A STUDY OF FACTORS ASSOCIATED WITH ORTHODONTIC TREATMENT OUTCOME

A thesis submitted to the University of Glasgow for the degree of Master of Science (Med. Sci.) in Orthodontics of the Faculty of Medicine.

Philip John Stockley Taylor
Department of Child Dental Care
1994

A study of factors associated with orthodontic treatment outcome

CONTENTS

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor W.J.S. Kerr, supervisor of this project, for his initial impetus, guidance, constructive criticism, and generous encouragement. Also to Mr J. McColl, Department of Statistics, University of Glasgow, for his direction and evaluation of the statistical analysis.

A large measure of thanks are due to Miss Alison Downing, who gave much of her time and patience in assisting with the error study.

The project would not have been possible without the support of Mr D.R. McCall, CADO, Dumfries and Galloway Health Board, to allow leave of absence to undertake this work. It also seems an opportune time to thank Mr J.C. Aird, Consultant in Orthodontics, Dumfries and Galloway Royal Infirmary, for investing so much time, energy and enthusiasm in initiating the training programme from which this thesis springs, and for his kind encouragement prior to and throughout the period of study.

There are many who have contributed in some way to this production and I am grateful to them all, but I am most in debt to my long-suffering wife, Jennifer, without whose selfless support this project could not have been contemplated.

DECLARATION

This thesis has been entirely the work of P.J.S. Taylor. No part of this thesis has been previously submitted in support of an application for any degree or qualification of any University.

P. J. S. Taylor

ABSTRACT

Objective measures of malocclusion have been developed in response to concerns over treatment standards and provision of unnecessary treatment, the rise in the level of clinical audit and the need for effective use of health resources.

To determine which factors, including treatment need and severity of malocclusion measured by occlusal indices, have most influence on quality of result, improvement, and resource commitment in terms of treatment duration, case notes and pre- and post-treatment study casts of 161 cases treated within a dental hospital orthodontic department were examined and assessed.

Differences were found between cases treated with full fixed appliances and other appliance types with regard to severity of malocclusion, extraction pattern, duration of treatment and number of appointments required to complete treatment. There was evidence that different criteria in appliance selection were in use for males and females.

Stepwise multiple linear regression analysis identified the initial severity of malocclusion as being highly influential and predictive for fixed appliance therapy in all the outcomes under study. Other important information included whether or not the treatment plan had been altered through poor co-operation, whether a first permanent molar had been extracted, and whether or not an anterior crossbite had been present initially.

Accurately predictive models were produced for improvement as measured by change in PAR score, but predictive models for post-treatment PAR and duration of treatment were less accurate.

A study of factors associated with orthodontic treatment outcome

Section A

REVIEW OF THE LITERATURE

Section contents Page

Chapter 1

Measurement of occlusion, treatment outcome and the need for occlusal indices

1.1 Introduction

Assessment of a malocclusion, the need for treatment, the likely difficulty involved, the potential resource commitment required and subsequent change in occlusion and assessment of success of orthodontic treatment have been largely a matter of subjective opinion of individual clinicians.

Any assessment of factors affecting the outcome of treatment must take account of the initial severity of the condition and the improvement produced by treatment.

Whereas it is possible to calculate duration of treatment, number of teeth extracted, number of appliances and archwires used *etc.,* objective measurement of improvement due to treatment has, in the past, been problematic.

The need for greater objectivity in occlusal assessment has been engendered by recently increased attention to certain aspects of orthodontic diagnosis and treatment.

These include concern over the provision of unnecessary treatment, the increased level of clinical audit, anxieties over orthodontic treatment standards generally, and the efficacious utilisation of progressively more constrained public health resources.

1.2 <u>The determination of need for treatment</u>

The existence of a need for orthodontic treatment is based on the premise that there is an accepted norm which comprises a group of individuals "who function better, have fewer defects, better development and adapt better than a group of individuals who have morphological or functional traits which deviate from the accepted norm." (Report Of The Occlusal Index Committee, 1987)

The W.H.O. Expert Committee on Dental Health (1962) considered that treatment was appropriate where the patient suffered from disfigurement or a functional defect but as the assessment of dento-facial anomaly was dependent on the subjective opinion of the examiner, no classification of the anomaly was made other than whether or not the patient required treatment. Inevitably this leads to disagreement on the need for treatment in some cases. Richmond *et al.* (1994) found little agreement among a group of 74 dentists in subjective treatment need assessment and in some instances intra-examiner agreement was purely by chance, although there was greater consistency and closer agreement among orthodontically trained clinicians compared to non-specialist practitioners.

In 1986 the Committee of Enquiry into Unnecessary Dental Treatment was set up by the Minister of Health in response to concerns expressed by the dental profession that abuse of the NHS system of remuneration was occurring by the prescription of unnecessary treatment. This was bringing the profession into disrepute having been drawn to public attention by extensive television and press coverage.

The Committee reported that evidence concerning orthodontics was more worrying than that related to other forms of treatment as regards unnecessary prescribing.

Unnecessary treatment in orthodontics concerns not only cases where only a marginal benefit to the patient could possibly be achieved even with the ideal conclusion of the treatment plan, but also cases where the practitioner cannot reasonably expect to achieve a beneficial result because of the improbability of good patient co-operation.

Bearing this in mind it is notable that within the General Dental Service during 1979, the patient failed to complete the proposed course of treatment in 25% of cases, (Haynes, 1982).

Under the Terms of Service, practitioners within the NHS are required to have as their objective "dental fitness" for their patients. Under regulation 2(1) of the NHS(General Dental Services) Regulations dental fitness is defined as a reasonable standard of dental efficiency and oral health as is required to safeguard general health.

The Committee of Enquiry recommended that there should be no change in this definition, therefore any treatment which was provided purely for cosmetic reasons seemed excluded from treatment necessary to provide dental fitness.

McLain and Proffit (1985) stated that "problems cannot be defined solely in physical terms. The impact of dental and facial characteristics on an individual's well-being is important, perhaps more important than the impact on physical health."

Small variations in tooth position can have a significant effect on the overall aesthetics of the face, (Sergl & Stodt, 1970). The face and oral region appears to be of prime importance in determining an individual's

attractiveness in inter-personal contact. Visible, disfiguring, dental features which differ greatly from the norm may stigmatise, impede career advancement and peer group acceptance, encourage negative stereotyping and have a negative effect on self-concept, (Cons *et al.*, 1986).
Individuals with normal incisor relationships have been rated significantly higher for personal friendliness, social class, popularity and intelligence than those displaying unattractive dental features although background dental attractiveness has been found more influential than dental configuration in many cases, (Shaw *et al.*, 1985).

Persons with some degree of facial unattractiveness are likely to have lower academic expectations held for them by their schoolteachers and lower ratings of behaviour and personality, (Salvia *et al.*, 1977).

Also the physically unattractive actually achieve less educationally and are less socially skilled. There is considerable evidence that the prejudice physically unattractive children suffer constitutes a negative social feedback and significantly affects self image and behaviour, (Rosenthal & Jacobsen, 1968).

Unattractive children are more likely to be the victims of bullying whilst teasing among schoolchildren has been found to be more hurtful if it occurred because of dental features as opposed to other physical characteristics, (Shaw *et al.*, 1991).

Prahl-Anderson stated that individuals with more severe disfigurement tend to elicit pity rather than ridicule. Personal reactions to these individuals may be predictable and the individuals concerned and their parents are able to develop compensatory mechanisms for psychological protection, such as the development of norms different to those generally accepted.

In cases of less severe dento-facial disfigurement, other persons' reactions to an individual are less predictable, such that compensatory mechanisms are less well developed. These people tend to be objects of amusement rather than pity and are more inclined to develop behavioural disorders than those with more severe disfigurement.

On the other hand little evidence was found by Kenealy *et al.* (1989) that malocclusion was disadvantageous for the social and psychological well-being of a child, although studies of cleft lip and palate cases indicate a general impairment in the quality of life for these more seriously affected individuals. Shaw, Addy and Ray (1980) cast doubt on the requirement for perfect and regular teeth to fulfil the emotional and psychological needs of the patient.

The Committee of Enquiry into Unnecessary Dental Treatment did however accept that certain malocclusions can cause disfigurement and be deleterious to a child's long term psychological and social development in terms of self confidence and achievement but concluded that "individuals with other than extreme cosmetic oral defects will be unlikely to have serious emotional difficulties".

The requirement for perfect and regular teeth to achieve the objectives of future dental fitness and efficiency, and oral health leading to good general health has also been questioned, (Shaw *et al.,* 1980).

With regard to the relationship between orthodontics and the achievement of dental fitness as defined by the NHS Terms of Service, evidence is lacking. Such evidence which is available is strong in only a few areas and is controversial in others.

Early correction of prominent incisors reduces the risk of damage to these teeth, although the risk decreases with age, (Jarvinen, 1979).
Dental health implications exist also in extreme situations such as cases of traumatic occlusion, very deep overbite causing gingival trauma, and gross displacement of individual teeth, (Shaw, 1991). There also appears to be an association between excessive overjet and pathological migration of the incisor teeth in adults, (Thilander, 1984).

Studies on the relationship between malocclusion and periodontal disease have produced conflicting results, with some studies finding a causal link and some finding no association. This may have been occasioned by differences in the groups selected for study. In several cases the samples comprised dental students who might be expected to have high standards of oral hygiene, whatever their dental arrangements.

Where oral hygiene is either very poor or excellent, any effect of dental irregularity on the periodontal condition is likely to be masked. In studies comparing patient groups with dental student groups (Alexander & Tipnis, 1970; Sandli, 1973) there was a significant association between dental irregularity and gingival inflammation in the patient groups but not in the student groups.

Sandli confirmed the possibly detrimental effect of lower incisor crowding on the periodontal tissues but the relationship is not consistent. An increased susceptibility to periodontal disease has been found in Class III malocclusions, where there is an anterior crossbite and where there is a deep traumatic overbite, (Sergl, 1977).

Regarding the association between caries and malocclusion the reports are inconclusive and contradictory. Miller and Hobson (1961) found the DMF rate to be increased in a group of 14-year-olds with malocclusion, but found no association between dental irregularity and DMF in groups of 12-year-olds. Where indices of malocclusion have been used no connection between incisor crowding and interproximal caries has been found, (Katz, 1978). In a 22 year longitudinal study Helm and Petersen found no relationship between caries incidence and degree of malocclusion, (Helm & Petersen, 1989).

Concerning mandibular dysfunction the association between orthodontics and temporomandibular dysfunction is again controversial. Mohlin and Thilander (1984) found weak associations between occlusal variations and dysfunction but TMD has a complex and multifactorial nature and no strong link between the presence of malocclusion and the incidence of TMD or between orthodontic treatment and reduction in symptoms of TMD has been established, (Rendell *et al.,* 1992).

The Committee of Enquiry into Unnecessary Dental Treatment stated that further research was necessary into what treatment should be provided under the NHS and which can be justified because of its contribution to dental fitness.

1.3 Utilisation of health resources

The demand for orthodontic treatment has been steadily increasing. The number of courses of orthodontic treatment provided has risen from 75,780 in 1963 to 178,620 in 1973, to 293,920 in 1983. Robertson and Hoyle (1983), suggested that there would be a 0.5% annual growth in active orthodontic treatment in the General Dental Services and a continuing increase in demand for treatment in response to higher standards has been predicted, (Stephens, Orton & Usiskin, 1985).

There may also be a substantial pool of unmet need. In a survey of 269 14-year-old Paisley school children Luffingham and Campbell (1974) found by subjective assessment that 50% were in need of appliance therapy whilst only 15% of those in need were receiving treatment.

This study did demonstrate that professionally assessed need for treatment does not necessarily equate with demand for treatment. Only 17% of those needing treatment but not receiving it accepted the offer of treatment leaving only 6% of those needing treatment and amenable to treatment not actually receiving it. The authors believed that this particular population of school children demonstrated a low level of dental awareness thus the discrepancy between need and demand may be greater than that in other populations.

The BSSO reported to the Committee of Enquiry into Unnecessary Dental Treatment (1986) that 30% children aged 15 needed treatment but had not received it, thus if this need was to be met, additional money would have to be made available or fees for other forms of dental treatment would have to be reduced.

There was concern that there might be a large potential market for treatment of minor malocclusions where only marginal benefit could be expected and also that this market would be largely driven by the operators themselves.

The Committee considered evidence for and against the development of an index of occlusion to enable malocclusions to be graded in order of severity and need for treatment, and referred to the Swedish system as an example, (Linder-Aronson, 1974).

Elderton and Clark (1983) reported that such an index was necessary as a starting point for establishing a method of comparison but that there were many difficulties involved.

To this end it was recommended that an index for the use of the Dental Estimates Board should be developed to facilitate Board decisions in regard of giving prior approval for treatment plans under the NHS. At that time half of the dental advisers to the Dental Estimates Board worked exclusively scrutinising estimates for prior approval orthodontic treatment although orthodontic estimates comprised only 5% of all estimates.

