BJ. Pauker SG. Incorporation of patients values in medical decision making. In: McNeil BJ, Carvelho FD eds. Critical issues in medical technology. Boston MA: Auburn House, 1982.

243 Acton JP. Evaluating public programs to save lives: the case of heart attacks. Report R-950-RC. Santa Monica: The Rand Corp, 1973.

244 Bush JW, Chen MM, Patrick ML. Health status index in cost effectiveness: analyses of PICU programs. In: Health Status Indexes. Berg RL (ed) pp 172-208.

245 Maynard AK. What price life? THS Health Summary 1984;2:4.

246 Williams AH. Coronary artery bypass grafting: an economic appraisal. Br Med J 1985;291:326-9.

247 Taylor SE. On inferring one's attitude from one's behaviour: some determining conditions. J Personality Soc Psychol. 1966;4:400-8.

248 Annual Report of the Chief Medical Officer. On the State of the Public Health. Br Dent J 1987;162:281.

249 Fyffe HE, Pitts NB. Origin, training and subsequent practice location of Scotland's general and community dentists. Comumunity Dent Oral Epidemiol 1989;17:325-9.

APPENDIX I

The letter to dentists requesting participation

UNIVERSITY OF GLASGOW
Department of Community Medicine

Henry Mechan Chair of Public Health
& Head of Department
PROFESSOR J. MCEWEN

2 Lilybank Gardens
GLASGOW, G12 8RZ

Telephone: 041-339 8855
International: -44-41-339 8855
Telex: 777070 UNIGLA
Fax: 041-330 4808

EJK/MS

Dear

You will be aware that the regulations governing dental radiography are becoming increasingly stringent. In order that we are able to specify the precise importance of bitewing radiography to our patients, I am undertaking a research project, in which I would very much appreciate your assistance.

I was wondering whether it would be possible for me to visit you at your practice, to ask you some questions concerning bitewing radiography. It would take approximately 30 - 45 minutes to complete the interview which I have planned. I do realise that this is a great imposition on your goodwill, and that many students ask for your co-operation, but I hope you will feel that this research subject is important enough for you to 'spare' me some of your time.

All of the results of my study will be in <u>strictest</u> confidence and no-one (including myself) will be able to identify your personal views from the results.

If the study is to achieve its aims, it is essential that my results pertain to successful general dental practitioners within Glasgow rather than NHS hospital or academic dentists. I do hope you will feel able to help me. If so, please return the slip below in the reply paid envelope at your earliest convenience.

I look forward to hearing from you.

Yours sincerely,

Elizabeth J. Kay
Lecturer in Community Dental Health
PhD Student

I would/would not (delete as necessary) be happy for you to come to my surgery to discuss bitewing radiography. Please telephone me to arrange a suitable time.

Name .... Address.......................

....

.... Phone No. ....

Please return this slip to me in the attached pre-paid envelope.

Thank you.

APPENDIX II

Rating Scale Data Collection Form

Please comment on the quality of this radiograph compared with those you normally use. ____________________

____

____

RADIOGRAPH No. (right)

Please score the approximal surfaces 4 distal to 7 mesial for this radiograph, according to the criteria below:

Score 1 if you would <u>definitely</u> place a restoration in the tooth due to a lesion in the surface in question.

Score 2 if you would <u>probably</u> place a restoration in the tooth due to a lesion in the surface in question.

Score 3 if you <u>might consider</u> placing a restoration in the tooth due to a lesion in the surface in question.

Score 4 if you <u>might consider</u> leaving the tooth unrestored.

Score 5 if you would <u>probably</u> leave the tooth unrestored.

Score 6 if you would <u>definitely</u> leave the tooth unrestored.

	SCORE		SCORE
Upper 4 distal		Lower 4 distal	
Upper 5 mesial		Lower 5 mesial	
Upper 5 distal		Lower 5 distal	
Upper 6 mesial		Lower 6 mesial	
Upper 6 distal		Lower 6 distal	
Upper 7 mesial		Lower 7 mesial	

APPENDIX III

Questionnaire concerning depths of lesions treated

In a sixteen year old, with moderate caries experience, whom you are likely to see again within a year, do you think it is most appropriate to restore approximal caries lesions when:-

(Please tick the most appropriate response)

(a) the lesion has penetrated up to half of the enamel ☐

(b) the lesion has penetrated into the deeper half of the enamel but has not reached the amelo-dentinal junction ☑

(c) the lesion has reached the amelo-dentinal junction but has not penetrated dentine ☐

(d) the lesion extends into dentine ☐

(e) the lesion extends well into dentine ☐

APPENDIX IV

True positive/negative and false positive/negative treatment decisions, at each rating scale level of decision-making, and each microscopic validating criteria

Appendix IV(a) True positive/negative and false positive/negative treatment decisions, using caries to dentine as the validating criteria and rating 1 as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	12	(3.5)	291	(84.6)	9	(2.6)	32	(9.3)
2	8	(2.3)	296	(86.0)	4	(1.2)	36	(10.5)
3	17	(4.9)	285	(82.8)	15	(4.4)	27	(7.8)
4	21	(6.1)	277	(80.5)	23	(6.7)	23	(6.7)
5	14	(4.1)	292	(84.9)	8	(2.3)	30	(8.7)
6	16	(4.7)	286	(83.1)	14	(4.1)	28	(8.1)
7	6	(1.7)	297	(86.3)	3	(0.9)	38	(11.0)
8	14	(4.1)	282	(82.0)	18	(5.2)	30	(8.7)
9	9	(2.6)	288	(83.7)	12	(3.5)	35	(10.2)
10	19	(5.5)	286	(83.1)	14	(4.1)	25	(7.3)
11	2	(0.6)	296	(86.0)	4	(1.2)	42	(12.2)
12	18	(5.2)	286	(83.1)	14	(4.1)	26	(7.6)
13	4	(1.2)	298	(86.6)	2	(0.6)	40	(11.6)
14	18	(5.2)	286	(83.1)	14	(4.1)	26	(7.6)
15	12	(3.5)	284	(82.6)	16	(4.7)	32	(9.3)
16	15	(4.4)	289	(84.0)	11	(3.2)	29	(8.4)
17	23	(6.7)	264	(76.7)	36	(10.5)	21	(6.1)
18	5	(1.5)	291	(84.6)	9	(2.6)	39	(11.3)
19	16	(4.7)	291	(84.6)	9	(2.6)	28	(8.1)
20	25	(7.3)	278	(80.8)	22	(6.4)	19	(5.5)

Appendix IV(b) True positive/negative and false positive/negative treatment decisions, using caries into dentine as the validating criteria, and rating 1+2 as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	24	(7.0)	252	(73.3)	48	(14.0)	20	(5.8)
2	8	(2.3)	295	(85.8)	5	(10.5)	36	(10.5)
3	21	(6.1)	278	(80.8)	22	(6.4)	23	(6.7)
4	24	(7.0)	267	(77.6)	33	(9.6)	20	(5.8)
5	22	(6.4)	267	(77.6)	33	(9.6)	22	(6.4)
6	18	(5.2)	283	(82.3)	17	(4.9)	26	(7.6)
7	9	(2.6)	293	(85.2)	7	(2.0)	35	(10.2)
8	17	(4.9)	273	(79.4)	27	(7.8)	27	(7.8)
9	13	(4.1)	281	(81.7)	19	(5.5)	30	(8.7)
10	22	(6.4)	278	(80.8)	22	(6.4)	22	(6.4)
11	15	(4.4)	284	(82.6)	16	(4.7)	29	(8.4)
12	20	(5.8)	284	(82.6)	16	(4.7)	24	(7.0)
13	6	(1.7)	295	(85.8)	5	(1.5)	38	(11.0)
14	24	(7.0)	270	(78.5)	30	(8.7)	20	(5.8)
15	23	(6.7)	270	(78.5)	30	(8.7)	21	(6.1)
16	18	(5.2)	287	(83.4)	13	(3.8)	26	(7.6)
17	27	(7.8)	257	(74.7)	43	(12.5)	17	(4.9)
18	10	(2.9)	278	(80.8)	22	(6.4)	34	(9.9)
19	19	(5.5)	279	(81.1)	21	(61.1)	25	(7.3)
20	26	(7.6)	268	(77.9)	32	(9.3)	18	(5.2)

Appendix IV(c) True positive/negative and false positive/negative treatment decisions, using caries into dentine as the validating criteria, and rating 1,2,+3, as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	33	(9.6)	177	(51.5)	123	(35.8)	11	(3.2)
2	8	(2.3)	293	(85.2)	7	(2.0)	36	(10.5)
3	27	(7.8)	260	(75.6)	40	(11.6)	17	(4.9)
4	30	(8.7)	233	(67.7)	67	(19.5)	14	(4.1)
5	28	(8.1)	228	(66.3)	72	(20.9)	16	(4.7)
6	19	(5.5)	282	(82.0)	18	(5.2)	25	(7.3)
7	13	(3.8)	287	(83.4)	13	(3.8)	31	(9.0)
8	24	(7.0)	242	(70.3)	58	(16.9)	20	(5.8)
9	17	(4.9)	266	(77.3)	34	(9.9)	27	(7.8)
10	26	(7.6)	264	(76.7)	36	(10.5)	18	(5.2)
11	25	(7.3)	263	(76.5)	37	(10.8)	19	(5.5)
12	26	(7.6)	261	(75.9)	39	(11.3)	18	(12.8)
13	15	(4.4)	284	(82.6)	16	(4.7)	29	(8.4)
14	26	(7.6)	247	(71.8)	53	(15.4)	18	(5.2)
15	29	(8.4)	241	(70.1)	59	(17.2)	15	(4.4)
16	23	(6.7)	279	(81.1)	21	(6.1)	21	(6.1)
17	33	(9.6)	221	(64.2)	79	(23.0)	11	(3.2)
18	19	(5.5)	246	(71.5)	54	(15.7)	25	(7.3)
19	24	(7.0)	273	(79.4)	27	(7.8)	20	(5.8)
20	28	(8.1)	267	(77.6)	33	(9.6)	16	(4.7)

Appendix IV(d) True positive/negative and false positive/negative treatment decisions, using caries into dentine as the validating criteria, and rating 1,2,3,+4 as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	38	11.0	100	29.1	200	58.1	6	1.7
2	10	2.9	290	84.3	10	2.9	34	9.9
3	28	8.1	254	73.8	46	13.4	16	4.7
4	36	10.5	207	60.2	93	27.0	8	2.3
5	31	9.0	211	61.3	89	25.9	13	3.8
6	20	5.8	281	81.7	19	5.5	24	7.0
7	15	4.4	280	81.4	20	5.8	29	8.4
8	26	7.6	227	66.0	73	21.2	18	5.2
9	17	4.9	266	77.3	34	9.9	27	7.8
10	27	7.8	262	76.2	38	11.0	17	4.9
11	39	11.3	83	24.1	217	63.1	5	1.5
12	27	7.8	260	75.6	40	11.6	17	4.9
13	16	4.7	283	82.3	17	4.9	28	8.1
14	30	8.7	223	64.8	77	22.4	14	4.1
15	30	8.7	219	63.7	81	23.5	14	4.1
16	23	6.7	278	80.8	22	6.4	21	6.1
17	33	9.6	215	62.5	85	24.7	11	3.2
18	27	7.8	181	52.6	119	34.6	17	4.9
19	24	7.0	273	79.4	27	7.8	20	5.8
20	28	8.1	266	77.3	34	9.9	16	4.7