However recommendation 10 of the Report states that dental staff at the DEB who deal with orthodontic references should have an orthodontic qualification. Also where treatment was proposed on purely cosmetic grounds additional information may be requested by the DEB in order to satisfy themselves that the patient or parents were "seriously concerned" about the condition.

The above recommendations would seem to demand that a degree of flexibility needs to be built into the system such that a malocclusion which does not reach a certain grade of severity is not arbitrarily denied treatment where there are other factors present.

In the light of the report of the Committee of Enquiry into Unnecessary Dental Treatment, the Occlusal Index Committee was set up in June 1986. The brief of the Occlusal Index Committee was to make recommendations on the development of an occlusal index which might be used by the Dental Estimates Board.

The type of index which was thought most relevant to the functions of the Dental Estimates Board was an index of orthodontic treatment need.

The Occlusal Index Committee stated that any index of treatment need must include some measure of social function handicap, namely dental aesthetics. Orthodontic treatment is sought by patients more often for aesthetic reasons than on functional grounds. However it was stressed that there is great personal variation in adjustment to dental disfigurement and so the index must be sensitive to the individual.

The Committee concluded that an index of orthodontic treatment need could in principle be of use in the processing of orthodontic estimates and in communications between practitioners, their patients and the Dental Estimates Board. Also the development of an index pertaining to treatment difficulty and the level of expertise required for successful completion of treatment in individual cases was recommended.

1.4 Treatment standards

The situation in the United Kingdom in the early 1980's was one of a lack of a policy on treatment standards, there was a financial disincentive to the use of fixed appliances, there were no guidelines to the clinical competence of operators, no case load restraints and unrestricted prescription of treatment, (Shaw, O'Brien & Richmond. Paper presented to the British Orthodontic Conference, 1992).

A Steering Committee towards a Board of Certification was set up in 1975 because of a feeling that there was a need for United Kingdom orthodontic treatment standards to equal the level prevailing in the United States and Europe, (Bass, 1988). At that time there was no recognised level of skill to which an individual could aspire.

A subsequent Working Party reported that a Certification Board should be developed and supported by the British Association of Orthodontists and during 1978 and 1979 individuals were invited to form initial Examining Panel and Management Committees. In 1983 only certificated clinicians were permitted to be Board members.

The aims of the Board were to encourage the attainment of excellence in clinical practice, improve standards of orthodontic treatment and provide certification as recognition of that attainment. Certification however was not intended as a requisite for specialist registration nor to cause clinical limitation nor reflect deprecatingly on the competence of any clinician.

It is fair to say that the Certification Board did not have the universal support of the speciality as the criteria for certification bore no relationship to already established training programmes and examination structure. Nevertheless the existence of the Certification board provided a spur to the speciality to actively co-operate in moves towards a specialist register in the form of a Joint Working Party set up in 1989, (Joint Working Party on Specialist Registration in Orthodontics, 1990).

Elderton and Clark (1984) using the Occlusal Index (Summers, 1966) with Scottish General Dental Services patients found that only one-third of cases showed a substantial improvement and one-third showed little or no improvement.

In 1984 the Senior Dental Officer responsible for orthodontics in England and Wales disagreed significantly with the proposed treatment in 58% of cases referred to the Dental Estimates Board.

The annual report of the Dental Estimates Board indicated that 21% of treatments started were not completed, however it is possible that substantial improvements may have been achieved in those cases even the courses of treatment were not concluded.

The 1983 OPCS survey of children's dental health (Todd & Dodd, 1985) reported that 30% of 15 year old children who had previously completed a course of orthodontic treatment were still in need of treatment in the opinion of a group of General Dental Practitioners and Community Dental Officers.

Shaw, in a comparison of orthodontic treatment services in England and Wales with those in Scandinavia, found that the impression given by direct clinical observation was that the treatment standards in England and Wales fell short of those achieved elsewhere, and that the standard of finished treatments submitted to the Dental Estimates Board was known to be a matter of concern notwithstanding the fact that in 1980 only 9% of cases were treated with fixed appliances and 75% of British Orthodontic Association members said less than 40% of their cases were treated by fixed appliances, (Shaw, 1983).

The British Orthodontic Standards Working Party was launched in 1977 with representatives from the BSSO, BAO, COG and AUTO with a view to the improvement of treatment standards in the United Kingdom.

The working party agreed that the quality of treatment provided by the orthodontic treatment services may be bettered by attention to the undergraduate training curriculum, by encouraging continuing post-graduate orthodontic training for General Dental Practitioners including an increase in the number of clinical attachments and nationally graded courses enabling progression on to carrying out treatments of increasing complexity.

Also it was considered that treatment quality may be improved by attention to the training of specialist orthodontists including the establishment of a specialist register, and by fostering better treatment standards generally, ensuring the maintaining of the standards taught during training.

It was noted that within the General Dental Services practitioners are contracted to produce satisfactory results and those results may be examined by the Dental Estimates Board. However, because of the criteria employed in the decision to reimburse practitioners for orthodontic treatment, payments were made in cases where a satisfactory result had not been achieved.

The Dental Estimates Board considered 49% of treatments completed during 1984 to be unsatisfactory.

A later study by Richmond *et al.* (1993) found that out of 200 non-prior approval and 1010 prior approval cases which represented a 5% sample of cases dealt with by the Dental Estimates Board a third of both non-prior approval and prior approval cases remained in need of treatment after appliance therapy. Furthermore 42% of non-prior and 21% of prior approval cases were found to have shown no improvement through treatment.

The British Orthodontic Standards Working Party was reconstituted in 1983 with an increased membership including a representative of the Community Orthodontists Group.

In its Second Report (BOSWOP, 1986), the working party recognised that although the quality of some United Kingdom treatments was excellent, overall standards were lower than those in North America and the Northern European countries but at that time there was only limited information available about the level of UK orthodontic treatment standards.

Case load was identified as a factor related to treatment standards such that a reduction in a clinicians' case load to manageable levels was regarded as desirable. The Working Party noted that modern treatment techniques can be time consuming whilst demand for treatment was increasing. It was believed that at least in the short term the demand for treatment would not be met by an increase in orthodontically trained manpower, but a priority index may be used to limit unnecessary or questionable treatments.

It was felt that General Dental Practitioners should be able to contribute to an increase in treatment standards by having a critical approach to the judgement of need of treatment and of when and to whom to refer cases for more complex treatment. Also by adopting a critical attitude towards the standards achieved by their specialist colleagues.

However graduates often felt inadequately prepared for the clinical management of orthodontic cases, (Pender, 1984).

Much of the Report concerned undergraduate training, postgraduate training for Hospital, Community and Specialist orthodontic practice and the establishment of a Specialist Register.

The Working Party also expressed the opinion that ways in which objectives and corresponding treatment outcomes within the General Dental Services should be assessed together with their implications for a system of payment, as financial considerations were believed to affect case loads and standards achieved. Systems employed in other countries for determining treatment priority and case load control and methods of matching of treatment complexity with the appropriate level of operator training should also be examined.

1.5 Audit

The application of medical audit is regarded as a fundamental principle in the Government White Paper "Working for Patients" (1989) and it was suggested that the development of an effective system of audit would have the benefit of reassuring professionals, public and management that the highest achievable quality of care is maintained within the resources available.

Medical audit pertaining to dentistry has been a relatively recent development yet audit has been accepted as necessary by the medical profession for some years, (Ashley-Miller, 1977).

There has been some uncertainty within the profession about the meaning of the terms peer review, clinical audit and dental audit, (Mouatt & Heap, 1991).

Audit has been variously defined as;

> "the systematic, critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcome and quality of life for the patient" (Working for Patients, 1989), also as;

> "a systematic approach to the peer review of medical care in order to identify opportunities for improvement and provide a mechanism for realising them", (Shaw & Costain, 1989).

Audit in clinical practice must include the following features:

> definition of criteria, standards, targets or protocols for good practice

data obtained must be objective and be gathered in a systematic fashion.

comparison of results against a standard or among peers.

identification of deficiencies in clinical practice and their correction.

monitoring the effects of any action taken, (Hughes & Humphrey, 1990).

Stratford (1991), in commenting on medical audit within hospital orthodontic departments, understood that some colleagues might consider the occasional meeting with fellow clinicians to discuss interesting cases as audit. However the systematic nature of and objective approach to audit are recurring elements in the above definitions.

The author believed that the meaning of quality control needed to be clarified. She felt that high quality treatment could be taken to mean "the best treatment individually tailored for a specific patient" rather than a good result in comparison with some morphological ideal and this should take into account the patient's own wishes, capabilities and limitations, (Parkhouse, 1987).

The adoption of this individualised objective may have implications for any measure of treatment success, in that measurements based on absolute morphological criteria may be inappropriate in some cases.

Stratford concluded that Medical audit should not be confused with attendance statistics, resource management or cost effectiveness, but concerns the quality of treatment provided and should be assessed through agreement with fellow professionals.

On the other hand, quality in health care has been alternatively described as concerning not only the achievement of professionally defined treatment goals (Professional quality) but also with patient satisfaction (Customer quality) and with the efficiency of the method used to achieve those goals (Process quality). This indicates that treatment standards, resource management and the use of the most efficient treatment regimes need to be considered in any system of audit. The process quality element of health services, which has largely been ignored in discussions of quality assurance and quality programmes, has been central to most commercial organisations' strategies, (Ovretveit, 1990).

Vig, Weintraub, Brown and Kowalski (1992) have intimated the need for evaluation not only of the effects of orthodontic treatment but also the need to gather objective data on treatment efficacy. They suggest that "increasing pressures for cost containment in health care concurrent with the need and desire to provide treatment of optimum quality dictates that research endeavors[*sic*] address the cost-effectiveness of alternative clinical policies and decisions and their outcomes".

Quality of medical care has been described by Donabedian (1966) as comprising 3 distinct parts; structure, process and outcome.

> Structure is concerned with facilities, equipment, available materials, organisation, record keeping and the appropriateness of the operator in terms of level of skill and training. These factors have an indirect bearing on the benefits accrued by the patient from the health care provided.
>
> Process refers to the decisions and actions taken by the practitioner and relates to the skill, knowledge

and attitude of the operator and the appropriateness and limitations of the treatment technique used. Process has a more direct effect on the resolution of a patient's problem than audit concerned with structure.

Outcome may be defined as the changes to a patient's current health status and future health prospects as a result of the health care provided. The outcome relative to the initial problem describes the effectiveness of the administered treatment.

In orthodontic terms where often the reason for a patient seeking treatment is aesthetics, where the problems are the possible socio-psychological effects and negative effects on self-image of a disfiguring malocclusion, a successful definitive outcome would be the psychological and social development of the individual to his or her maximum potential and absence of behavioural problems.

In this case measurement of outcome would be impossible in the short term and difficult in the long term because of the myriad other factors involved in an individual's overall development.

The term "intermediate outcome" may be useful in describing the stage between process and true definitive outcome, (Irvine & Irvine, 1991). Intermediate outcomes are easier to measure and may be assumed to have a bearing on the final outcome.

In the field of orthodontics, process may be considered for example to refer to the particular treatment technique used, duration of treatment, amount of chairside time required, laboratory commitment, level of training of the operator, degree of discomfort and inconvenience to the patient, whereas intermediate outcome might refer to the improvement in aesthetics and the degree of residual irregularity at the end of treatment.

In summary, clinical audit in orthodontics requires dentists to objectively question their working practices and to justify their actions to themselves and their colleagues, and if deficiencies are found then to modify their practice and monitor the effects of the changes made. Thus the attempt should be made to provide the highest possible quality care within the limitations imposed by the patient's suitability and capability, and within financial constraints.

It is reasonable to assume that the use of an index for occlusal assessment using well defined and objectively measured criteria would aid clinical audit in orthodontics, (O'Brien, Shaw & Roberts, 1993). Wide acceptance and calibration in the use of such an index would facilitate both self-assessment and comparison between peers although differing circumstances, for example case-mix and local facilities, would have to be considered.

Chapter 2

Occlusal indices: General aspects

2.1 The desirable properties of an occlusal index

The indexing of a malocclusion was defined by the Occlusal Index Committee (1987) as the recording of the traits of malocclusion in numerical or categorical format which might involve direct physical measurement, the recognition of different morphological traits, functional traits, or a more general appraisal of the dentition.

The properties which an ideal index should possess have been suggested by various authors and these may be listed: (Summers, 1966; W.H.O., 1966; Occlusal Index Committee, 1987)

1 Reliability in clinical use

2 Clinical validity and validity over time

3 Reproducibility in everyday practice

4 Sensitivity to individual patient needs

5 Acceptability to both the profession and the public

6 Administrative simplicity in operation

7 Equal sensitivity throughout the range of malocclusion

8 The index should require a minimum of subjective judgement and any decisions made should be between mutually exclusive conditions based on well defined criteria

9 Ability to be performed both on patients in the clinical situation and on study casts

10 The index should be amenable to statistical analysis

Given the infinite variation seen in dento-facial morphology, in individuals' limitations with regard to treatment possibilities and psychological and sociological needs, it is unlikely that an index could be developed which does not produce the occasional anomaly in use.