Appendix IV(e) True positive/negative and false positive/negative treatment decisions, using caries into dentine as the validating criteria, and rating 1,2,3,+4 and 5 as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	44	12.8	4	1.2	296	86.0	0	0
2	11	3.2	290	84.3	10	2.9	33	12.8
3	30	8.7	246	71.5	54	15.7	14	4.1
4	40	11.6	137	39.8	163	47.4	4	1.2
5	38	11.0	156	45.3	144	41.9	6	1.7
6	20	5.8	280	81.4	20	5.8	24	7.0
7	16	4.7	271	78.8	29	8.4	28	8.1
8	27	7.8	223	64.8	77	22.4	17	4.9
9	17	4.9	265	77.0	35	10.2	27	7.8
10	29	8.4	253	73.5	47	13.7	15	4.4
11	44	12.8	0	0	300	87.2	0	0
12	27	7.8	260	75.6	40	11.6	17	4.9
13	19	5.5	272	79.1	28	8.1	25	7.3
14	32	9.3	170	49.4	130	37.8	12	3.5
15	34	9.9	191	55.5	109	31.7	10	2.9
16	24	7.0	277	80.5	23	6.7	20	5.8
17	35	10.2	208	60.5	92	26.7	9	2.6
18	44	12.8	1	0.3	299	86.9	0	0
19	29	8.4	262	76.2	38	11.0	15	4.4
20	265	77.0	28	8.1	35	10.2	16	4.7

Appendix IV(f) True positive/negative and false positive/negative treatment decisions, using caries into inner dentine as the validating criteria, and with rating 1 as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	13	3.8	263	76.5	8	2.3	60	17.4
2	8	2.3	267	77.6	4	1.2	65	18.9
3	19	5.5	258	75.0	13	3.8	54	15.7
4	24	7.0	251	73.0	20	5.8	49	14.2
5	14	4.1	263	76.5	8	2.3	59	17.2
6	18	5.2	259	75.3	12	3.5	55	16.0
7	6	1.7	268	77.9	3	0.9	67	19.5
8	16	4.7	255	74.1	16	4.7	57	16.6
9	9	2.6	259	75.3	12	3.5	73	21.2
10	20	5.8	258	75.0	13	3.8	53	15.3
11	2	0.6	267	77.6	4	1.2	71	20.6
12	20	5.8	259	75.3	12	3.5	53	15.4
13	4	1.2	269	78.2	2	0.6	69	20.1
14	20	5.8	259	75.3	12	3.5	53	15.4
15	13	3.8	256	74.4	15	64.4	60	17.4
16	17	4.9	262	76.2	9	2.6	56	16.3
17	27	7.8	239	69.5	32	9.3	46	13.4
18	6	1.7	263	76.5	8	2.3	67	19.5
19	16	4.7	262	76.2	9	2.6	57	16.6
20	29	8.4	253	73.5	18	5.2	44	12.8

Appendix IV(g) True positive/negative and false positive/negative treatment decisions, using caries into inner enamel but not into dentine as the validating criteria with rating 1+2 as decisions to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	32	9.3	231	67.2	40	11.6	41	11.9
2	8	2.3	266	77.3	5	1.5	65	18.9
3	25	7.3	253	73.5	18	5.2	48	14.0
4	32	9.3	246	71.5	25	7.3	41	11.9
5	25	7.3	241	70.1	30	8.7	48	14.0
6	21	6.1	257	74.7	14	4.1	52	15.1
7	9	2.6	264	76.7	7	2.0	64	18.6
8	21	6.1	248	72.1	23	6.7	52	15.1
9	17	4.9	255	74.1	16	4.7	56	16.3
10	25	7.3	252	73.3	19	5.5	48	14.0
11	17	4.9	257	74.7	14	4.1	56	16.3
12	23	6.7	258	75.0	13	3.8	50	14.5
13	6	1.7	266	77.3	5	1.5	67	19.5
14	28	8.1	245	71.2	26	7.6	45	13.1
15	26	7.6	244	70.9	27	7.8	47	13.7
16	20	5.8	260	75.6	1	3.2	53	15.4
17	32	9.3	233	67.7	38	11.0	41	11.9
18	12	3.5	251	73.0	20	5.8	61	17.7
19	24	7.0	255	74.1	16	4.7	49	14.2
20	32	9.3	245	71.2	26	7.6	41	11.9

Appendix IV(h) True positive/negative and false positive/negative treatment decisions, using caries into inner enamel but not into dentine as the validating criteria, and with rating 1+2+3 as decisions to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	48	14.0	163	47.4	108	31.4	25	7.3
2	9	2.6	265	77.0	6	1.7	64	18.6
3	36	10.5	240	69.8	31	9.0	37	10.8
4	40	11.6	214	62.2	57	16.6	33	9.6
5	36	10.5	207	60.2	64	18.6	37	10.8
6	22	6.4	256	74.4	15	4.4	51	14.8
7	14	4.1	259	75.3	12	3.5	59	17.2
8	35	10.2	224	65.1	47	13.7	38	11.0
9	23	6.7	243	70.6	28	8.1	50	14.5
10	34	9.9	243	70.6	28	8.1	39	11.3
11	31	9.0	240	69.8	31	9.0	42	12.2
12	32	9.3	238	69.2	33	9.6	41	11.9
13	17	4.9	257	74.4	14	4.1	56	16.3
14	33	9.6	225	65.4	46	13.6	40	11.6
15	39	11.3	222	64.5	49	14.2	34	9.9
16	28	8.1	255	74.1	16	4.7	45	13.1
17	44	12.8	203	59.0	68	19.8	29	3.4
18	27	7.8	225	65.4	46	13.4	46	13.4
19	30	8.7	250	72.7	21	6.1	43	12.5
20	34	9.9	244	70.9	27	7.8	39	11.3

Appendix IV(i) True positive/negative and false positive/negative treatment decision, using caries into inner enamel but not into dentine as the validating criteria, and with rating 1+2+3+4 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	% of total	N	% of total	N	% of total	N	% of total
1	62	18.0	95	27.6	176	51.2	11	3.2
2	11	3.2	262	76.2	9	2.6	62	18.0
3	38	11.0	235	68.3	36	10.5	35	10.2
4	49	14.2	191	55.5	80	23.3	24	7.0
5	43	12.5	194	56.4	77	22.4	30	8.7
6	23	6.7	255	74.1	16	4.7	50	14.5
7	17	4.9	253	73.1	18	5.2	56	16.3
8	37	10.8	209	60.8	62	18.0	36	10.5
9	23	6.7	243	70.6	28	8.1	50	14.5
10	36	10.5	242	70.3	29	8.4	37	10.8
11	63	18.3	78	22.7	193	56.1	10	2.9
12	33	9.6	237	68.9	34	9.9	40	11.6
13	18	5.2	256	74.4	15	4.4	55	16.0
14	40	11.6	204	59.3	67	19.5	33	9.6
15	46	13.4	206	59.9	65	18.9	27	7.8
16	28	8.1	254	73.8	17	4.9	45	13.1
17	46	13.4	199	57.8	72	20.9	27	7.8
18	41	11.9	166	48.3	105	30.5	32	9.3
19	30	8.7	250	72.7	21	6.1	43	12.5
20	73	21.2	4	1.2	267	77.6	0	0

Appendix IV(j) True positive/negative and false positive/negative treatment decision, using caries into inner enamel, but not dentine as the validating criteria, and with rating 1+2+3+4 +5 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	73	21.2	4	1.2	267	77	0	0
2	12	3.5	262	76.2	9	2.6	61	17.7
3	42	12.2	229	66.6	42	12.2	31	9.0
4	58	16.9	126	36.6	145	42.2	15	4.4
5	57	16.6	146	42.4	125	36.3	16	4.7
6	23	6.7	254	73.8	17	4.0	50	14.5
7	20	5.8	246	71.5	25	7.3	53	15.4
8	38	11.0	205	59.6	66	19.2	35	10.2
9	23	6.7	242	70.3	29	8.4	50	14.5
10	39	11.3	234	68.0	37	10.8	34	9.9
11	73	21.2	0	0	271	78.8	0	0
12	33	9.6	237	68.9	34	9.9	40	11.6
13	24	7.0	248	72.1	23	6.7	49	14.2
14	47	13.7	156	45.3	115	33.4	26	7.6
15	51	14.8	179	520	92	26.7	22	6.4
16	29	8.4	253	73.5	18	5.2	44	12.8
17	52	15.1	196	57.0	75	21.8	21	6.1
18	73	21.2	1	0.3	270	78.0	0	0
19	36	10.5	240	69.8	31	9.0	37	10.8
20	36	10.4	244	70.9	27	7.8	37	10.8

Appendix IV(k) True positive/negative and false positive/negative treatment decisions, using any enamel/dentine lesion as the validating criteria, and with rating 1 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	15	4.4	242	703	6	1.7	81	23.5
2	8	2.3	244	70.9	4	1.2	88	25.6
3	21	6.1	237	68.9	11	3.2	75	21.8
4	26	7.6	230	66.9	18	5.2	70	20.3
5	14	4.1	240	69.8	8	2.3	82	23.8
6	20	5.8	238	69.2	10	2.9	76	22.1
7	6	1.7	245	71.2	3	0.9	90	26.2
8	19	5.5	235	68.3	13	3.8	77	22.4
9	11	3.2	238	69.2	10	2.9	85	24.7
10	22	6.4	237	68.9	11	3.2	74	21.5
11	4	1.2	244	70.9	4	1.2	94	27.3
12	22	6.4	238	69.2	10	2.9	74	21.5
13	5	1.5	247	71.8	1	0.3	91	26.5
14	20	5.8	236	68.6	12	3.5	76	22.1
15	15	4.4	235	68.3	13	3.8	81	23.5
16	19	5.5	241	70.1	7	2.0	77	22.4
17	29	8.4	218	63.4	30	8.7	67	19.5
18	7	2.0	241	70.1	7	2.0	89	25.9
19	18	5.2	241	70.1	7	2.0	78	22.7
20	33	9.6	234	68.0	14	4.1	63	18.3

Appendix IV(I) True positive/negative and false positive/negative treatment decisions, using any enamel/dentine lesion as the validating criteria and with ratings 1+2 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	37	10.8	213	61.9	35	10.2	59	17.2
2	8	2.3	243	70.3	5	1.5	88	25.6
3	28	8.1	233	67.7	15	4.4	68	19.8
4	34	9.9	225	65.4	23	6.7	62	18.0
5	27	7.8	220	64.0	28	8.1	69	20.1
6	24	7.0	237	68.9	11	3.2	72	20.9
7	10	2.9	242	70.3	6	1.7	86	25.0
8	24	7.0	228	66.3	20	5.8	72	20.9
9	19	5.5	234	68.0	14	4.1	77	22.4
10	28	8.1	232	67.4	16	4.7	68	19.8
11	20	5.8	237	68.9	11	3.2	76	22.1
12	25	7.3	237	68.9	11	3.2	71	20.6
13	8	2.3	245	71.2	3	0.9	88	25.6
14	31	9.0	225	65.4	23	6.7	65	18.9
15	30	8.7	225	65.4	23	6.7	66	19.2
16	22	6.4	239	69.5	9	2.6	74	21.5
17	35	10.2	213	61.9	35	10.2	61	17.7
18	14	4.1	230	66.9	18	5.2	82	23.8
19	27	7.8	235	68.3	13	3.8	69	20.1
20	36	10.5	226	65.7	22	6.4	60	17.4

Appendix IV(m) True positive/negative and false positive/negative treatment decisions, using any enamel/dentine lesion as the validating criteria, and with ratings 1+2+3 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	60	17.4	152	44.2	96	27.9	36	10.5
2	10	2.9	243	70.3	5	1.5	86	25.0
3	39	11.3	220	64.0	28	8.1	57	16.6
4	47	13.7	198	57.6	50	14.5	49	14.2
5	43	12.5	191	55.5	57	16.6	53	15.4
6	25	7.3	236	68.6	12	3.5	71	20.6
7	15	4.4	237	68.9	11	3.2	81	23.5
8	40	11.6	206	59.9	42	12.2	56	16.3
9	26	7.6	223	64.8	25	7.3	70	20.3
10	38	11.0	224	65.1	24	7.0	58	16.9
11	34	9.9	220	64.0	28	8.1	62	18.0
12	35	10.2	218	63.4	30	8.7	61	17.7
13	20	5.8	237	68.9	11	3.2	76	22.1
14	37	10.8	206	59.9	42	12.2	59	17.9'
15	44	12.8	204	59.3	44	12.8	52	15.1
16	30	8.7	234	68.0	14	4.1	66	19.2
17	52	15.1	188	54.7	60	17.4	44	12.8
18	34	9.9	209	60.8	39	11.3	62	18.0
19	33	9.6	230	66.9	18	5.2	63	18.3
20	38	11.0	225	65.4	23	6.7	58	16.9

Appendix IV(n) True positive/negative and false positive/negative treatment decisions, using any enamel/dentine lesion as the validating criteria, and with ratings 1+2+3+4 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	79	23.0	89	25.9	159	46.2	17	4.9
2	13	3.8	241	70.1	7	2.0	83	24.1
3	41	11.9	215	62.5	33	9.6	55	16.0
4	59	17.2	178	51.7	70	20.3	37	10.8
5	51	14.8	179	52.0	69	20.1	45	13.1
6	26	7.6	235	68.3	13	3.8	70	20.3
7	19	5.5	232	67.4	16	4.7	77	22.4
8	44	12.8	193	56.1	55	16.0	52	15.1
9	26	7.6	223	64.8	25	7.3	70	20.3
10	40	11.6	223	64.8	25	7.3	56	16.3
11	79	23.0	71	20.6	177	51.1	17	4.9
12	36	10.5	217	63.1	31	9.0	60	17.4
13	21	6.1	236	68.6	12	3.5	75	21.8
14	46	13.4	187	54.4	61	17.7	50	14.5
15	53	15.4	190	55.2	58	16.9	43	12.5
16	30	8.7	233	67.7	15	4.4	66	19.2
17	54	15.7	184	53.5	64	18.6	42	12.2
18	52	15.1	154	44.8	94	27.3	44	12.8
19	33	9.6	230	66.9	18	5.2	63	18.3
20	39	11.3	225	65.4	23	6.7	57	16.6

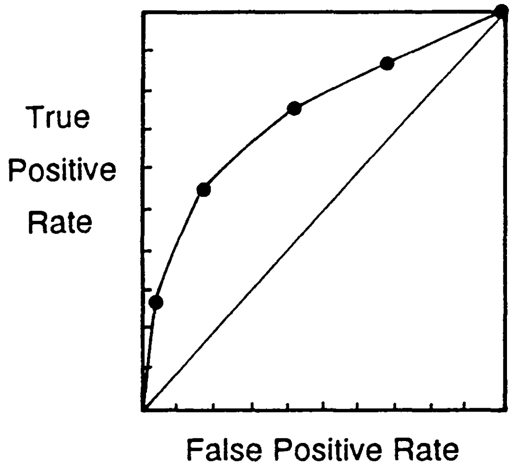
Appendix IV(o) True positive/negative and false positive/negative treatment decisions using any enamel/dentine lesion as the validating criteria, and with ratings 1+2+3+4+5 taken as a decision to restore

	TRUE POSITIVE		TRUE NEGATIVE		FALSE POSITIVE		FALSE NEGATIVE	
Dentist	N	%	N	%	N	%	N	%
1	96	27.9	4	1.2	244	70.9	0	0
2	14	4.1	241	70.1	7	2.0	82	23.8
3	46	13.4	210	61.0	38	11.0	50	14.5
4	74	21.5	119	34.6	129	37.5	22	6.4
5	67	19.5	133	38.7	115	33.4	29	8.4
6	26	7.6	234	68.0	14	4.1	70	20.3
7	22	6.4	225	65.4	23	6.7	74	21.5
8	45	13.1	189	54.9	59	17.2	51	14.8
9	27	7.8	223	64.8	25	7.3	69	20.1
10	44	12.8	216	62.8	32	9.3	52	15.1
11	96	27.9	-	-	248	72.1	-	-
12	36	10.5	217	63.1	31	9.0	60	17.4
13	28	8.1	229	66.6	19	5.5	68	19.8
14	55	16.0	141	41.0	107	31.1	41	11.9
15	59	17.2	164	47.7	84	24.4	37	10.8
16	31	9.0	232	67.4	16	4.7	65	18.9
17	61	17.7	182	52.9	66	19.2	35	10.2
18	96	27.9	1	0.3	247	71.8	-	-
19	39	11.3	220	64.0	28	8.1	57	16.6
20	40	11.6	225	65.4	23	6.7	56	16.3