2.2 The validity and reliability of occlusal indices

An index may be said to be valid if it measures what it purports to measure.

Before the validity of an index can be tested, the specific use to which the index is to be put must be clearly established, although it is possible that an index may be designed with more than one use in mind, (Carlos, 1970).

The evaluation of the validity of an index involves the estimation of its sensitivity *i.e.* the probability that it will correctly identify cases which demonstrate the characteristic being measured, and its specificity *i.e.* the probability that it will correctly identify cases which do not show the characteristic.

This requires an independent and previously validated means of identifying cases with the characteristic against which the index can be tried and its tendency for indicating false positives and false negatives evaluated.

Unfortunately regarding malocclusion, the indices themselves are the only means available. There is no independent, physical indicator of need for orthodontic treatment available against which a need for treatment index can be set. Validation of orthodontic need indices has therefore been by the criteria of subjective clinical opinion. This approach to validation is flawed in two ways:

1 The judgement of the clinician is not independent of the index being evaluated. The information on which the judgement is made and that on which the index is based is precisely the same.

2 It has not been shown that the subjective clinical judgement of handicapping malocclusion has any validity in itself.

There is merit in testing validity by comparing decisions made with an index with decisions made by subjective judgement where an attempt is to be made to substitute subjective clinical judgement by the index scores in ranking treatment priority.

Where an index is used in an epidemiological survey an attempt is made to establish the distribution of index scores in a population sample to enable description or comparison. Such an index is valid if the investigator provides a clear statement of what is measured and how it is measured. *i.e.* definitions of conditions and components included in the index are chosen on informed clinical judgement. Such an index has "construct validity" and a question of measurement of such validity is not meaningful.

Carlos equates reliability with precision, *i.e.* reproducibility and testing of reliability is only concerned with inter- and intra-examiner agreement on the same material. Reliability is regarded as being distinct from validity.

In testing reliability of priority indices only the ranking of each case is important. It matters not that all the index scores are too high or too low as long as they are consistent and the rank order remains unchanged.

2.3 Categorisation of occlusal indices

Indices may be categorised by the use to which they are to be put or for which they were designed. The areas in which occlusal indices may be useful are:

Diagnostic classification.

Epidemiology.

Evaluation of need for treatment and ascribing priority.

Evaluation of treatment success.

Prediction of treatment complexity.

Other uses, for example the recording of a specific trait within a dentition.

It is unlikely that an index could be developed which would be universally applicable for all purposes, however it is possible that an index designed for one use may be adapted for another although ideally an index should be developed with a particular use in mind.

Chapter 3

Occlusal indices: Review of reported indices

3.1 Diagnostic classifications

3.1.1 Introduction

These are means by which the vast range of malocclusions may be grouped into a few broad categories essentially for the purpose of communication between practitioners, but they are also useful in the formulation of diagnoses and treatment planning objectives.

3.1.2 Angle's Classification (Angle, 1899)

This classification was based on the fallacious argument that the first permanent molars maintained an unchanging relationship with their supporting skeletal structures. Angle modified the classification to include the relationship between the upper and lower incisors antero-posteriorly, however the arrangement of molar occlusion and incisor occlusion could not always be fitted to one of Angle's categories.

A modification of this classification was suggested by Gravely and Johnson (1974) who believed that the inclusion of a Class II uncertain category increased the classification's ability to discriminate between certain cases. They found Angle's classification unsuitable for epidemiological use as there was a high degree of inter-examiner variability.

This was confirmed by Jago (1974) who believed Angle's classification to be a useful treatment prescription device but who, in a review of the use of this classification in 45 studies internationally across 18 different countries, found a wide variation in the assessment of malocclusions. This was related to imprecision in the terminology used to define the cut-off levels between categories. The low reproducibility of the classification was corroborated by Salzmann (1969) and the limitation of the classification to the antero-posterior relationship of the upper and lower teeth has also been criticised, (Isaacson, 1975).

Angle's classification however did become very widely used and Katz (1978) indicated that it best expressed the patient's own degree of dental self-satisfaction compared to other indices as it is closely related to facial profile and is unaffected by dental features which do not have implications for the facial profile.

3.1.3 The British Standards Incisor Classification (British Standards Institute, 1983)

The incisor classification evolved from the work of Backlund (1963) and was introduced by Ballard & Wayman (1964). It is based entirely on the antero-posterior relationship of the upper incisors to the lower incisors and has been widely adopted by the profession.

It suffers in that limited data is provided and categorisation is difficult in certain circumstances due to ill defined cut-off points requiring subjective judgements.

A modification to the classification has been suggested to increase inter-examiner agreement in some more contentious cases, (Williams & Stephens, 1992).

3.2 Indices for epidemiological use

3.2.1 Introduction

These indices are designed to investigate and record the prevalence of occlusal traits and malocclusion within a population.

The indices attempt to record all dental morphological features and therefore tend to be detailed and comprehensive, but have been shown to often have a high degree of reproducibility, (Helm, 1977).

3.2.2 A Method For The Epidemiological Registration Of Malocclusion (Bjork, Krebs & Solow, 1964)

This complex index, which was more a coded description than a statement of severity of malocclusion, assigned occlusal traits to 567 individual code items and was divided into three parts:

1. Dental anomalies, *i.e.* abnormalities of tooth morphology, eruption and malalignment

2. Anomalies of occlusion

3. Discrepancies in spacing/crowding.

Some features *e.g.* overjet overbite and molar relationship were graded arbitrarily for severity and not with any regard to dental health or aesthetic implications.

Although some categories were not mutually exclusive and some category cut-off points were not well defined, for example in the extent of molar eruption or differentiating between "late, but otherwise normal shedding" and "persistent deciduous" teeth, resulting in some scope for inter- and intra-examiner variability, Helm *et al.* (1977) showed 80 per cent agreement between examiners in the use of this index.

The index included a section concerning need for treatment, the most appropriate timing of the treatment *i.e.* within the subsequent two years or not for at least two years, and the estimated length of time that the treatment would take.

The treatment need section was totally subjective and dependent on the experience of the examiner.

3.2.3 A Method For Measuring Occlusal Traits (Baume *et al.,* 1973)

This index was developed by the F.D.I. Commission on Classification and Statistics for Oral Conditions (COCSTOC) Working Group 2 and was intended to present a simple objective method for measuring occlusal traits. It was not designed to include any attempt at establishing or measuring need for treatment. Indeed it was stated that with the current knowledge at that time it was thought impossible to establish meaningful cut-off points to distinguish those who needed treatment from those who did not.

The method involved three general categories of measurement:

1. Dental measurements, including anomalies in number or shape, impacted teeth and retained primary teeth.

2. Intra-arch measurements *i.e.* crowding/spacing.

3 Inter-arch measurements including antero-posterior, vertical and transverse measurements of the incisal and buccal segments and also the presence of upper palatal or lower labial gingival trauma due to the occlusion of opposing incisors.

The inclusion of this sole soft tissue measurement might indicate that there was a view to incorporating some element of need assessment.

The individual traits were recorded as a single code letter together with the FDI code number of the particular tooth concerned. There was no method of summarising the measurements although the development of a summary method of need assessment was the ultimate goal. The index stopped short of being an assessment of need because methods of assessing the psycho-sociological impact of malocclusion and for determining the combinations of traits that resulted in various facial appearances had not been developed.

However the method was simple in use and some reports using this index have been published.
(Cons *et al.,* 1978)

3.3 Indices of treatment need and priority

3.3.1 Introduction

This type of index attempts to categorise malocclusion or to reduce a malocclusion to a number which can then be compared with a scale of norms. Thus one malocclusion may be compared with another and ranked for severity and therefore degree of treatment need.

3.3.2 Handicapping Labio-Lingual Deviations Index (Draker, 1960)

The aim of this index was to complement and perhaps substitute for clinical judgement. It was born of a need to select those who were eligible for treatment under the New York State Dental Rehabilitation program and those who weren't, for which it was stated "the presence or absence of a demonstrable handicap is the only factor of public health interest."

Nine conditions of malocclusion are recorded and scored and different weightings applied to the individual condition scores. The conditions observed are cleft palate, severe trauma or gross pathology, overjet, overbite, mandibular protrusion, open bite, ectopic eruption, anterior crowding and "labio-lingual spread". Labio-lingual spread is an expression of crowding present whereby the distance between the incisal edges of the most lingually placed and most protruding anterior teeth is measured. All measurements are made using a Boley gauge. The measurement of labio-lingual spread may be difficult if the most lingual and most labial teeth are not adjacent.

A final score of more than thirteen is regarded as indicating the presence of a physical handicap.

That there are only two possible results of this index shows inflexibility and insensitivity to individual patient needs, particularly significant as the index was intended to be use by public health dentists rather than orthodontically trained specialists.

3.3.3 The Dental Aesthetic Index (DAI) (Cons *et al.,* 1986)

The DAI was designed to assess the relative social acceptability of dental appearance based on the public's perception of dental aesthetics.

It was developed using monochrome photographs of 200 sets of study casts viewed from the front and from either side with stylised masks to eliminate viewing non-essential and possibly confusing parts of the plaster casts. The sets of photographs were examined by adolescents and non-professional adults and were rated by them for social acceptability.

The designated scores were related to the measurements of certain occlusal traits of each of the occlusal configurations by stepwise regression and factor analysis.

The occlusal traits measured were those 10 recommended by the Commission on Classification and Statistics for Oral Conditions and regression weighting factors were applied to each occlusal trait measurement as follows:

Missing visible teeth	6
Crowding	1
Spacing	1
Diastema	3
Largest anterior irregularity:upper	1
Largest anterior irregularity:lower	1

Anterior maxillary overjet	2
Anterior mandibular overjet	4
Vertical anterior openbite	4
Antero-posterior molar relation	3

The DAI score for a particular dentition was the sum of the products of the 10 measurements and their weightings plus a constant of 13.

The Social Acceptability Scale of Occlusal Conditions was drawn up by the ranking of 1337 study casts. This involved examination of samples from several different countries with different racial and cultural backgrounds by non-professional examiners. No difference in rankings was found between countries.

In prioritising cases there are no pre-set cut off levels between grades, but DAI scores for cases can be compared against this normal scale and discriminatory levels for a particular population may be set according to available resources. Thus the index is proffered as an administrative and management tool.

The index focuses on the aesthetic aspect of malocclusion with no allowance for any oral health components. The rationale which is offered is that presented by Isaacson (1985) who stated that "orthodontic care can not be justified on the basis that negative sequelae would occur without it." It is the view of the author that an index which attempts to combine assessment of factors which may have oral health implications with assessment of aesthetics "will only be confounded."

It is interesting to note the presence of antero-posterior molar relation in the final regression equation. This variable might have been expected to have had a weak link with aesthetics.

The aesthetic appreciation of dental configurations has been examined with no reference to the face in which the dentition resides yet a certain dental appearance which might be acceptable in one subject with a particular soft tissue arrangement and in a particular social situation may be unacceptable in another face and in different circumstances. Asher-McDade *et al.* (1992) have shown that judges rating the lip aesthetics in cleft lip and palate cases can not disassociate their judgement of a localised facial area from an assessment of general facial attractiveness.

The DAI does however attempt to provide a potential instrument for future research into the psycho-social impact of malocclusion.

3.3.4 Handicapping Malocclusion Assessment (Salzmann, 1968)

This index was intended to record any occlusal traits which might jeopardise the health of the dentition or general well being.

It is similar to other indices in that weightings are applied to certain measured deviations and a summary score produced.

Deviations found in the upper anterior segment were weighted by a factor of two. If the total upper anterior segment score exceeded a certain value then further points were added to indicate the presence of an aesthetic impairment. No rationale was presented for the level of the "aesthetic handicap" threshold. Presumably this was an arbitrary decision.

The value of the weightings were also arbitrarily decided upon and thus depended on subjective clinical experience and belief concerning the dental health implications of the traits rather than any objective evidence of oral health sequelae.

Grewe and Hagan (1972) found a greater level of systematic error in use of this index than with the TPI or OI and the lowest reproducibility of these three indices.

3.3.5 Treatment Priority Index (Grainger, 1967)

This index was based on the study of inter-relationships of ten manifestations of malocclusion as they occurred in 375 12-year-old Ontario children who had no history of orthodontic treatment.

The ten features were; increased overjet, reverse overjet, increased overbite, anterior open bite, congenital absence of incisors, disto-occlusion, mesio-occlusion, posterior crossbite, scissors bite and tooth displacement.

It was designed to assess the effect of preventative orthodontic treatment and the prevalence and severity of malocclusion in children by considering five component parts and weighting each part according to the severity of the measured trait.

Turner (1983) suggested alterations to the index weightings, pointing out that underscoring was likely in the presence of submerging second deciduous molars, where lower first permanent molars were missing or of poor prognosis and where extraction of deciduous teeth had been carried out in order to transfer crowding from labial to buccal segments. Overscoring was found to be likely in the presence of bilateral crossbites, bimaxillary proclinations or where the malocclusion was very mild.