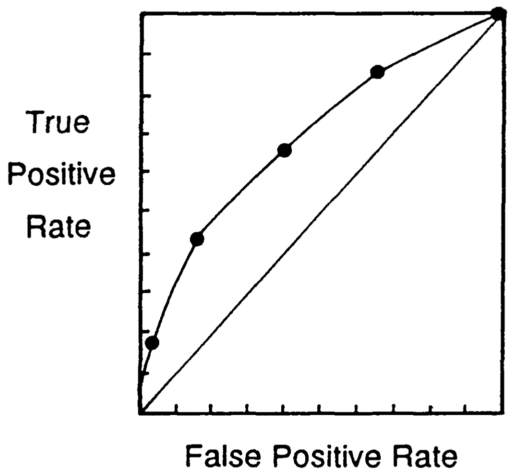
APPENDIX V

ROC curves for each dentist using three validating criteria

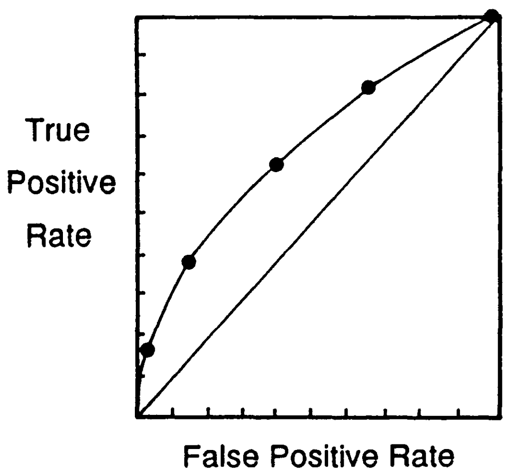
DENTIST NO: 1



Caries into dentine as the validating criterion

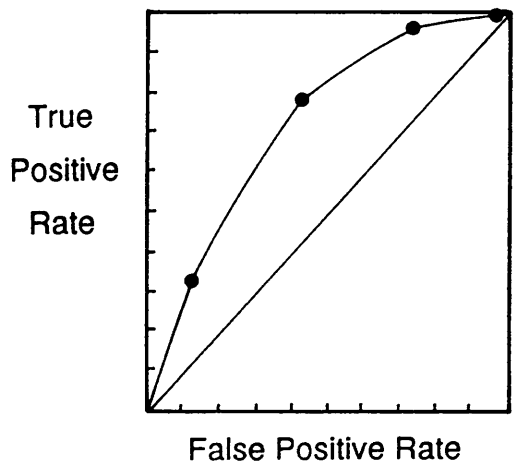


Caries into inner enamel or dentine as validating criterion

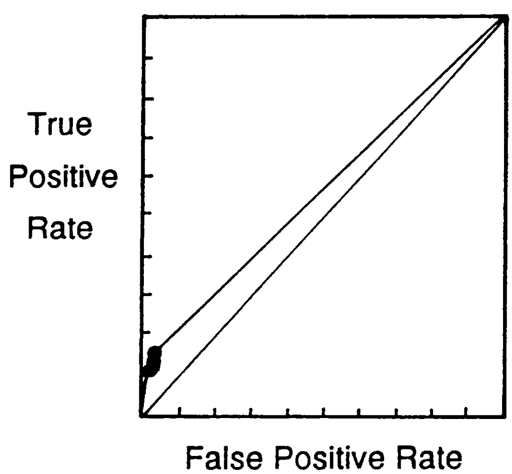


Caries in any part of enamel or dentine as validating criterion

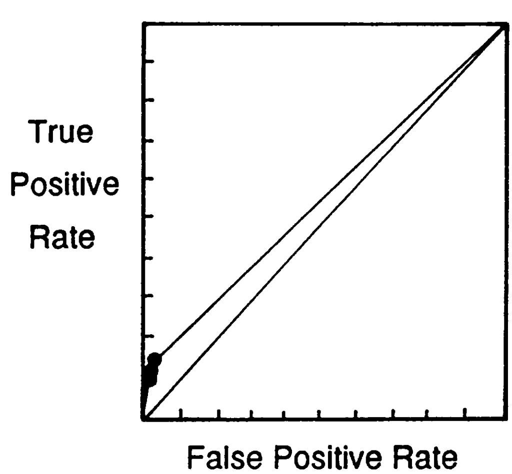
DENTIST NO: 2



Caries into dentine as the validating criterion

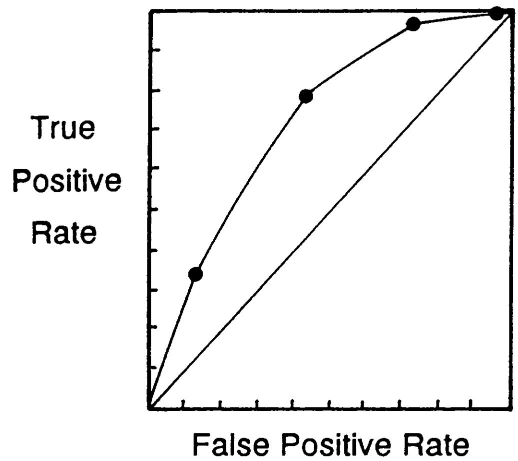


Caries into inner enamel or dentine as validating criterion

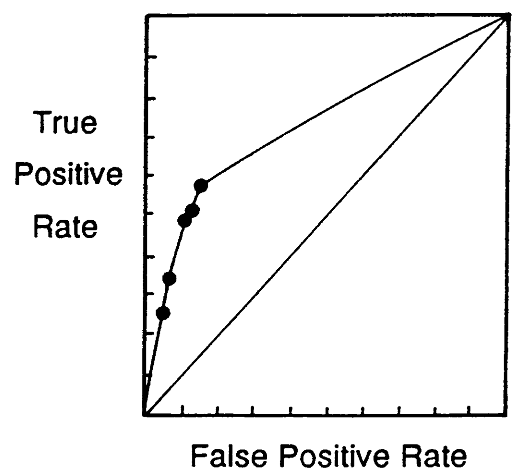


Caries in any part of enamel or dentine as validating criterion

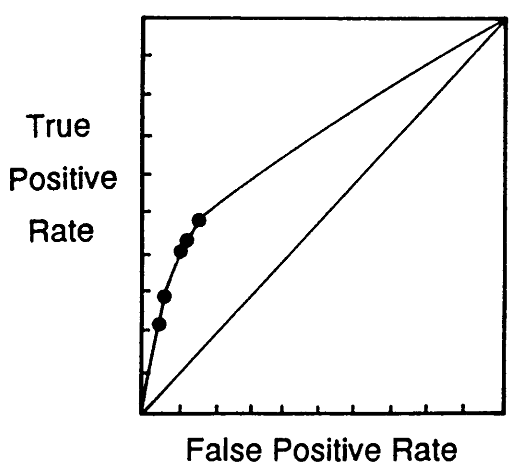
DENTIST NO: 3



Caries into dentine as the validating criterion

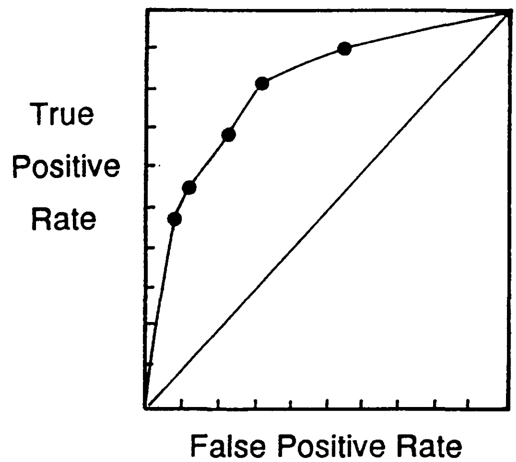


Caries into inner enamel or dentine as validating criterion

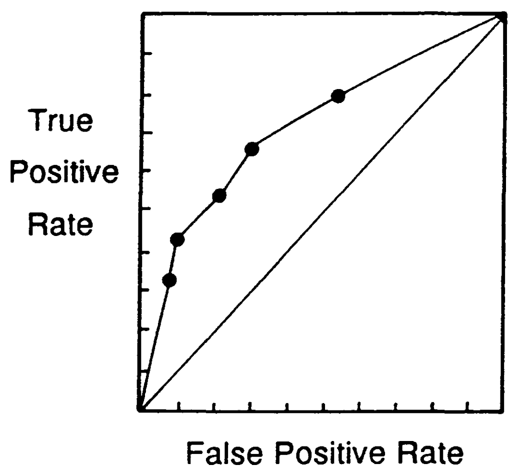


Caries in any part of enamel or dentine as validating criterion

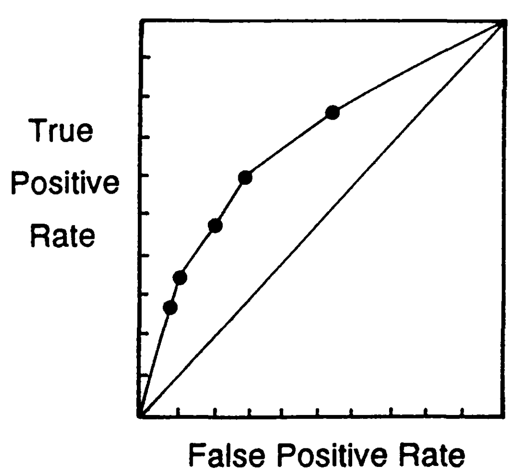
DENTIST NO: 4



Caries into dentine as the validating criterion

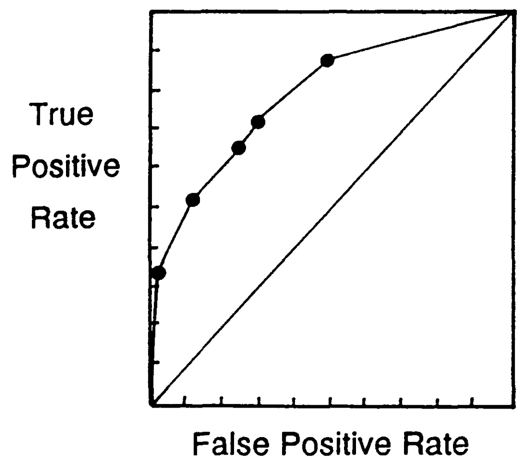


Caries into inner enamel or dentine as validating criterion

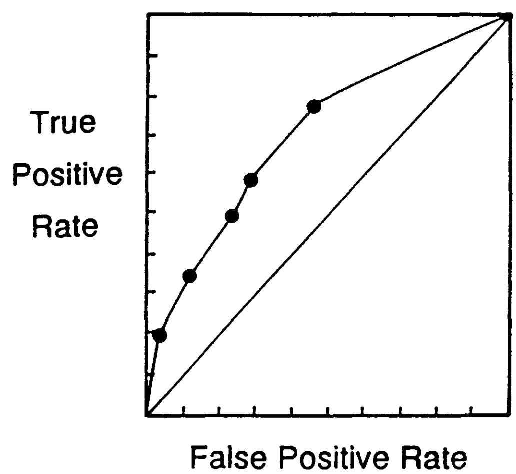


Caries in any part of enamel or dentine as validating criterion

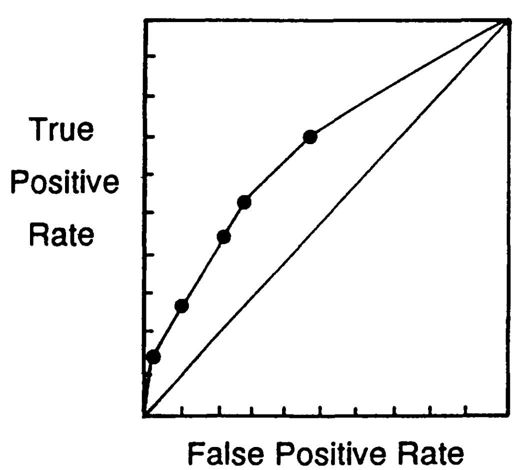
DENTIST NO: 5



Caries into dentine as the validating criterion

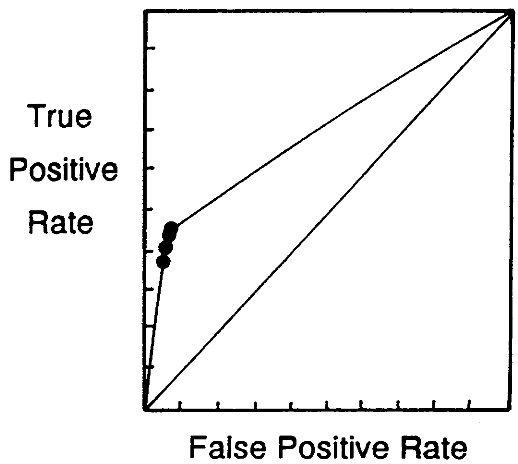


Caries into inner enamel or dentine as validating criterion

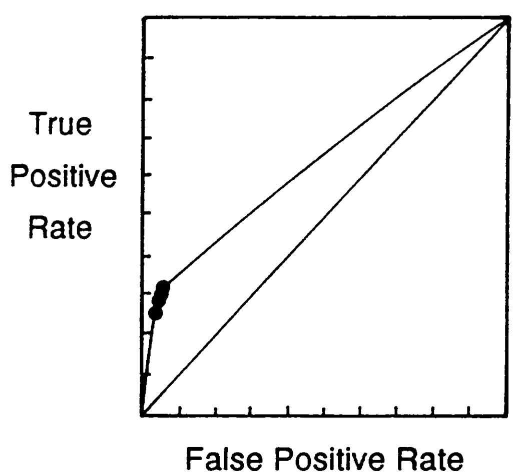


Caries in any part of enamel or dentine as validating criterion

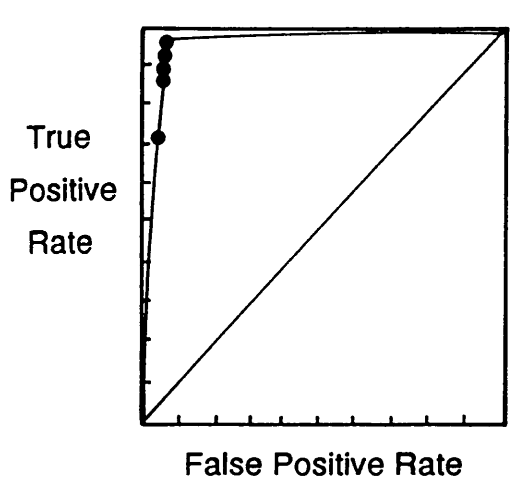
DENTIST NO: 6



Caries into dentine as the validating criterion

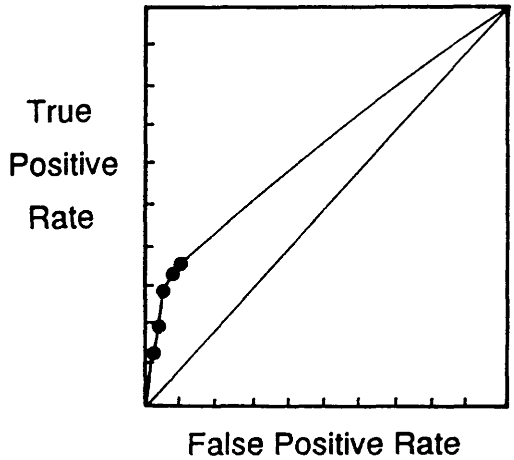


Caries into inner enamel or dentine as validating criterion

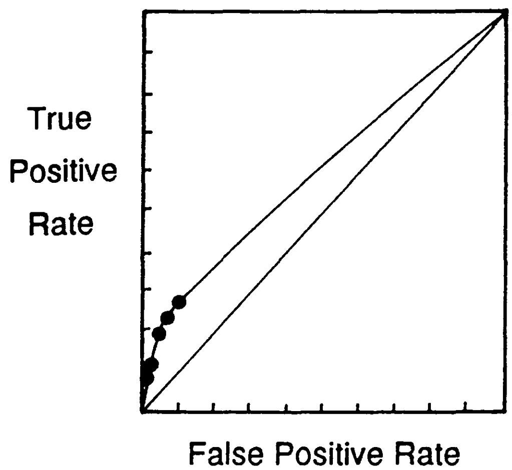


Caries in any part of enamel or dentine as validating criterion

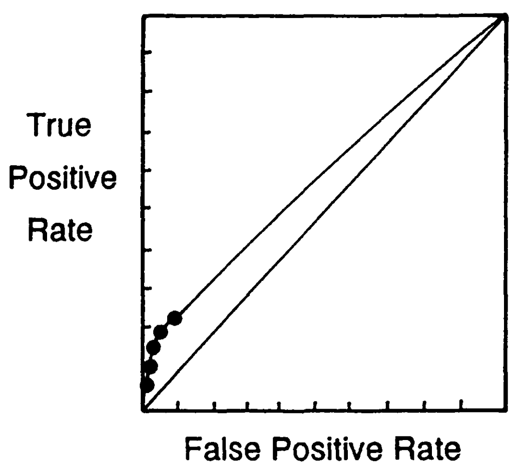
DENTIST NO: 7



Caries into dentine as the validating criterion



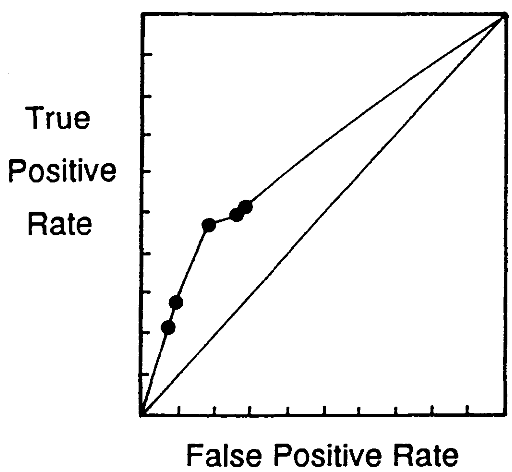
Caries into inner enamel or dentine as validating criterion



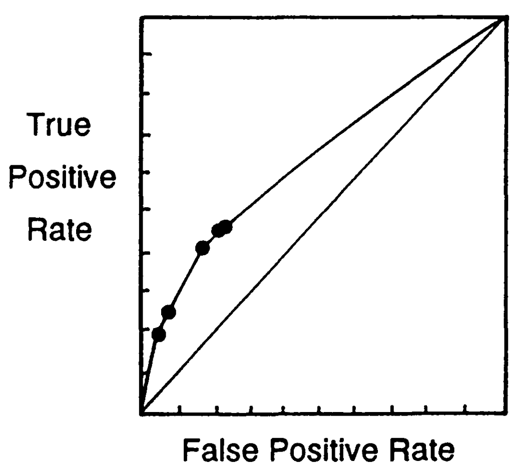
Caries in any part of enamel or dentine as validating criterion