Reproducibility with the TPI has been found to be higher in orthodontically trained examiners than in non-specialist dental personnel. Turner (1990) observed that the inter-examiner agreement between five non-orthodontist Community Dental Officers improved with practice although with the level of reproducibility demonstrated in that study at least one third of examined children would be scored incorrectly. However the TPI is easier to use in the mixed dentition than the Occlusal Index (Summers, 1966) and would therefore be better for orthodontic screening of ten-year-olds as recommended by the 1967 Court Report, (Turner, 1990).

3.3.6 Occlusal Index (OI) (Summers, 1966)

The Occlusal Index was an attempt to refine and develop the TPI as an epidemiological tool and is based directly on that index.

Summers believed that previous attempts to measure occlusion had not been universally accepted because of variations in terminology and definitions.

The author also states that the index should be sensitive to the "basic defect" but not unduly sensitive to normal developmental changes so that in the absence of any intervening treatment the index score of a dentition should remain constant or should increase over time but should not decrease with time. This statement does seem to ignore the wide variation between individuals in the timing, rate and amount of change during development which is normally observed.

Also if the "basic defect" can be measured and variation in growth ignored presumably the index score should remain constant and not decrease or increase.

Nine characteristics are scored in calculating the Occlusal Index. These are dental age, molar relationship, overbite, overjet, posterior crossbite, posterior open bite, tooth displacement, midline relationship and missing permanent teeth.

Overbite and overjet are measured in millimetres.

Dental age — This is categorised from 0 (birth to the eruption of the first tooth) through to VI (complete permanent dentition - 2nd molars may or may not have erupted)

Molar relation — When both dental age and molar relation are considered, "normal" molar relation at one age *e.g.* cusp-to-cusp in the early mixed dentition (DA III) would be abnormal at another *e.g.* cusp-to-cusp in the permanent dentition (DA VI). The molar relationship is described in different dental ages with a different code number being given for each molar relationship/DA combination.

Posterior crossbite — Differentiation is made between crossbites of dental, functional and skeletal origin, and the number of teeth in each type of bucco-lingual relationship is counted.

Posterior open bite — Present or not.

Tooth displacement

Tooth displacement is scored differently for mixed and non-mixed dentitions. In a non-mixed dentition a score of 0, 1 or 2 is given for every tooth depending on the deviation or rotation of that tooth from "normal arch alignment". In the mixed dentition tooth displacement is divided into that not associated with space deficiency, when scoring is as for non-mixed dentitions, and that where displacement is due to lack of space. In this case a mixed dentition analysis is carried out and the degree of space deficiency estimated in millimetres.

Centrelines

Any discrepancy is regarded as a mandibular deviation regardless of whether dental or skeletal factors are involved. Only upper/lower midline discrepancies greater than 3 mm are scored.

Missing permanent teeth

Only missing upper incisor teeth which have not been replaced by a prosthesis are scored.

There are separate weighting mechanisms for each dental age and separate calculating forms for deciduous, mixed and permanent dentitions.

In the Occlusal Index there are two divisions and seven syndromes designed after those used by Grainger in the TPI.

Correct determination of the molar relationship is necessary so that the appropriate occlusal syndrome can be chosen. The score for each item of assessment is placed in the chosen syndrome. Once the division has been decided upon all scores are placed in the syndromes of that division.

To calculate the total score the sum of the highest scoring syndrome is taken and to that is added half of the sum of the remaining syndromes in the division.

Six judges ranked the 60 sets of study casts used in the validation exercise. The subjective criteria used in ranking the casts were accorded relative degrees of importance. The criteria were:

1 Aesthetics	50%
2 Function	35%
3 Treatment difficulty	15%

A subjective classification into treatment need groups according to index score was suggested:

Good occlusions	0.0 to 2.5
No treatment	2.6 to 4.5
Minor treatment	4.6 to 7.0
Definite treatment	7.0 to 11.0
Worst occlusions	11.1 to 16.0

An indication of the type of treatment which would be required was included with these classification guidelines such that cases in the "Minor treatment" group could be treated with simple space regainers or removable appliances, "Definite treatment" cases would require multi-banded appliances and "Worst occlusions" would necessitate "major treatment".

The validity of the index was tested by comparison of the Occlusal index scores with the clinical standard rankings. Good correlation (r=0.920) was found between the standard rankings and the ranked OI scores and there was also good inter-examiner agreement.

However this method of establishing validity may not be appropriate. Carlos (1970) has suggested that the selection of statistical techniques for index evaluation requires an unambiguous prior decision based upon the proposed use of the index.

In this case the index is attempting to combine epidemiological data recording, assessment of treatment need and probable treatment difficulty into a single summary score and might therefore be criticised on this basis.

The index depends on the accurate assessment of molar relationship which may be difficult if one or more molars are missing. Also some authors have felt that the criteria employed by the OI have not been adequately defined, (Pickering & Vig, 1976).

The assessment of tooth displacement is difficult. There is 3 times as much error in measurement of tooth displacement than in other Occlusal Index measurements, (Elderton & Clark, 1983).

Buccal occlusions which show discrepancies of a full unit antero-posteriorly but which may be perfectly acceptable and functional are penalised by the OI.

The directions for use of the index have been criticised as being difficult to understand initially, (Hermanson & Grewe, 1970).

The index considers morphological features only, and does not take into account aesthetic or functional factors.

3.3.7 Index Of The Swedish Health Board (Linder-Aronson, 1974)

This index was formulated in 1966 by the orthodontic section of the Swedish Dental Society and the Swedish Medical Board in order to prioritise cases. An index of treatment need, together with a knowledge of the frequency of dental anomalies within the population and also the treatment time involved, is cited by Linder-Aronson as being necessary to determine orthodontic manpower requirements in a public health system.

Unlike the above indices of treatment need which require direct measurement and calculation of a summary score, this index requires a dentition to be graded into one of four categories dependent on the worst feature of the malocclusion.

The categories are:

Grade 1. (Little need) This includes minor deviations from ideal occlusion *e.g.* prenormal occlusions with little or negative overjet, mild anterior open bite, crossbite without displacement, mild crowding or spacing.

Grade 2. Aesthetically or functionally disturbing proclined incisors, infra-occlusion of deciduous molars or permanent teeth, severe crowding or spacing, moderate rotations of anterior teeth, increased overbite without gingival trauma.

Grade 3. Prenormal forced bite, increased overbite with gingival trauma, extremely open bite, posterior crossbite with displacement, scissors bite, severe anterior crowding or spacing, retained canines, cosmetically or functionally disturbing rotations.

Grade 4. (Very urgent need) Cosmetic and functionally handicapping anomalies, *e.g.* cleft lip and palate, extreme postnormal and prenormal occlusion, retained incisors, extensive aplasia.

In effect these are a set of guidelines and the decision to grade a dentition into one or other of the categories is entirely subjective. The cut-off points to each grade are not defined and there is wide scope for interpretation by clinicians. On the other hand, this allows considerable flexibility in individual cases where use of public funds is being allocated.

The large difference in frequency of malocclusion between two counties found with this index was explained by the author in terms of differences in available treatment facilities rather than possible inter-examiner variation in interpretation of the index criteria.

3.3.8 Malocclusion Severity Index (Hill, 1992)

The MSI is similar to several of the above indices in that weighting scores are applied to various occlusal traits and a single summary score calculated. The weightings were established following reference to the literature linking malocclusion with dental health and to previous indices which employed weighting scores.

Weighted items are either linear and measured in millimetres or a number of teeth showing a particular trait *e.g.* crossbite, incisor rotations over 30°, number of slipped incisor contacts.

Overall scores are graded in a similar fashion to those in the Occlusal Index but differ in that the statements of severity are associated with desirability of treatment rather than the type of treatment likely to be involved. Also the number of categories is reduced to four. These are:

0-7	Ideal occlusion or minimal malocclusion. No treatment need
8-17	Moderate malocclusion. Treatment elective
18-33	Severe malocclusion. Treatment desirable
>33	Very severe malocclusion. Treatment essential.

Validity was established by comparison of scores taken from 50 sets of study casts against rankings assigned to the study casts by 3 orthodontic consultants.

Inter- and intra-examiner reproducibility with MSI was found to be good. Agreement between pairs of consultants in ranking the study casts was not high when assessed with Kendall's correlation coefficient, this may have been due to this test of agreement being preferred where there are many tied ranks. As a large number of ranks were involved the likelihood of a large proportion of them being tied may have been reduced and thus lower correlations were reported, (Norman & Streiner, 1986). However Kendall's coefficient of concordance which is more suitable in situations with multiple raters did indicate good agreement, supporting the validity of the index.

The index was used in a survey of 793 children aged 9, 12, and 15 years and a high prevalence of malocclusion was found with 72% of 9-year-olds falling into the moderate malocclusion category and possibly requiring treatment. Overall a quarter of the children surveyed had previous experience of appliance wear although this group comprised more females than males. 23.5% of 15-

year-olds who had appliance therapy previously were found still to require treatment. This was lower than the 33% reported by Todd and Dodd, (1985).

3.3.9 Index Of Orthodontic Treatment Need (IOTN) (Brook & Shaw, 1989)

The Index of Orthodontic Treatment Need consists of two component parts, the Aesthetic Component and the Dental Health Component.

1 Dental Health Component (DHC)

The DHC is based on the guidelines suggested by the Index of the Swedish Health Board and comprises five grades. A dentition may be categorised into one of the five grades according to the worst feature of the malocclusion.

Details of the criteria for each grade are given in Appendix 1.

In the development of this index a literature review was done to find information which might allow rational cut-off points between grades to be defined for each occlusal trait potentially hazardous to the health of the dentition. This information, as previously indicated, is sparse and often controversial so that grades may have been defined in a somewhat subjective manner and the validity of the index may be open to question.

A refinement of the index in an attempt to provide further descriptive information concerning the nature of the "worst" occlusal trait of a dentition has been the addition of a letter code to the grade number. This may offer additional data of an epidemiological nature.

2 Aesthetic component (AC)

The AC is based on the Standardized Continuum of Aesthetic Need (Brook & Shaw, 1987) which was developed by the panel rating of dental photographs of 1000 subjects for dental attractiveness, and from the visual analogue scale so constructed the taking of a sub-sample of ten cases separated by equal intervals along the scale.

The dental photographs of this sub-sample was used to illustrate a ten point scale which represented a wide range of dental attractiveness, (Appendix 1).

Selection of AC grade for a particular dental configuration was designed to be according to position on a continuum of attractiveness rather than similarity with any specific feature demonstrated by any of the ten photographs.

In testing the Standardized Continuum of Aesthetic Need good inter-examiner agreement was found between five orthodontists and eighty per cent of all individual ratings were found to be within one scale point of the five judges' mean rating.

IOTN was intended to be used in the clinical setting, however the index has been adapted for use with study casts. Morphological features such as posterior crossbite may have a functional association, the presence of which cannot be determined from study casts. In cases such as this the functional feature is assumed to be present and the case is correspondingly graded. When used with study casts AC has good reproducibility, but correlation between panel attractiveness ratings from casts and from photographs is only moderate, (Woolass & Shaw, 1987).

IOTN has been used by Holmes (1992) in an investigation into the prevalence of unmet orthodontic treatment need amongst a sample of 955 12-year-old schoolchildren. A potential level of treatment need of 36.3% was found when individuals having a DHC greater than 3 or an AC greater than 5 were regarded as being in need of treatment. The index was found to be quick and simple to use.

The convenience of IOTN was confirmed by So and Tang (1993) in a comparison with the Occlusal Index, however the authors considered IOTN to be oversensitive to contact point displacement and missing teeth when used as an epidemiological tool.

3.4 Indices of treatment success

3.4.1 The evaluation of treatment results

Any discussion of the evaluation of treatment results should take into account not only the occlusal arrangement of the teeth achieved but also functional aspects, aesthetics both in the opinion of the orthodontist and the patient, potential stability and any iatrogenic damage which has occurred as a result of treatment, (Berg, 1991). However assessment of treatment success has largely focused on morphological aspects and how far final tooth positions differed from the ideal. This ideal is a hypothetical situation which may exist in the minds of orthodontists but which seldom occurs even in occlusions considered to be successfully treated, (Andrews, 1972).

The majority of assessments of treatment results have been through highly subjective anecdotal case reports where there may have been arbitrary selection of cases. The value of the information presented in this type of report is dependent on the method of selection, but where no analysis of the sample group is made then it is impossible to estimate the worth of the report.

Berg (1991) has stated that the method of selection of cases for evaluation is of fundamental importance and that the sample produced should be as representative as possible in order for generally applicable conclusions to be made.

Sample selection of cases for evaluation may be prospective or retrospective.

Prospective sampling which establishes samples prior to treatment does not allow for drop-outs, which may affect sample size, but also is not a guarantee against bias. Where different treatment modalities are being compared it is possible that subjects may be selected for whom the treatment in question may be particularly appropriate.

In retrospective sampling, cases where treatment was discontinued may be excluded and thus bias the sample towards more successful results. This would be significant in comparing treatment schemes where differing degrees of patient co-operation were required. This effect may be found where cases are selected at random but from consecutively finished cases.

3.4.2 Assessments of treatment results

Myrberg and Thilander (1973) carried out a retrospective assessment of a group of 1486 cases treated between 1958 and 1968. A classification comprising five grades ranging from "good" to "poor" and including a "no effect" category was used by two examiners who agreed in 85% of cases. In addition, the examiners gathered information concerning appliance type, duration of treatment, amount of relapse following treatment and patients opinion regarding efficacy of their treatment. These last two items of data were drawn from a small proportion of the main sample.

91% of cases were graded good or acceptable. Patients' views of the success of their treatment were often found not to concur with the gradings allocated by the authors.

Eismann (1974) used an index of occlusion which considered 15 morphological criteria to evaluate 200 cases treated with removable appliances.

In Eismann's index each of the 15 items under consideration was measured, either in millimetres, degrees of rotation or inclination, or number of teeth and each measurement was categorised and points awarded for that particular item. The scoring was based on the author's estimate of the need for treatment taking into account his opinion on the effect of the measured traits on aesthetics, function and possible harmful effects on general dental health. The author assumed that increased incisor irregularity would result in increased caries incidence. Allocation of scores was therefore arbitrary with little scientific justification.

Treatment success was indicated by the size of the difference between the pre- and post-treatment study cast scores.

It was felt that there was scope for this form of assessment to be used in comparison of treatment techniques, not only in effectiveness but also in duration and cost.

Gottlieb (1975) devised a simple system of grading considering 10 factors which he suggested were generally accepted criteria for orthodontic correction. This system provided a percentage achievement related to the factors requiring correction and the success of the correction. The grading was biased towards improvement as for each trait considered 5 points could be gained for complete correction whereas only 1 point could be deducted if the condition was worsened.

The Occlusal Index (Summers, 1966) has been used in studies involving comparison of different appliance types in terms of effectiveness.

As well as attempting to provide an objective assessment of 321 cases treated over a nine year period, Pickering and Vig (1976) wished also to correlate change in Occlusal Index scores with whether fixed or removable appliances were employed. In excluding cases not completed from the study it is possible that patient co-operation factors associated with the two different forms of treatment were not taken into account. The authors concluded that fixed appliance therapy produced a greater improvement and a slightly superior result than removable appliance treatment.

Tang and Wei (1990) came to the same conclusion in a similar study but mean pre-treatment OI score was significantly greater in their fixed appliance group than in their removable appliance group.

Elderton and Clark's use of the OI in an assessment of orthodontic treatment as carried out in the General Dental Services (1984) and the finding of only one third of cases being substantially improved has been noted previously.

Berg's report (1979) was unusual as not only were morphological aspects of occlusions considered but radiographs to assess root resorption and photographs were also examined. He graded his cases as either grade

A - fulfilled all criteria, or grade B - criteria not fulfilled. His sample of 264 cases were treated mainly with fixed appliances, 43.2% of which he felt were grade A.

Berg and Fredlund (1981) carried out an evaluation of 30 cases taken at random from each of two orthodontic practices. A treatment priority index developed by a local study group of Norwegian orthodontists was applied to study casts. Post-treatment casts had been taken at least one year out of retention. The index comprised 28 parameters involving linear measurements and judgements concerning occlusal relationships. A score of 0 indicated a normal or near-normal occlusion and this was achieved in 36 cases.

The authors felt that the evaluation of the degree of improvement achieved by a course of treatment seemed to be a realistic way of assessing a treatment's efficacy. Despite the shortcomings of morphologically based analyses and the difficulty in assessing psychological and aesthetic effects, the authors believed that "the evaluation of orthodontic treatment results based on morphological criteria will continue to be central to the discussion of treatment benefits, standards, cost and efficiency".

3.4.3 Peer Assessment Rating (PAR) (Richmond *et al.,* 1992a; Richmond, Shaw, Roberts & Andrews, 1992b)

The PAR index was designed to provide a single summary score for all the occlusal anomalies which might be found in a malocclusion. PAR was developed in conjunction with the British Orthodontic Standards Working Party to record malocclusion at any stage of development.

200 sets of study casts were examined and discussed in order to decide which occlusal traits should be assessed to arrive at an estimate of alignment of malocclusion. Eleven separate components were incorporated into the index:

1 Upper right segment
2 Upper anterior segment
3 Upper left segment
4 Lower right segment
5 Lower anterior segment
6 Lower left segment
7 Right buccal occlusion
8 Overjet
9 Overbite
10 Centreline
11 Left buccal occlusion

Details of the scoring system for each component are given in Appendix 2.

74 dentists participated in a validation study. This panel comprised 22 consultant orthodontists, 22 specialist practitioners, two members of staff of the Dental Practice Board of England and Wales, two junior hospital staff members, 15 General Dental Practitioners and 11 Community Dental Officers. This panel was intended to represent all the various groups carrying out orthodontic treatment.

The panel examined study casts of 272 cases drawn equally from cases treated at two dental schools, prior approval and non-prior approval cases from the DPB, and untreated cases from a study of 1000 Cardiff school children.

In order to arrive at appropriate weightings for each of the PAR components to best predict average opinion concerning the importance of each aspect, multiple regressions were made of the components of PAR against the panels mean score. The weightings which were found to give the highest correlations between subjective opinion and PAR score were as follows:

Overjet	6
Centrelines	4
Overbite	2
Right buccal occlusion	1
Left buccal occlusion	1
Upper anterior segment	1
Lower anterior segment	1

Examiners' subjective assessments of occlusal changes between pre-treatment and post-treatment study casts were subjected to discriminant analysis and a nomogram produced which categorized cases as "Worse/No different", "Improved" or "Greatly improved".

There seemed to be two systems operating in assessing improvement. For a case to be classified as "Improved" there needed to be a 30% reduction in total PAR score. For a "Greatly improved" classification there needed to be a reduction by at least 22 weighted PAR points.

As well as use of the nomogram classification, change in occlusion due to treatment could also be expressed as the change in number of weighted PAR points and as percentage change in PAR score.

The index was intended to be used with study casts rather than by direct measurement in a clinical setting, and as such the accuracy of an index score is dependent on the quality and condition of the models and their correct articulation.

Within the anterior segments, displacements are recorded between the contact points of adjacent teeth parallel to the occlusal plane. Buccal segment displacements are not recorded because of the probable low reproducibility due to the broad contact areas in molar and premolar teeth.

Crowding in the buccal segments is not specifically recorded, however crossbites which may arise because of crowding are recorded. Potential crowding in the mixed dentition is crudely assessed by measurement between the lateral incisor and the first permanent molar tooth and calculating space discrepancy using average mesio-distal widths. Only where the difference between measured space available and total average widths is greater than 4 mm is a score awarded and it is assumed to result in upper canine impaction and therefore scored in the anterior segment component. In this way only moderate degrees of potential buccal segment crowding are scored and there is little differentiation between this and more severe crowding or less severe crowding which may have treatment implications if space closure following premolar extractions is required.

In the Overjet and Overbite sections the measurement is taken for the tooth showing the most severe deviation. Although the presence of an anterior tooth in crossbite will add to the overjet score, the overjet score does not indicate the presence or not of an anterior crossbite.

Good inter-examiner and intra-examiner reproducibility has been found with this index, (Richmond *et al.,* 1992a). The index has also been shown to be reproducible when applied to holographic images of study casts although a greater degree of random error with hologram measurement was found compared to study casts measurement. However in comparing holograph PAR scores with corresponding study cast scores the scores obtained from holograms were significantly lower than those from direct study cast measurements, (McGuinness & Stephens, 1993).

Richmond, Turbill and Andrews (1993) found that it was possible to teach individuals with no dental qualification and a minimum of dental training to use PAR with a level of reliability comparable with dentally and orthodontically qualified personnel. The use of PAR by non-dental staff in clinical audit procedures and third party payment agencies was suggested as a means of economising in terms of wages and clinical time. However, other treatment effects such as soft tissue changes, aesthetic considerations, likely stability and iatrogenic damage also have a bearing on the quality of result achieved, (Berg, 1991) such that change in PAR score should not be regarded as the sum total of treatment changes.

The possibility of non-orthodontically trained staff using only PAR in decisions concerning remuneration should perhaps be viewed with disquiet by the speciality considering the present state of development of the index and the infinite variety of clinical problems. PAR is heavily biased towards change in overjet due to the weighting that component receives. It is not inconceivable that cases requiring treatment but where overjet change would be minimal would not be accepted for treatment for financial reasons. There might also be a tendency for treatment to be delayed until the established permanent dentition denying the patient any potential benefits which might be gained from interceptive treatment.

Knowledge of dental development and orthodontic procedures in the mixed dentition would also seem to be necessary to interpret change in PAR score correctly. Kerr, Buchanan and McColl (1993) have indicated that although PAR may be useful for assessing improvement due to treatment in established dentitions, in the assessment of mixed dentition treatments, where definitive treatment is not possible, and also in treatments with limited treatment objectives use of PAR is inappropriate. From a sample of 150 cases treated with removable appliances, in 14 out of 16 cases falling into the "Worse/no different" category treatment was designed to address only one specific feature of the malocclusion. In 10 of these cases the treatment objective was achieved as judged by precisely defined and agreed criteria.

The authors also found that change in PAR was closely correlated with pre-treatment PAR (r=0.79) and highlighted the fact that occlusions with an initial PAR score of less than 22 could not become "Greatly improved" under the nomogram system. By removing all cases with pre-treatment PAR scores of less than 22 the proportion of cases falling into the "Worse/No different" group was reduced by 38 percent. It was suggested that the term "Not significantly improved" might be a more accurate label for this category.

The nomogram might also be considered a coarse instrument for determining quality of result and success of treatment as it only allows classification into one of three categories.

The recent use of PAR in personal clinical audit has been reported by Fox (1993) and Richmond (1993). In an audit of 100 consecutively started cases Fox found the use of full fixed appliances type to be the only factor significantly related to change in PAR. It should be noted however that a similar average change in PAR to the full fixed appliance group was achieved by the functional appliance group although the small number of cases in that group did not allow the use of functional appliances to show significance. An examination of those cases which fell into the "Worse/No different" category added further weight to the consideration of PAR as inappropriate in mixed dentition and adjunctive orthodontic treatment.

Richmond (1993) proposed that a good standard of practice was achieved when the mean percentage reduction in PAR score was over 70 percent and the proportion of patients in the "Worse/No different" category was less than 5 percent of an individual clinician's workload. These figures are comparable to those achieved in four Norwegian orthodontic practices visited by the author and which were attributed to the high proportion of cases treated by full fixed appliances rather than any difference in workload as suggested by the British Orthodontic Standards Working Party, (BOSWOP, 1986). The standards proposed by the author were without reference to possible variation in case-mix which might have considerable effect on a practice's PAR profile due to the large overjet weighting and unsuitability for mixed dentitions and adjunctive therapy noted previously.

In a comparison of the effectiveness of seventeen hospital orthodontic departments assessed using IOTN and PAR, O'Brien *et al.* (1993) found mean percentage change in PAR due to treatment ranging from 50.9% to 77.9%. There was also great variation in initial need for treatment as measured by IOTN, ranging from 98.0% to 70.6% of cases in definite need of treatment. Treatment method varied widely, with one department using two arched fixed appliances in 80% of cases and another using removable appliances in 41% of cases.

The factors found to have the greatest effect in influencing treatment effectiveness were the choice of appliance and the level of operator training. Two arched fixed appliance therapy was found to be more effective than single arched fixed appliances which were in turn more effective than removable appliances. Consultants and senior registrars were more effective than clinical assistants and registrars. These two factors may not be independent. It is likely that cases with more severe initial malocclusions involving inter-maxillary discrepancies, and therefore larger pre-treatment PAR scores, would be treated with two arched fixed appliances, and would also be considered more complex, requiring an operator with a higher level of training.

In considering change in PAR and pre-treatment IOTN, variation between departments was found to be an overriding effect. This variation was not explained by the social status of the population served by each department and the authors concluded therefore that the "clinical attitudes and aspirations" of the staff were the cause of the differences.

The number of clinicians in each department was taken into consideration by the calculation of a standardised treatment activity for each department although it is not clear whether staff training levels within each department and case-mix, except with regard to type of appliance used, were allowed for.

A similar conclusion regarding the superiority of full fixed appliances over other treatment methods was reached by Richmond *et al.* (1993) in their study of 1210 cases treated within the General Dental Services of England and Wales. The use of upper and lower fixed appliances was shown to have the greatest influence on treatment changes as assessed by PAR and IOTN. The mean percentage reduction for full fixed appliance treatments was 71% compared with 49% for removable appliance cases and 39% for non-prior approval cases. The level of payment was also found to have a slight effect on the treatment outcome, a fee in excess of £350 improving the outcome over cases attracting a fee below £300. Treatments in the high fee group often involved the use of removable or functional appliances in addition to full fixed appliances. Once appliance type had been controlled for, neither the level of operator training nor the degree of specialisation of the operator had any significant effect on treatment outcome.