DENTIST NO: 8

Caries into dentine as the validating criterion

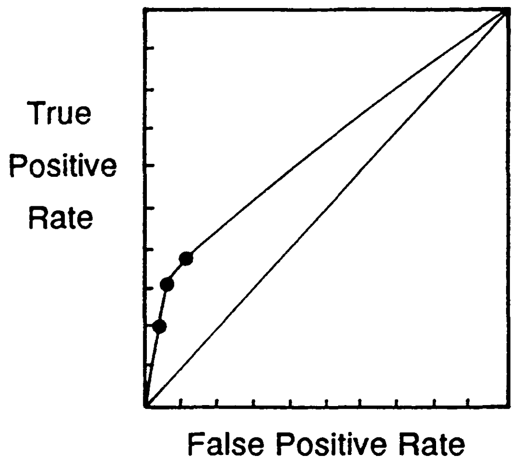


Caries into inner enamel or dentine as validating criterion

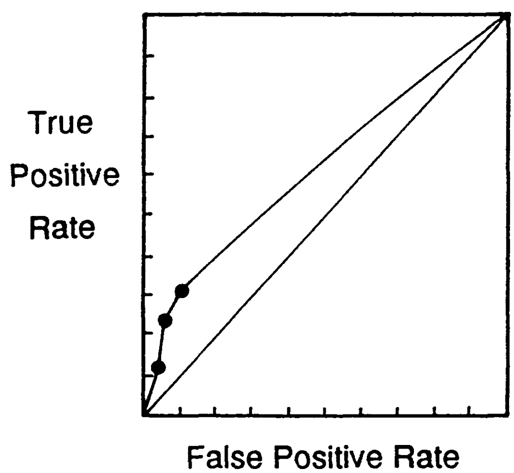


Caries in any part of enamel or dentine as validating criterion

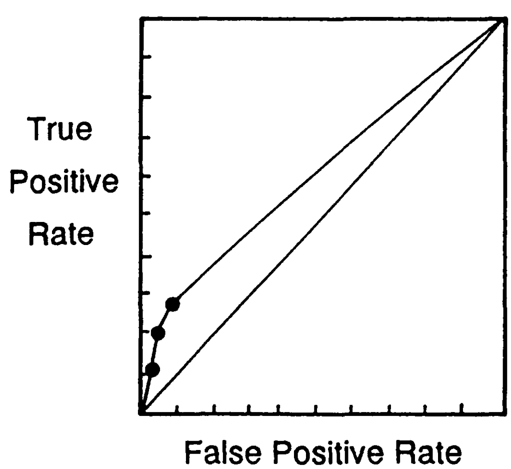
DENTIST NO: 9



Caries into dentine as the validating criterion

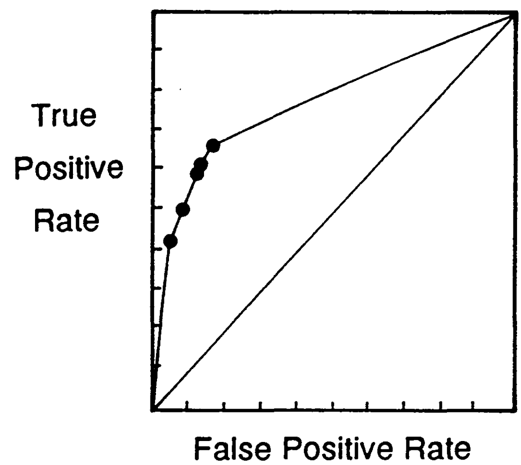


Caries into inner enamel or dentine as validating criterion

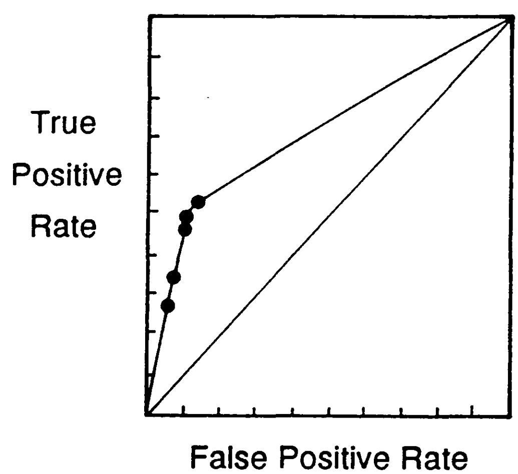


Caries in any part of enamel or dentine as validating criterion

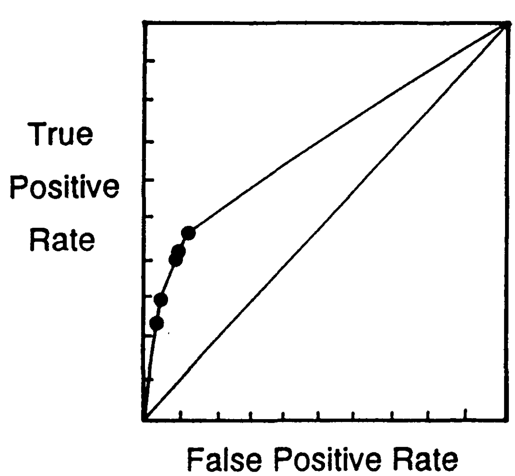
DENTIST NO: 10



Caries into dentine as the validating criterion

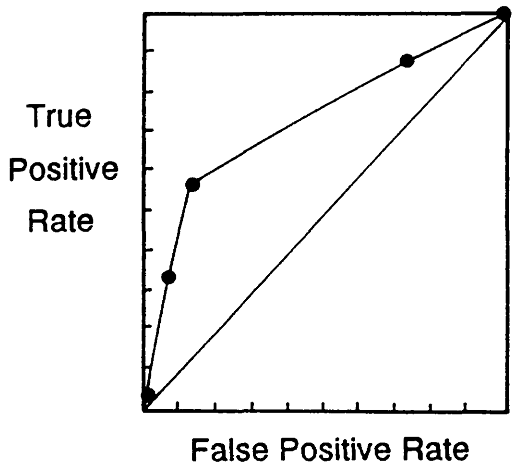


Caries into inner enamel or dentine as validating criterion

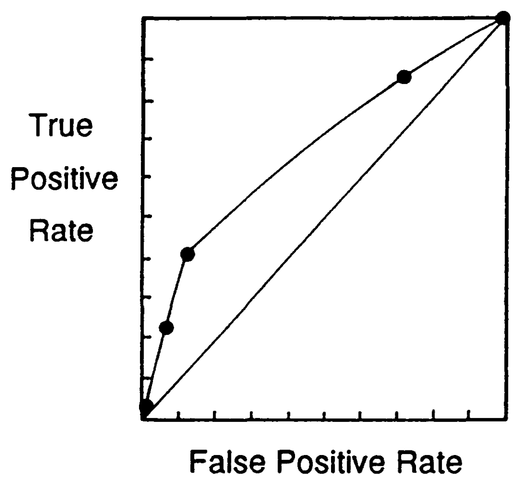


Caries in any part of enamel or dentine as validating criterion

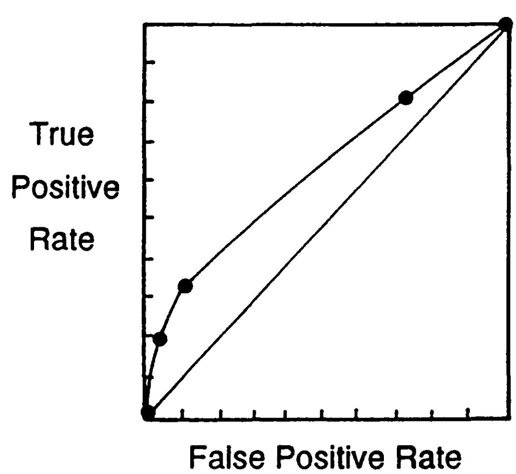
DENTIST NO: 11



Caries into dentine as the validating criterion

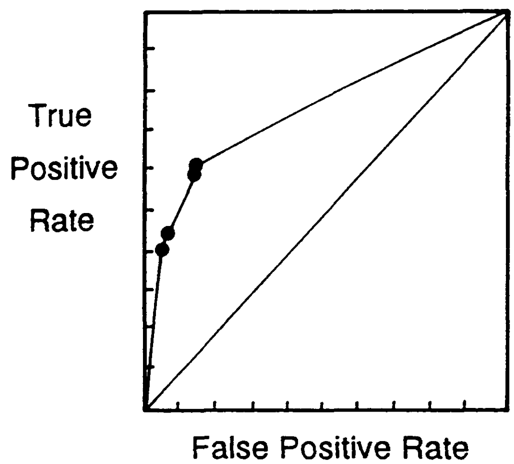


Caries into inner enamel or dentine as validating criterion

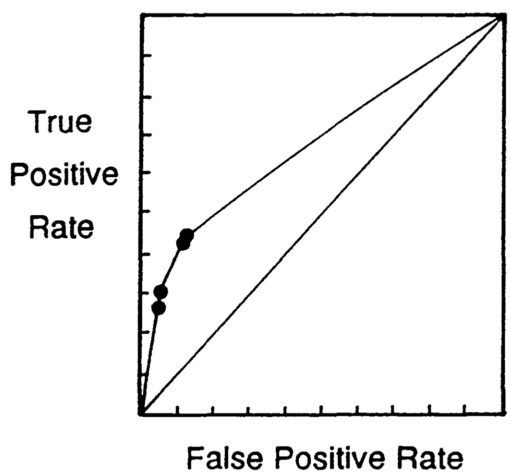


Caries in any part of enamel or dentine as validating criterion

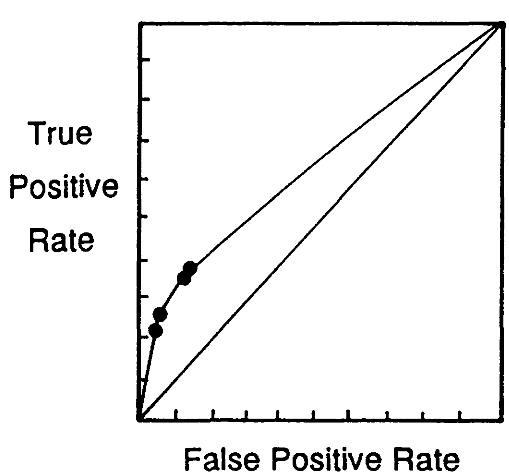
DENTIST NO: 12



Caries into dentine as the validating criterion

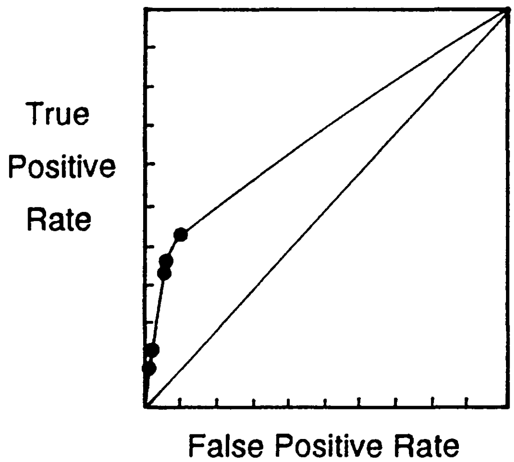


Caries into inner enamel or dentine as validating criterion

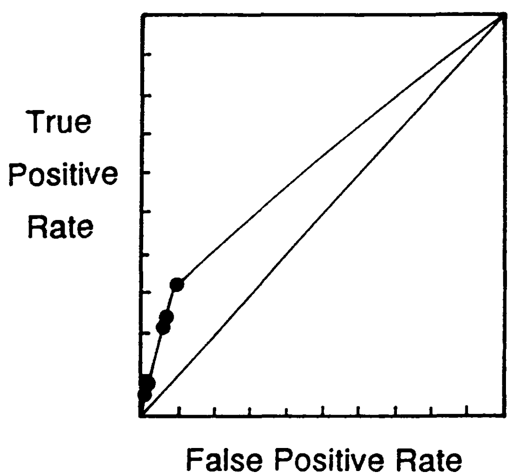


Caries in any part of enamel or dentine as validating criterion

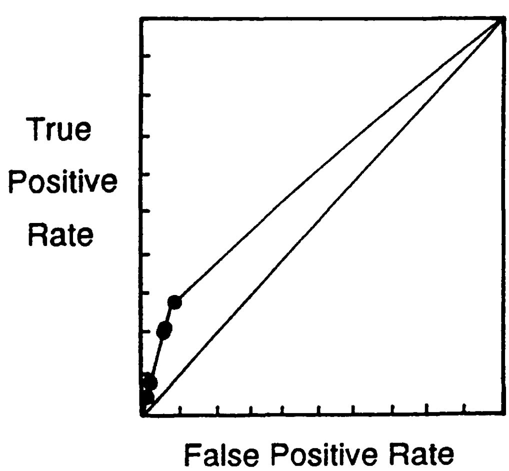
DENTIST NO: 13



Caries into dentine as the validating criterion

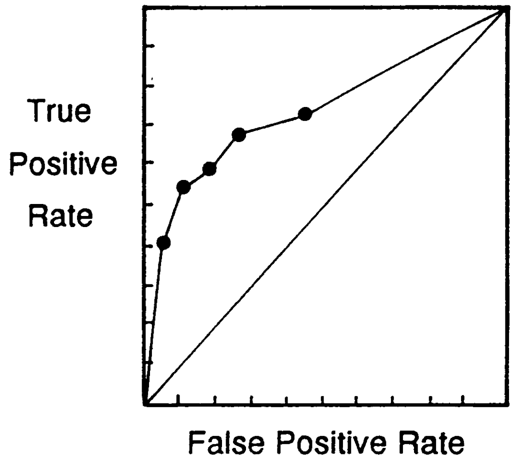


Caries into inner enamel or dentine as validating criterion

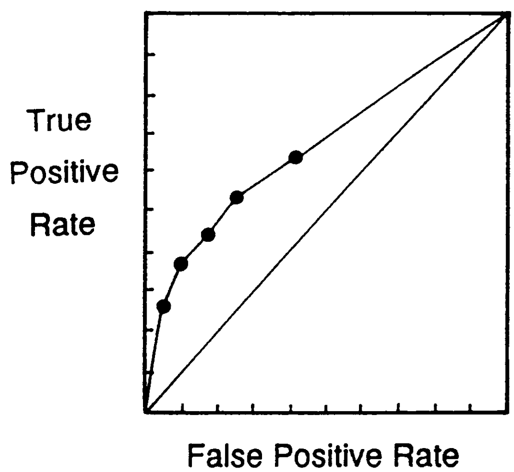


Caries in any part of enamel or dentine as validating criterion

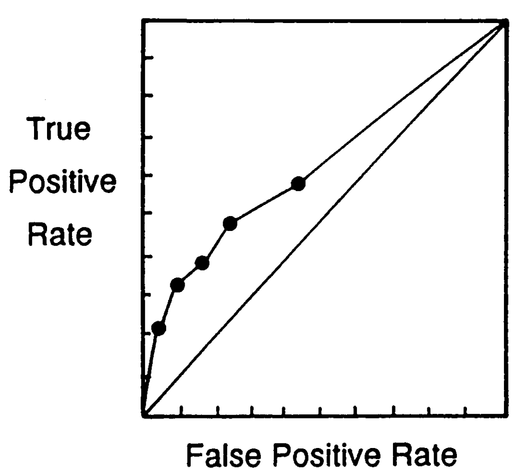
DENTIST NO: 14



Caries into dentine as the validating criterion

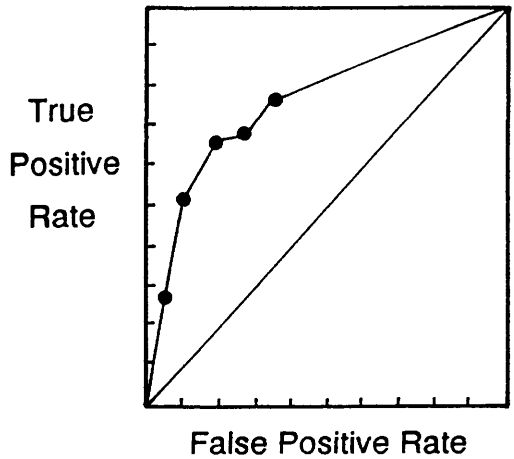


Caries into inner enamel or dentine as validating criterion

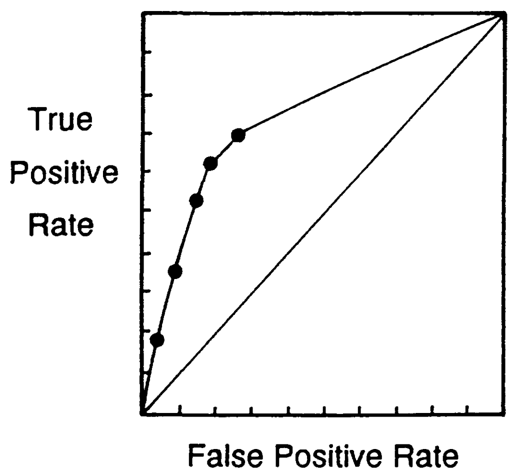


Caries in any part of enamel or dentine as validating criterion

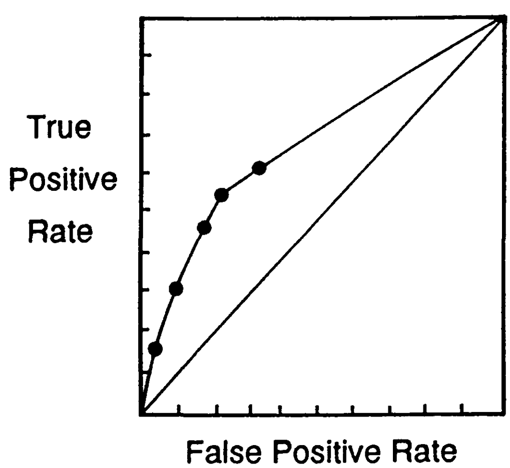
DENTIST NO: 15



Caries into dentine as the validating criterion

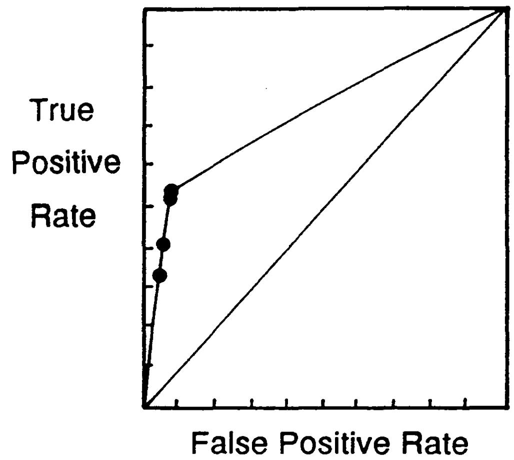


Caries into inner enamel or dentine as validating criterion

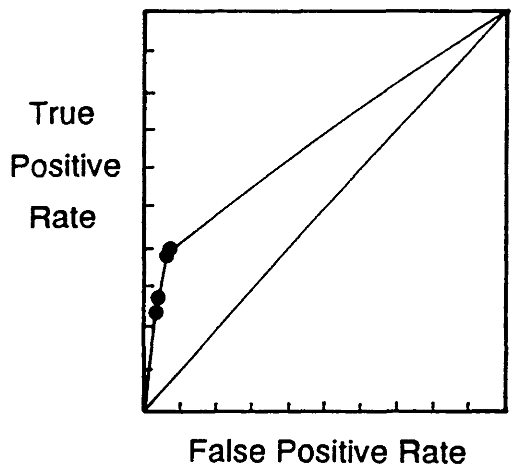


Caries in any part of enamel or dentine as validating criterion

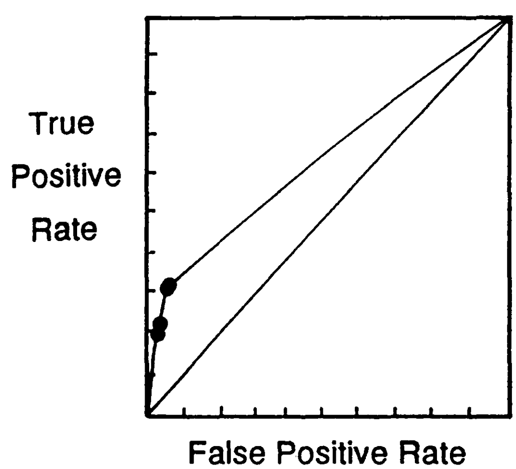
DENTIST NO: 16



Caries into dentine as the validating criterion

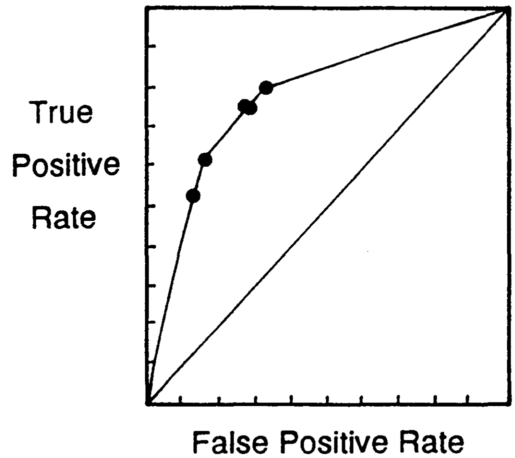


Caries into inner enamel or dentine as validating criterion

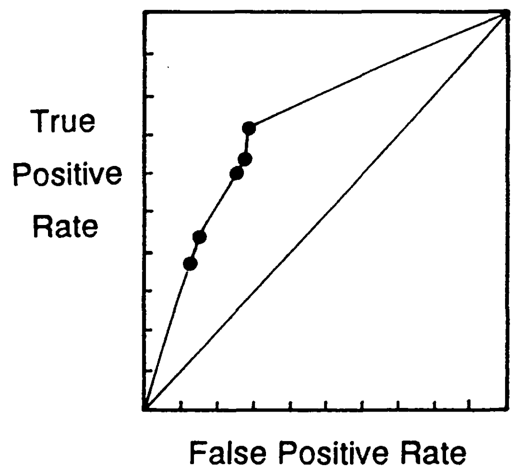


Caries in any part of enamel or dentine as validating criterion

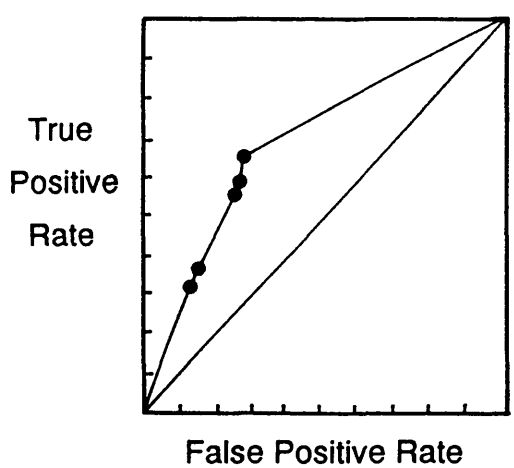
DENTIST NO: 17



Caries into dentine as the validating criterion

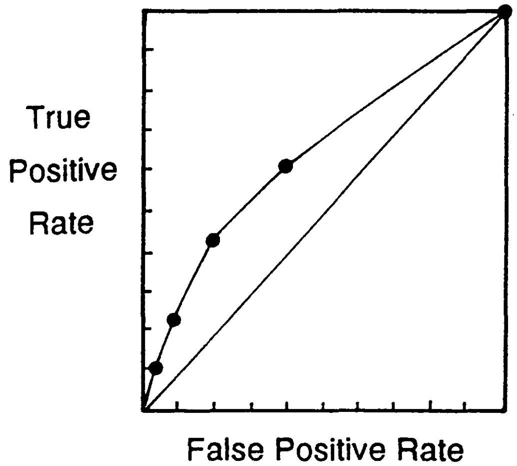