3.5 Indices of treatment complexity

3.5.1 Introduction

For planning of public health services and for management of resources it would be useful to possess an index which would indicate the level of operator training required for a particular case and would predict the likely difficulty and potential clinical commitment necessary in terms of duration of treatment and appliance type.

3.5.2 The concept of treatment difficulty and complexity

Difficulty of orthodontic treatment is a complex and multifactorial concept.

The factors which may be thought to have some bearing on treatment difficulty relate to the malocclusion, the patient, and to the treatment and may include the following:

1 Malocclusion factors

- Skeletal discrepancy
- Soft tissue configuration
- Unfavourable tooth inclinations
- Rotations
- Degree of crowding or spacing
- Ectopic teeth
- Tooth quality and prognosis

2 Patient factors

Age	Growth potential Co-operation
Co-operation	Oral hygiene Dietary control Appliance care Parental support School circumstances
Expectations	Both patient's and parents'

3 Treatment factors

Level of training and experience of operator

Treatment objective in terms of result quality *i.e.* nearness to ideal

Appliance type, including use of elastics and headgear

Extractions or non-extraction

Duration of treatment

The uncertainty inherent in some of the above factors *e.g.* growth, co-operation, together with the myriad possible interactions between the factors in these groups makes it unlikely that an index could be developed which would result in a simple yet meaningful summary numeric expression of overall difficulty in an individual case.

A definition of the term "difficulty" would be required before such an index could be developed. This might involve further definition of the ideal treatment plan and the quantifiable treatment objective for each case.

Difficulty implies concern with the likelihood of a desired treatment change being achieved. The difficulty of a particular malocclusion therefore varies with varying combinations of operator level and appliance type. For example, the same malocclusion which a consultant orthodontist employing a fixed appliance might regard as "easy" might be regarded as more difficult if treated with removable appliances by a less highly trained operator.

The term "complexity" may be more useful as it implies consideration of treatment by the most appropriate operator and most suitable appliance type.

In the opinion of Myrberg and Thilander (1973), before treatment is commenced the chances of obtaining an improvement in a dentition should be estimated, and that the clinical experience of the operator and the duration of treatment and retention, amongst other factors, are important in estimating the probability of improvement. The authors note that one way of evaluating a treatment outcome is to find out the patient's opinion of the treatment received, but aspects allied to this evaluation and which must also be considered include the initial demand for treatment and the duration of treatment.

The authors also cited estimated long duration as a frequent reason given for rejection of treatment.

Current assessment of operator appropriateness, type of treatment to be used, likely outcome, and duration of treatment is made entirely through the subjective clinical judgement of experienced consultants. This judgement will be greatly influenced by local treatment service factors. These may include availability of manpower and facilities, practitioner/population ratio, rate of referral from general dental practitioners, techniques employed within a department, and knowledge of the abilities of individual supporting clinicians and local practitioners. These factors vary regionally, (O'Brien & Corkhill 1990).

3.5.3 Assessments of complexity and duration

In a study by Brown *et al.* (1977), the likely treatment duration and difficulty in 50 cases was assessed by four similarly experienced and trained orthodontists through the examination of pre-treatment study casts. There was disagreement in more than half the cases examined which was explained by differing definitions as to what constitutes difficulty, the differences in treatment methods preferred by each orthodontist and the possible speculation on the part of the examiner as to morphological and functional aspects which may not have been discernible from the study casts.

A three category complexity grading scheme was used in an investigation into the change over time in the complexity of cases referred into the Bristol Orthodontic Department, (Stephens & Harradine, 1988). The cases were graded as either 1) suitable for treatment by a general practitioner with or without advice, 2) simple one arch fixed appliance therapy suitable for a clinical assistant level operator or 3) complex cases requiring specialist treatment including full multi-bracketed fixed appliances. The grading was entirely by subjective judgement of experienced clinicians. The authors found no difference in the complexity of referred cases between 1977 and 1985.

The Occlusal Index (Summers, 1966) was developed as a treatment priority index and as an epidemiologically useful index but included in its priority grading scheme is an indication of the specific appliance type which may be used in a case of a particular category together with an indication of the complexity of the required treatment.

Pickering and Vig (1976) felt that if the OI was to be used as an estimate of complexity additional weighting should be given to the enforced loss of one or more molars. Also incisors lost through trauma or caries should be scored as well as those congenitally absent.

When the mean pre-treatment OI score of cases treated with fixed appliances was compared with that of cases treated with removable appliances, no significant difference was found which may have been an indication that pre-treatment OI score used exclusively was not a good predictor of treatment complexity in terms of appliance type.

Devenish (1978) included assessments of canine and incisor angulation together with the OI in comparing OI assessments of difficulty with subjective panel judgements of difficulty. The upper incisor angulation assessment was found to be reproducible and it was felt that this feature could be included as an item in a treatment difficulty index. A higher correlation was found between subjective difficulty judgement and measured overjet than difficulty and OI score. However the sample size in this study was small at 31 cases. Also the level of disagreement found between clinicians in subjectively assessing likely treatment difficulty has been noted previously.

An attempt has been made by O'Brien to adapt the PAR index to use as an index of treatment complexity. Cases represented by 198 sets of study casts were rated for likely difficulty from 1(easy) to 5(very difficult) by an eleven strong panel of orthodontists. The difficulty ratings were correlated against unweighted PAR scores for the casts and a correlation of 0.54 was obtained.

Weightings were then applied to the different constituent parts of the PAR index and a higher correlation of 0.63 achieved. It should be noted that this study was carried out in the United States and it is possible that perceptions of treatment difficulty held there may differ from those held in this country due to varying treatment philosophies.

Regarding treatment duration, Myrberg and Thilander (1973) found no difference in the duration of courses of removable appliance therapy compared with courses of fixed appliance therapy. However only 34% of the sample in this Scandinavian study completed treatment within 12 months whereas 62.8% of cases in a contemporaneous British study finished treatment within that time, (Rose, 1974). This may be indicative of differing treatment modalities between the two groups and possibly differing treatment objectives. 40% of the Myrberg and Thilander group had fixed appliances compared with only 0.8% in the Rose sample.

Rose found the mean active treatment duration to be 13.1 months involving an average of 11.7 visits with a mean of 1.5 appliances being used to complete a case. The treatment result was deemed satisfactory in 74% of subjects. It is not clear whether the calculation of mean duration included cases which were curtailed or only completed cases. Inclusion only of cases which were completed whether satisfactorily or not, may have biased the mean duration towards a lower figure as the likelihood of a patient's withdrawing from treatment might increase as treatment time increases and the patient gets older, (Haynes, 1982).

The mean figures calculated in this survey are similar to those of Hooper (1966) at 12.4 months but low compared with those of Sheiham, Hobdell, Vig and Griffiths (1971) at 16.9 months and Fletcher (1958) at 20 months.

In an investigation of the influence of serial extractions on growth in Class I crowding cases, Ringenberg (1967) found the average duration of active treatment of an experimental group who had previous serial extractions to be 12.7 months. This compared to an average of 19 months of active treatment in a control group. It should be noted that the control group did appear to differ from the experimental group at least in group size and sex ratio.

Comparing the treatment of a fixed appliance group with a removable appliance group, Tang and Wei (1990) found an average treatment duration for the removable appliance group to be 13.4 months but duration was highly variable with a large range of from 3.1 to 23.7 months and a standard deviation of 10.3 months. The corresponding figures for the fixed appliance group were a mean of 20.2 months, a smaller range from 15.7 to 24.7 months and a standard deviation of 4.5 months. This indicates that although treatment duration with removable appliances tended to be shorter than with fixed appliances it was also less predictable.

The difference in duration might not have been entirely due to type of appliance used. The removable appliance cases had largely been treated by undergraduate students whilst the fixed cases had been treated by postgraduates. Also the severity of the pre-treatment malocclusions as measured by the Occlusal Index were greater in the fixed group than the removable group. This is in contrast with the sample in Pickering and Vig's study where no difference was found in pre-treatment severity between removable and fixed groups. It was suggested that the differences between the two studies may be attributable to differences in background and training of the authors and also in the facilities available.

A statistically significant difference in treatment duration (2 months) between a Begg and an Edgewise group was also found but this was not felt to be clinically significant. The average number of appliances used in the removable group was 2.38.

Fink and Smith (1992) attempted to evaluate causes of variation in the duration of treatment using the case records, lateral cephalometric radiographs and pre- and post-treatment study casts of 118 patients taken from 6 specialist orthodontic practices.

The initial severity of the malocclusion and degree of dental change through treatment was assessed with the Handicapping Malocclusion Assessment of Salzmann.

Cephalometric measurements included angle ANB, mandibular planes angle and upper/lower anterior face height ratio.

Information taken from individual case records included; treatment duration, number of visits, number of broken appointments, patient's age at start of treatment, patient's sex, whether headgear was worn or not, number of impacted canines, number and type of teeth extracted, type and length of retention and whether orthognathic surgery was part of the treatment. From this data was calculated the frequency of appointments, the frequency of broken appointments and the ratio of appointments attended to broken appointments.

Further information concerning the experience of each practitioner, number of staff, number of chairs and patient throughput was ascertained.

Multiple regression equations to predict treatment duration were calculated with three separate sub-sets of independent variables. The sub-sets were one of patient characteristics, one of treatment characteristics and one of practice characteristics. The significant independent variables from each group were then used to calculate a single general regression equation to predict treatment duration.

Analysis of only patient characteristics identified angle ANB, mandibular planes angle and pre-treatment Salzmann index as being significantly related to treatment duration. However multiple correlation was low (r=0.34).

Analysis of only treatment characteristics identified number of extracted premolars, number of broken appointments, and the use of headgear as significantly predictive of treatment duration.

No practice characteristics were found to be significant contributors to the initial regression model.

The six significant variables identified from the sub-sets of variables were then combined in a single general equation. This was a five step equation which included the number of premolars extracted, number of broken appointments, pre-treatment mandibular planes angle, pre-treatment ANB angle and pre-treatment Salzmann index (r=0.5).

Number of premolars extracted was found to be the single variable which most explained differences in treatment duration between patients. Each extracted premolar was found to add 0.9 months to the treatment duration. In this study non-extraction cases averaged 21.95 months, two premolar extraction cases averaged 25.0 months and four extraction premolar cases averaged 26.18 months.

The next most important variable was broken appointments. This was felt to closely equate with patient compliance and therefore was believed to be indicative of other manifestations of poor co-operation which might also affect treatment progress.

Increased treatment duration was positively correlated with increased angle ANB and increased Salzmann index, showing longer treatment time with increased severity, but negatively with mandibular planes angle. It was felt that this might be a reflection of the greater time required for treatment of deep overbite cases.

It is surprising that no practice characteristics were found to contribute to the initial regression model in view of the fact that the mean treatment durations predicted from the general regression equation for each practice differed by up to seven months from the actual mean treatment durations. This might suggest there may have been be practice characteristics involved which were not taken into account.

It was felt that the Salzmann index may not have been adequately sensitive to fine differences in detail of finished cases and that the differences in treatment duration may have been as a consequence of differing times devoted to finishing procedures.

The authors stated that a small set of variables had been found to account for a substantial portion of the variation in treatment duration. However, the value of R^2 for the stepwise regression including combined patient and treatment characteristics was reported as 0.249, which indicates that less than one quarter of the variation in treatment duration could be explained by the regression model. There was no indication as to how acceptable the assumptions of regression were in this case.

In a study limited to removable and functional therapy, Kerr *et al.* (1993) applied the PAR index to pre- and post-treatment study casts of 150 consecutively completed cases. The cases were divided into an upper removable appliance group and a functional/headgear group. Stepwise multiple linear regression was use to identify potential predictors of change in PAR score and of duration of treatment for both groups. The variables entered into the regressions were sex, age, number of malocclusion features treated, number of missed appointments, number of appliance breakages, number of appliances used and pre-treatment PAR score.

For change in PAR in the upper removable appliance group the only variable found to be predictive was pre-treatment PAR score (R^2=0.677) whilst in the functional/headgear group pre-treatment PAR score and sex contributed to the regression equation (R^2=0.712). It was found that, on average, females' change in PAR would be 3.13 points greater than in males.

For duration of treatment in the functional/headgear group, pre-treatment PAR and number of appliances used, *i.e.* whether or not three or more were used, were explanatory variables (R^2=0.603).

Only the number of appliances used was entered as an explanatory variable for the upper removable appliance group (R^2=0.559). The average treatment durations for single appliance, 2-appliance and 3-appliance treatments were found to be 6.4 months, 13.5 months and 19.8 months respectively. The authors suggested that using a second appliance would add between 4.4 and 9.8 months to the treatment time whilst using a third appliance would add between 3.1 and 9.5 months. It should be noted that there were only 7 subjects in the 3 or more appliances category.

Pre-treatment DHC was found to be a potential predictor of treatment duration but the variation in duration explained by a regression equation which included pre-treatment DHC and number of appliances was reduced to 47.5% (R^2=0.475).

Vig *et al.*(1990) also employed stepwise regression in their study of orthodontic treatment duration in extraction compared with non-extraction treatments.

438 subjects were selected from five local orthodontic practices. The five practices were selected as being at the extremes of the spectrum of local practices regarding the ratio of extraction to non-extraction treatments. The mean extraction rate across all practices was 39%, with a range of from 5% to 88%.

Student's *t* test did not indicate any significant difference between mean duration of treatment with extractions and mean duration of treatment without extractions when data from all the selected practices were analysed together. A significant difference between treatment types was found in one particular practice but in this case proportion of non-extraction treatments was particularly small. The authors regarded this practice as an outlier as it also had the lowest mean treatment duration. A significant difference between non-extraction and extraction groups indicated by the *t* test would in any case have to be interpreted with caution as it would be difficult to ensure that the groups differed only in respect of the single independent variable observed.

Stepwise analysis was used to investigate the relationship between duration of treatment and a number of independent variables.

These included patient gender, molar relationship, arches treated, age at start of treatment, number of phases of treatment, extraction or non-extraction, change in overbite, and the practice where the case was treated. Variables concerning patient co-operation such as missed appointments or requests for early appliance removal were not considered although these might be regarded as being potentially important factors in duration of treatment.

Also change in overbite and change in overjet were the only variables included which pertained to treatment success and in 141 cases this information was unobtainable. There was no data gathered which indicated the severity of the initial malocclusion or the dental arrangement at the end of treatment.

Regressions were carried out with all the data and also with the data from each individual practice.

Number of treatment phases, arches treated, molar relationship, age at start of treatment, and whether treatment was extraction or non-extraction were the variables found to be significantly correlated with duration of treatment in this study.

It was suggested that a regression model developed for an individual practice might aid the orthodontist in making more accurate predictions of treatment duration for individual patients than a model would which used pooled data.

The authors concluded that there were differences in treatment duration within individual practices between extraction and non-extraction groups. Extraction treatments were stated to take longer than non-extraction treatments. This conclusion was made despite there being no significant differences demonstrated and in the face of wide variation in durations within groups as indicated by the relatively large standard deviations.

The need for further research was expressed, using samples which were larger and more representative of the broad spectrum of orthodontic practice.

3.6 <u>Other occlusal indices</u>

Other indices have been produced which have been designed for limited and specific purposes for example the STRAIT index of Lau, Griffiths and Shaw (1984) to measure the alignment of individual teeth.

Chapter 4

Review of the literature: conclusions

Increasing accountability concerning the provision of unnecessary treatment, use of resources and the benefit gained as a result of orthodontic treatment has led to the development of a number of indices and systems of measurement designed to be objective and reproducible assessments of occlusion.

No single system has been found which adequately measures all aspects of occlusion and treatment of malocclusion in all circumstances. There has been particular difficulty in measurement and prediction of likely treatment complexity and clinical commitment required.

The Index of Orthodontic Treatment Need has now been widely accepted as an valid assessment of treatment need.

A recently developed index, PAR, has been useful in assessing comprehensive treatment of malocclusion in established dentitions, though has been shown less successful where mixed dentition and adjunctive treatments are concerned. Investigations using PAR into factors related to change in occlusion through treatment and the quality of treatment result have indicated that the use of fixed appliances was the most significant factor.

With regard to duration of treatment a wide range of average figures have been reported though it has been suggested that a regression model developed for an individual practice might aid the clinician in prediction of treatment duration for a particular patient.

Up to the time of writing there has been no published study which has endeavoured to identify features which could be used to quantitatively predict change in occlusion due to treatment, standard of result and duration of treatment within a fixed appliance group.

A study of factors associated with orthodontic treatment outcome

Section B

AIMS OF THE STUDY

Chapter 5

Aims and objectives

The purpose of this study was to describe a sample of cases treated within the orthodontic department of a teaching dental hospital and their orthodontic treatment outcomes. The study also intended to identify which treatment factors, including patient characteristics and occlusal features, might be associated with treatment outcome as assessed by the PAR index and by treatment duration.

The aim was to formulate linear regression models which could explain the variability in treatment outcome for a full fixed appliance treatment group and also for a combined group of removable, limited fixed and removable/limited fixed appliance combination cases. Further regression models were to be constructed using only that information which would be available at the start of treatment. The variables entered into these equations could be used to quantitatively predict outcome of treatment.

The objectives of the study were therefore:

1 To identify factors which were associated with post-treatment total PAR score, including factors potentially predictive for post-treatment PAR score.

2 To identify factors which were associated with change in PAR score through orthodontic treatment. Also to identify potentially predictive factors for the change in total PAR score through treatment.

3 To identify factors which were associated with the duration of active treatment, including pre-treatment factors which might be predictive of treatment duration.

A study of factors associated with orthodontic treatment outcome.

Section C

MATERIALS AND METHODS

Section contents

Chapter 6

Materials

6.1 Material for the study

The material for the study comprised the pre-treatment and post-treatment study casts and case notes of 161 patients treated within the Orthodontic Department of Glasgow Dental Hospital and School.

In accordance with the terms of the United Kingdom Consumer Protection Act, 1987, records of orthodontic patients treated at the Glasgow Dental Hospital are retained for at least 10 years plus 1 year to allow for the possible serving of a writ.

Study cast boxes in the main department box room which have not been accessed for one year are culled serially and transferred to a secondary storage area for long term storage.

Boxes are catalogued by name of patient and box number only.

6.2 Selection of sample

In order to establish the limits of the study sample it was decided to locate all cases held in secondary storage where the pre-treatment study casts had been taken during 1985.

The year 1985 was chosen for the following reasons:

1 Study cast assessment and data collection took place between January and April 1992. At the time of data collection a period would have elapsed since commencement of treatment to allow completion of

treatment, retention and possible long term post retention review.

2 A preliminary examination of the stored boxes indicated that there would be an adequate number of cases from that year for inclusion in the study.

3 Adequate time should have elapsed to ensure that all the relevant material had been transferred from the main box room to the secondary store.

4 Temporary removal of material from the records storage systems would not hinder the day to day running of the Orthodontic Department.

Cases were excluded from the sample in the following circumstances:

Where post treatment study casts were missing

Where damage to study casts would render measurement impossible

Where there was no reasonable indication as to the articulation of the casts

Where case notes could not be located

Where case notes were incomplete

Where not all treatment was carried out within the Orthodontic Department, for example cases transferred from outwith the department

Study casts and case notes were examined in order of their being found. As no particular system of cataloguing or ordering had been employed during storage it was reasonable to assume that the material was examined in random order.

Chapter 7

Methods

7.1 Case note data collection

The following information collected from the case notes of each selected case was entered on the case notes data capture form (Figure 1).

Patient details: Including sex and date of birth.

Oral hygiene: Cases were classified as "Good"(1), "Fair"(2), "Poor"(3) and "No indication"(4). The assessment of oral hygiene status was entirely subjective and dependent on the opinions of the diagnosing clinicians. Furthermore, where a term indicating the level of oral hygiene had been used other than " Good", "Fair", or "Poor" (for example, "could be better" or "needs improving") interpretation was required in order to place the case into one of the available categories.

Teeth missing: Teeth missing through developmental causes, excluding missing permanent third molars and permanent teeth which had been extracted prior to treatment planning, were recorded using the FDI system of notation, (W.H.O., 1987).

Ectopic teeth: Teeth noted in the case notes as being ectopic were recorded using the FDI system of notation.

Operator type: Operator type was categorised into "Consultant"(1), "Senior Registrar"(2), "Postgraduate student"(3) and "Undergraduate student"(4). Members of staff with a postgraduate orthodontic qualification and who were not consultants were grouped under category 2.

Date Start of Treatment: This was defined as date of insertion of the first removable appliance, or in the

CASE NOTES DATA FORM

surname

initials

SEX

D.O.B.

UNIT NUMBER

BOX 1

BOX 2

ORAL HYGIENE
1 Good 2 Fair 3 Poor 4 No indication

MISSING TEETH
Developmental
Previously xtd

ECTOPIC TEETH

OPERATOR TYPE

DATE START OF TREATMENT

EXTRACTIONS

TREATMENTS EMPLOYED

APPLIANCES

A. REMOVABLES

UPPER
Number of appliances (Not including lost/broken)
Number of breakages (Including lost)

LOWER
Number of appliances (Not including lost/broken)
Number of breakages (Including lost)

REMOVABLE/fixed combination Y/N

FEATURE TREATED

TREATMENT OBJECTIVE ACHIEVED

B. FUNCTIONALS

APPLIANCE TYPE

NUMBER OF BREAKAGES

C. FIXED

1 - Multiband 2 - "miniband"
Multi = attachments on 3 or more teeth in every quadrant

APPLIANCE TYPE
UPPER
LOWER

NUMBER OF A/W CHANGES

NUMBER OF BRACKETS AND BANDS REPLACED
(damage)

HEADGEAR USED Y/N?

DATE END OF **ACTIVE TREATMENT**

PLANNED APPOINTMENTS
(active treatment)

BROKEN/CANCELLED APPOINTMEMTS
(active treatment)

CASUAL VISITS (active treatment)

T/P ALTERED - LACK OF PROGRESS Y/N

RETENTION APPLIANCE

UPPER
Type Date end of retention
permanent retention planned

LOWER
Type Date end of retention
permanent retention planned

DATE at last visit

PLANNED APPOINTMENTS
(retention)

BROKEN/CANCELLED APPOINTMEMTS
(retention)

CASUAL VISITS (retention)

FORMAL DISCHARGE Y/N

Study casts **Treatment Stage**

1..
2..
3.................................... ....
4.................................... ....
5.................................... ....

1 = pre-treatment 2 = mid-treatment 3 = < 3mths after active treatment 4 = < 3mths out of retention 5 = < 1yr out of retention 6 = 1yr+ out of retention

Figure 1

case of fixed appliances, the date of insertion of separators or of first cementation of fixed components which ever was the earlier.

Extractions: Teeth extracted as part of the orthodontic treatment plan were identified and recorded using the FDI system of notation. Extractions of deciduous teeth and of third permanent molars were excluded.

Treatments Employed: A code number string for all treatment types used in each case was entered to facilitate data processing prior to analysis.

1 Extractions
2 Removable appliances
3 Fixed appliances
4 Functional appliances
5 Minor oral surgery
6 Restorative procedures

Removable appliances: The number of upper and lower removable appliances inserted and the number of breakages necessitating repair or replacement of appliances was recorded. It was noted whether or not an appliance had been worn concurrently with fixed components.

Removable appliances are restricted mainly to tipping movements and only a small number of teeth may be moved simultaneously. Removable appliance therapy objectives may be limited to the correction of a single feature or few features of a malocclusion rather than directed towards the attainment of a morphologically "ideal" occlusion.

The decision as to which features had been tackled specifically in each case was based on the stated treatment plan, appliances used and entries in the case notes regarding appliance adjustments and treatment progress.

Judgements as to which treatment objectives had been achieved were subjective and depended on the opinion of the examiner of the post treatment study casts and case sheet entries.

Functional appliances: Appliance type and number of breakages were recorded.

1 Bionator
2 Frænkel
3 Andresen
4 Twin-block
5 Other

Fixed appliances: Fixed appliances were divided arbitrarily into two groups, "multi-banded" where attachments were place on more than three teeth in every quadrant, and "mini-banded" where there were fewer than three teeth banded or bonded in every quadrant.

Appliance type for each arch, total number of archwire changes and number of brackets and bands replaced were noted.

1 Edgewise
2 Preadjusted
3 Begg
4 Beddtiot
5 Other wide
6 Other narrow

Headgear: Whether or not headgear was worn was noted.

Date end of active treatment: This was defined as the date of removal of remaining fixed components excluding fixed retaining appliances, or the date of insertion of a removable retaining appliance, or of withdrawal or deactivation of the final removable appliance.

Planned appointments: The number of planned appointments during active treatment was recorded.

Broken appointments: The number of occasions where the patient failed to attend for an appointment or had cancelled an appointment at short notice was recorded.

Treatment plan altered: Departures from the planned course of treatment were invariably due to co-operation factors such as inadequate oral hygiene, failure to wear elastics, failure to wear headgear as instructed and lack of care with appliances resulting in frequent damage.

Retention: Appliance type, either removable, fixed or functional and date of withdrawal of retaining appliance.

If a retaining appliance had been found to be broken or lost with no reasonable estimate of the date of loss or breakage and no replacement or repair was made, then the date of end of retention was taken to be the date of the previous appointment.

Information included the number of planned appointments, appointments which were missed or broken at short notice, and number of casual visits up to the date of the last visit.

Date of last visit: This was taken as the date of the last entry in the patient's case notes.

Formal discharge: Cases which were noted to have been discharged were done so in order to differentiate from cases where treatment had been terminated by failure of the patient to attend a review appointment.