Caries into inner enamel or dentine as validating criterion

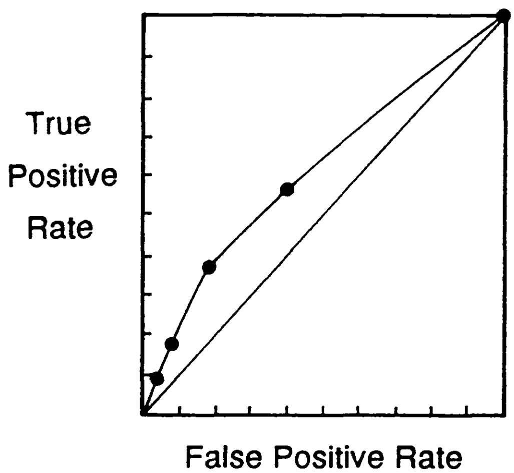


Caries in any part of enamel or dentine as validating criterion

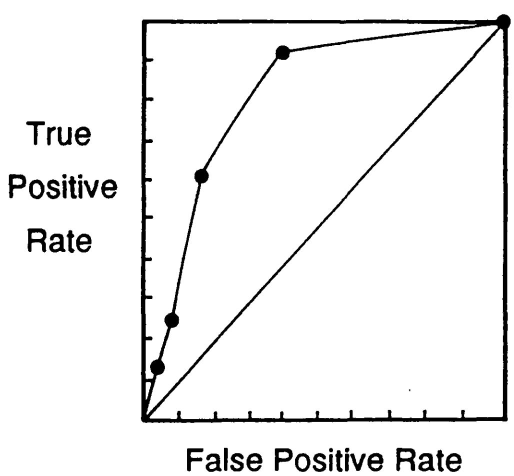
DENTIST NO: 18



Caries into dentine as the validating criterion

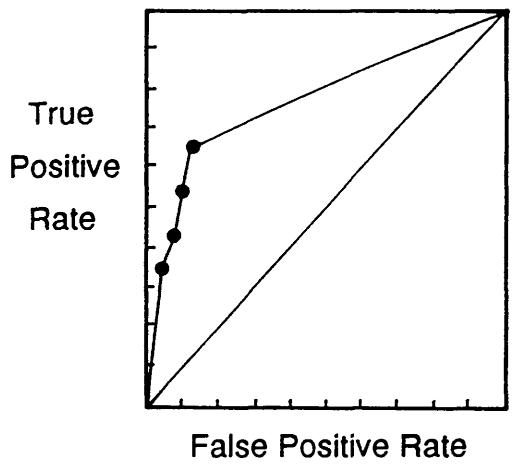


Caries into inner enamel or dentine as validating criterion

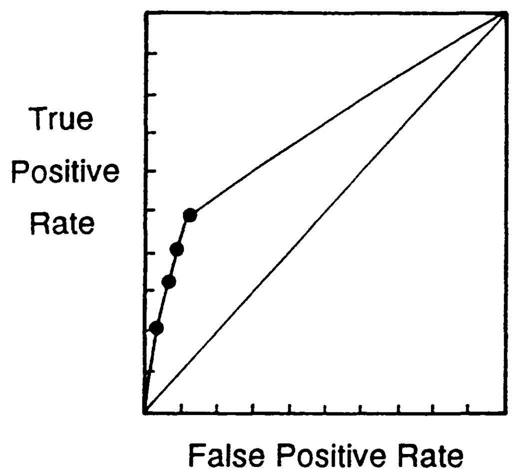


Caries in any part of enamel or dentine as validating criterion

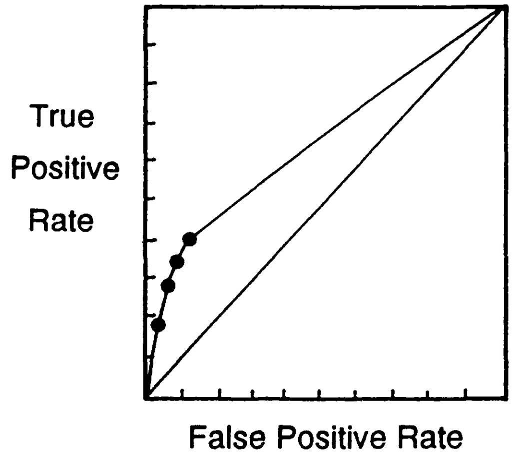
DENTIST NO: 19



Caries into dentine as the validating criterion

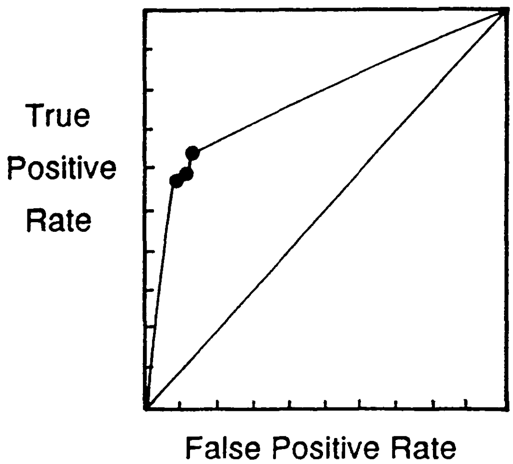


Caries into inner enamel or dentine as validating criterion

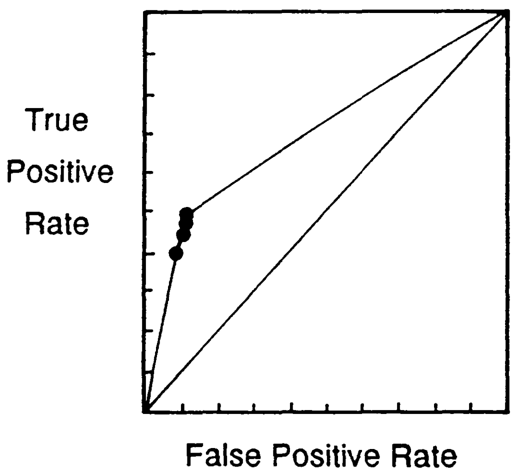


Caries in any part of enamel or dentine as validating criterion

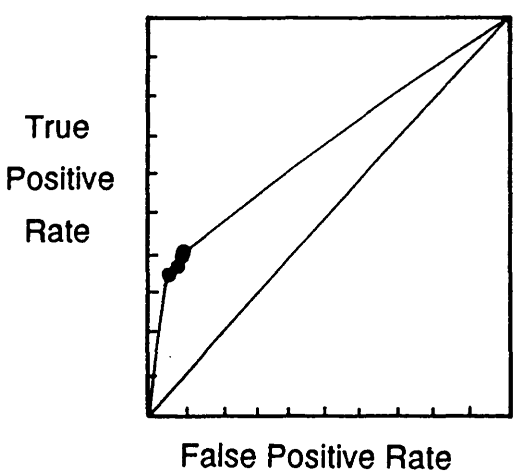
DENTIST NO: 20



Caries into dentine as the validating criterion



Caries into inner enamel or dentine as validating criterion



Caries in any part of enamel or dentine as validating criterion

APPENDIX VI

Questionnaire used for the quantitative assessment of values affecting treatment decisions

Listed below are some matters which might influence your clinical decision making. Please consider these issues then indicate how they influence your treatment planning. For each statement please tick either box A, B or C where:

In relation to this issue it is:

a) More important to <u>always</u> fill all carious teeth.

OR

b) More important to <u>never</u> unnecessarily fill sound teeth.

OR

c) Avoiding errors of any kind are of equal importance.

		A	B	C
1.	Value to my self-esteem	☐	☐	☐
2.	Patients prefer to receive/not receive treatment and are therefore more likely to become attenders if I take this course of action.	☐	☐	☐
3.	Value to my reputation with my colleagues.	☐	☐	☐
4.	Value to my patient of correct diagnosis	☐	☐	☐
5.	Value to my reputation with patients.	☐	☐	☐
6.	Value to society of correct diagnosis	☐	☐	☐
7.	Value to profession of achieving correct diagnosis.	☐	☐	☐
8.	Value to my feeling of professional responsibility.	☐	☐	☐
9.	Professional fee.	☐	☐	☐
10.	Value to my ethical conscience.	☐	☐	☐

APPENDIX VII

Visual analogue questionnaire used in the assessment of dental health state utilities (Study 1)

Q U E S T I O N N A I R E

We are trying to qualify what views members of the public hold on various states of dental health. We would be very grateful for your help by filling out this questionnaire.

1. AGE (Tick where appropriate)

18-25	----	46-55	----
26-35	----	56-65	----
36-45	----	over 65	----

2. SEX (m/f) ----

3. OCCUPATION

4. How often do you attend your dentist. (Tick where appropriate).

   a) On a routine basis ----

   b) Only when in pain or have a problem ----

   c) Other - Please specify ----

5. Please rate the value you would place on a tooth which is filled, but which will require to be refilled in 5 years

   Front tooth - Mark with an A

   Back tooth - Mark with a B

a tooth which is to be extracted ---------------------------------- A completely healthy tooth

0 5 10

6. Please rate the value which you would place on a tooth which is filled but otherwise healthy.

a) Front tooth - Mark with an A

b) Back tooth - Mark with a B

a tooth which is to be extracted | A completely healthy tooth

----

0 5 10

7. Please rate the value which you would place on a tooth which was decayed and painful.

a) Front tooth - Mark with an A

b) Back tooth - Mark with a B

a tooth which is to be extracted | A completely healthy tooth

----

0 5 10

8. Please rate the value which you would place on a tooth which was decayed, but <u>not</u> painful

a tooth which is to be extracted | A completely healthy tooth

----

0 5 10

9. What is the most important characteristic of your front teeth (Tick where appropriate)

a) For biting ____

b) How they look ____

c) General health ____

d) None ____

10. What is the most important characteristic of your back teeth (Tick where appropriate)

a) For chewing ____

b) How they look ____

c) General health ____

d) None ____

APPENDIX VIII

Standard gamble questionnaire used in assessment of dental health state utilities (Study 2)

DENTAL HEALTH SURVEY

SEX: Male ___ Female ___

AGE: 0-15 ___ 15-30 ___ 30-45 ___ 45-60 ___ 60+ ___

OCCUPATION: ....

When was the last time you actually received treatment from a dentist?

1-3 months ___

3-6 months ___

6-9 months ___

Over 1 year ___

Never ___

Would you say you find going to the dentist:

Enjoyable ___

Routine ___

Necessary ___

Bearable ___

Traumatic ___

This questionnaire will ask you to consider certain conditions of your teeth, and will ask you to choose between them. Please consider the options carefully before placing a tick in the box on the right hand side of the page.

This section deals with a choice between:

a) A certainty of having a decayed tooth, but one which will last until you die, although you may experience pain from it.

OR

b) Where you have a chance of having a completely healthy tooth for the rest of your life, but there is also a possibility that the tooth becomes so painful that must it be extracted immediately.

Indicate your preference with a tick.

A	B	Prefer A	Prefer B
100% chance of having decayed painful tooth for rest of life.	100% chance of having sound tooth for rest of life. 0% chance of immediate extraction.	____ ____	____ ____
100% chance of having decayed painful tooth for rest of life.	99% chance of having sound tooth for rest of life. 1% chance of immediate extraction.	____ ____	____ ____
100% chance of having decayed painful tooth for rest of life.	95% chance of having sound tooth for rest of life. 5% chance of immediate extraction.	____ ____	____ ____
100% chance of having decayed painful tooth for rest of life.	90% chance of having sound tooth for rest of life. 10% chance of immediate extraction.	____ ____	____ ____
100% chance of having decayed painful tooth for rest of life.	80% chance of having sound tooth for rest of life. 20% chance of immediate extraction.	____ ____	____ ____
100% chance of having decayed painful tooth for rest of life.	70% chance of having sound tooth for rest of life. 30% chance of immediate extraction.	____ ____	____ ____

A	B	Prefer A	Prefer B
100% chance of having decayed painful tooth for rest of life.	60% chance of having sound tooth for rest of life. 40% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	50% chance of having sound tooth for rest of life. 50% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	40% chance of having sound tooth for rest of life. 60% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	30% chance of having sound tooth for rest of life. 70% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	20% chance of having sound tooth for rest of life. 80% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	10% chance of having sound tooth for rest of life. 90% chance of immediate extraction.	____	____
100% chance of having decayed painful tooth for rest of life.	0% chance of having sound tooth for rest of life. 100% chance of immediate extraction.	____	____

This section is similar to the previous section except that this time the choice is between:

a) A certainty of having a tooth filled but thereafter it being healthy and will last until you die.

OR

b) Where you have a chance of having a completely healthy tooth for the rest of your life but there is also a possibility that the tooth must be extracted immediately.

Indicate your preference with a tick.

A	B	Prefer A	Prefer B
100% chance of having a filling which will then last for the rest of your life.	100% chance of having sound tooth for the rest of life. 0% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	99% chance of having sound tooth for the rest of life. 1% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	95% chance of having sound tooth for the rest of life. 5% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	90% chance of having sound tooth for the rest of life. 10% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	80% chance of having sound tooth for the rest of life. 20% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	70% chance of having sound tooth for the rest of life. 30% chance of immediate extraction.	____ ____	____ ____

A	B	Prefer A	Prefer B
100% chance of having a filling which will then last for the rest of your life.	60% chance of having sound tooth for the rest of life. 40% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	50% chance of having sound tooth for the rest of life. 50% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	40% chance of having sound tooth for the rest of life. 60% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	30% chance of having sound tooth for the rest of life. 70% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	20% chance of having sound tooth for the rest of life. 80% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	10% chance of having sound tooth for the rest of life. 90% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will then last for the rest of your life.	0% chance of having sound tooth for the rest of life. 100% chance of immediate extraction.	____ ____	____ ____

This section is similar to the previous ones except that this time the choice is between:

a) A certainty of having a tooth filled but the tooth will need to be filled again at least once within the next 5 years.

OR

b) Where you have a chance of having a completely healthy tooth for the rest of your life but there is also a possibility that the tooth must be extracted immediately.

Indicate your preference with a tick.

A	B	Prefer A	Prefer B
100% chance of having a filling which will need to be replaced within 5 years.	100% chance of having sound tooth for the rest of life. 0% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	99% chance of having sound tooth for the rest of life. 1% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	95% chance of having sound tooth for the rest of life. 5% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	90% chance of having sound tooth for the rest of life. 10% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	80% chance of having sound tooth for the rest of life. 20% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	70% chance of having sound tooth for the rest of life. 30% chance of immediate extraction.	____ ____	____ ____

A	B	Prefer A	Prefer B
100% chance of having a filling which will need to be replaced within 5 years.	60% chance of having sound tooth for the rest of life. 40% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	50% chance of having sound tooth for the rest of life. 50% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	40% chance of having sound tooth for the rest of life. 60% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	30% chance of having sound tooth for the rest of life. 70% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	20% chance of having sound tooth for the rest of life. 80% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	10% chance of having sound tooth for the rest of life. 90% chance of immediate extraction.	____ ____	____ ____
100% chance of having a filling which will need to be replaced within 5 years.	0% chance of having sound tooth for the rest of life. 100% chance of immediate extraction.	____ ____	____ ____

This section is similar to the previous ones except that this time the choice is between:

a) A certainty of having a decayed tooth but one which will last until you die, but it will not be painful.

OR

b) Where you have a chance of having a completely healthy tooth for the rest of your life but there is also a possibility that the tooth must be extracted immediately.

Indicate your preference with a tick.

A	B	Prefer A	Prefer B
100% chance of having a decayed non-painful tooth for the rest of life.	100% chance of having sound tooth for the rest of life. 0% chance of immediate extraction.	____	____
100% chance of having a decayed non-painful tooth for the rest of life.	99% chance of having sound tooth for the rest of life. 1% chance of immediate extraction.	____	____
100% chance of having a decayed non-painful tooth for the rest of life.	95% chance of having sound tooth for the rest of life. 5% chance of immediate extraction.	____	____
100% chance of having a decayed non-painful tooth for the rest of life.	90% chance of having sound tooth for the rest of life. 10% chance of immediate extraction.	____	____
100% chance of having a decayed non-painful tooth for the rest of life.	80% chance of having sound tooth for the rest of life. 20% chance of immediate extraction.	____	____
100% chance of having a decayed non-painful tooth for the rest of life.	70% chance of having sound tooth for the rest of life. 30% chance of immediate extraction.	____	____

100% chance of having a decayed non-painful tooth for the rest of life.	60% chance of having sound tooth for the rest of life. 40% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	50% chance of having sound tooth for the rest of life. 50% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	40% chance of having sound tooth for the rest of life. 60% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	30% chance of having sound tooth for the rest of life. 70% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	20% chance of having sound tooth for the rest of life. 80% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	10% chance of having sound tooth for the rest of life. 90% chance of immediate extraction.	___ ___	___ ___
100% chance of having a decayed non-painful tooth for the rest of life.	0% chance of having sound tooth for the rest of life. 100% chance of immediate extraction.	___ ___	___ ___