Repeat Y/N
STUDY MODEL DATA FORM

BOX 1

BOX 2

STUDY MODELS SET 1

DATE

DEVELOPMENT STAGE

INCISOR CLASSIFICATION

OVERJET mm

INCISAL CROSSBITE Y/N (any tooth)

SPACE ANALYSIS (MIXED DENTITION)

UPPER right left
LOWER right left

IOTN GRADE DHC AC

PAR SCORE U L RBO LBO OJ OB CL

INCISOR ANGULATION (< 100. >110 in CL III)
CANINE ANGULATION (< 90. >100 in CL III)
MULTIPLE ROTATIONS (2+ teeth > 25°)

STUDY MODELS SET 3

DATE

DEVELOPMENT STAGE

INCISOR CLASSIFICATION

OVERJET mm

INCISAL CROSSBITE Y/N (any tooth)

SPACE ANALYSIS (MIXED DENTITION)

UPPER right left
LOWER right left

IOTN GRADE DHC AC

PAR SCORE U L RBO LBO OJ OB CL

PAR DEGREE OF CHANGE (NOMO)

STUDY MODELS SET 2

DATE

DEVELOPMENT STAGE

INCISOR CLASSIFICATION

OVERJET mm

INCISAL CROSSBITE Y/N (any tooth)

SPACE ANALYSIS (MIXED DENTITION)

UPPER right left
LOWER right left

IOTN GRADE DHC AC

PAR SCORE U L RBO LBO OJ OB CL

PAR DEGREE OF CHANGE (NOMO)

STUDY MODELS SET 4

DATE

DEVELOPMENT STAGE

INCISOR CLASSIFICATION

OVERJET mm

INCISAL CROSSBITE Y/N (any tooth)

SPACE ANALYSIS (MIXED DENTITION)

UPPER right left
LOWER right left

IOTN GRADE DHC AC

PAR SCORE U L RBO LBO OJ OB CL

PAR DEGREE OF CHANGE (NOMO)

Figure 2

7.2 Study cast data collection

The following information collected from the study casts of each selected case was entered on the study casts data capture form (Figure 2).

Development stage: Development stage was categorised into three groups (modified from Summers, 1966):

1. Early mixed dentition - up to emergence of the last permanent incisor.
2. Later mixed dentition - up to the eruption of the last successional tooth into occlusion.
3. Permanent dentition

Incisor classification: British Standards Incisor Classification (British Standards Institute, 1983).

Overjet: The overjet component of PAR does not necessarily indicate the size of overjet present but may include edge-to-edge and anterior crossbite relationships from the canines anteriorly. Overjet was measured as the distance in the occlusal plane radial to the dental arch from the midpoint of the incisal edge of the most prominent central incisor to the labial surface or tangent to the labial surface of the corresponding lower incisor.

Incisal crossbite: This was recorded where any incisor tooth was found to be in an edge-to-edge or crossbite relationship to permit correct interpretation of the PAR overjet component.

Space analysis: In mixed dentition cases both PAR and IOTN record potential crowding only when the space available for premolars and canines is equal or less than 17 mm in the lower arch and 18 mm in the upper arch and is regarded as severe enough to be graded "5i" and to score 5 PAR points as an impaction. Lesser degrees of potential crowding which may give rise to an increased score in the permanent dentition through transverse buccal discrepancies or anterior segment contact point discrepancies are not allowed for.

In mixed dentitions in this study the space available between the distal surface of the permanent lateral incisor and the mid point of the mesial surface of the first permanent molar for upper and lower arches on both left and right was measured with a Boley gauge.

IOTN: Both the Aesthetic Component rating and the Dental Health Component grade and descriptive code letter for IOTN were noted. Details of the index are given in Appendix I.

IOTN scores were awarded mainly on the evidence of study casts although the presence or absence of unerupted teeth was confirmed by reference to the case notes or available radiographs. The monochrome Aesthetic Component photographs intended for use in the assessment of study casts were available during the examinations. The conventions pertaining to the use of IOTN with study casts were observed such that the presence of crossbites of any kind scored 4c, thus assuming the presence of a mandibular displacement greater than 2 mm, and overjets between 3.5 mm and 6 mm were graded 3a disregarding lip competence.

A single IOTN ruler was used throughout the period of data collection to eliminate the effect of any variation in length between rulers which might have occurred during manufacture.

The examiner had previously been calibrated in the use of IOTN. Further details concerning calibration are given in Appendix III.

PAR: Unweighted scores for the seven individual elements of PAR were entered on the data capture form.

A single PAR ruler was used throughout the period of data collection to eliminate the effect of any variation in length between rulers which might have occurred during manufacture.

The examiner had previously been calibrated in the use of PAR. Further details concerning calibration are given in Appendix III.

In addition to the above measurements and descriptions, pre-treatment study casts were assessed for upper incisor angulation, upper canine angulation and number of rotated anterior teeth.

Incisor angulation: Incisor angulation was considered to be a potential factor in treatment complexity. Overjet correction in cases where teeth can simply be tipped into position may be more readily achieved than in cases where bodily movement and root torqueing is necessary.

The angulation of the upper incisor teeth is not recorded by PAR or by IOTN except in that excessive proclination may contribute to the size of overjet and that proclination or retroclination may influence the assessment of the Aesthetic Component.

Incisor angulation was crudely categorised as being less than 100° or equal to or greater than 100° to the occlusal plane in Class I and Class II cases or greater than 110° or equal to or less than 110° to the occlusal plane in Class III cases. The "worst" central incisor *i.e.* most retroclined in Class I and II or most proclined in Class III, was recorded with the aid of a clear plastic guide marked with lines at the appropriate angles.

Canine angulation: This was categorised in a similar manner to incisor angulation noting upper canines retroclined at less than 90° to the occlusal plane in Class I and II cases and more than 100° in Class III cases.

Multiple rotations: A dentition was deemed to have multiple rotations where two or more anterior teeth were rotated at more than 25° to the line of the dental arch.

7.3 Data processing

Information from the two sets of data capture forms was stored in two computer database files sorted and related by box number. DBase III Plus[1] programs were written to calculate and store patients' ages at start of treatment, duration of active treatment and duration of retention expressed in months. Further programs calculated and stored total PAR score, percentage change in PAR score and categorisation into "Worse/No different", "Improved" and "Greatly improved" groupings for each case having applied the appropriate weightings to overjet, overbite and centreline.

[1]Ashton-Tate UK Ltd. Oaklands, Bath Rd. Maidenhead

Chapter 8

Data analysis

8.1 Categorisation of variables

It was possible to group the variables under scrutiny into patient factors, treatment factors and outcome variables.

8.1.1 Patient factors:

These were further sub-divided into personal factors, occlusal factors and co-operation factors:

1 Personal factors

- Sex
- Age at start of treatment

2 Occlusal factors

- Incisor classification
- Developmental stage
- Pre-treatment PAR
- Pre-treatment IOTN of grade 4 or 5
- Size of overjet
- Presence of anterior crossbite
- Incisor angulation
- Canine angulation
- Presence of rotations
- Presence of ectopic teeth

3 Co-operation factors

- Oral hygiene
- Number of broken appointments
- Number of removable appliances broken or lost
- Number of bands/bonds dislodged or archwires broken

Whether or not the original treatment plan was altered

8.1.2 Treatment factors:

Appliance type

1 Full fixed appliances
2 Removable appliances only
3 "Mini-fixed" appliances
4 Removable + "Mini-fixed"

Operator type

Extraction pattern

1 Non-extraction
2 Extraction of a premolar from each quadrant
3 Where loss of a first permanent molar was included in the extraction pattern
4 Other

Number of archwires used

Number of removable appliances used

Whether or not headgear was worn

Number of appointments during active treatment

Duration of active treatment

8.1.3 Outcome variables:

Post-treatment PAR

Change in PAR score

Nomogram change in PAR

Duration of active treatment

Number of appointments during active treatment

8.2 Preliminary analysis

Most of the statistical analysis of the data in this study was performed using the SPSS for Windows v5.0.2[2] computer package.

Preliminary analysis of the data was by direct examination and investigation of the frequency and distribution of the variables included for study, firstly for the total sample of all 156 cases and then with the sample divided into sub-groups according to type of treatment employed. Chi-square testing for categorical data and analysis of variance for interval data was performed to explore any associations between the variables and the treatment type groupings. A Scheffé multiple comparisons procedure was used where appropriate to compare groups in analysis of variance. This is a more severe test of significance between groups than calculating the least significant difference by t-testing between pairs of groups or other commonly used criteria.

In order to produce groups suitable in size for regression analysis the non-full fixed appliance groups were combined into an "Other" treatment type group.

This grouping was made on clinical grounds in that the use of full fixed appliances might be regarded as being maximum intervention towards an ideal result, controlling the positions of most or all of the teeth. On the other hand the alternative treatments might be regarded as less than maximum intervention, being adjunctive, interceptive, or possibly compromise measures where the positions of a limited of teeth were controlled. This grouping does not mean to imply that the treatment used for a particular case was inappropriate.

---

[2]SPSS Inc., 444 North Michigan Avenue, Chicago

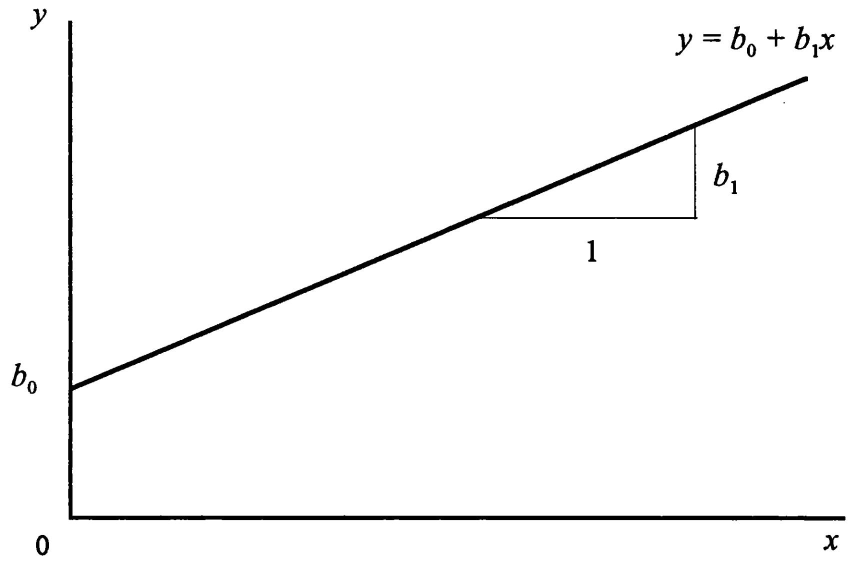
Furthermore the use of full fixed appliances has previously been differentiated from other forms of treatment in terms of quality of result, and has been identified as the most influential factor towards the benefit gained from treatment, (O'Brien *et al.*, 1993).

8.3 <u>Regression analysis</u> (Draper & Smith, 1981)

Linear regression analysis was employed to identify possible predictors of treatment outcome and duration from among the variables examined.

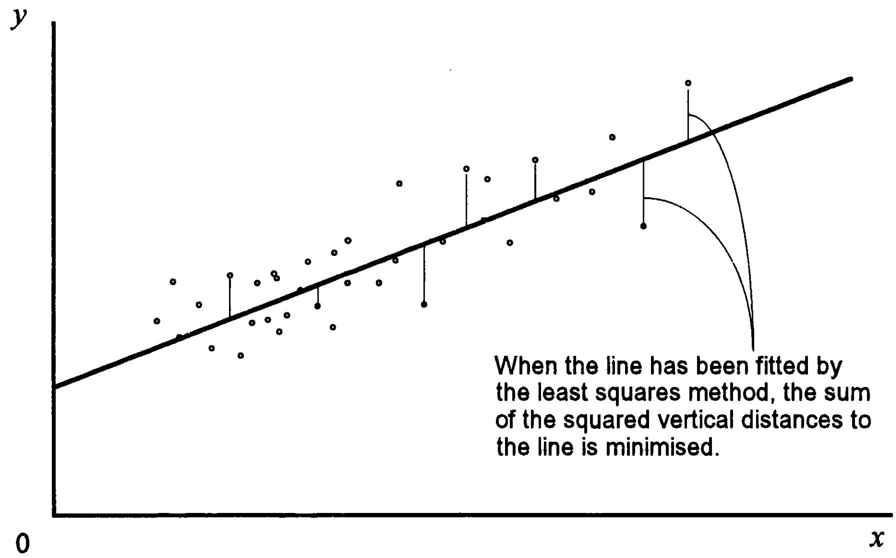
8.3.1 Simple linear regression

Simple linear regression involves the construction of a straight "line of best fit" to relate an independent or predictor variable, x, to a dependent or response variable, y.



The equation of a line is $y = b_0 + b_1x$ where b_0 is the intercept and b_1 the slope of the line. The slope is the increase in y for a unit increase in x.

In reality two variables may show a linear relationship however the plotted points will not fall precisely on a straight line but be clustered around it. In fact many lines might be drawn through the data points but the "least squares" method selects a line such that the sum of the squared vertical distances of the plotted points of y and x to the fitted line is minimised. This "line of best fit" could be regarded as that which best describes the association between the two variables.



This line may be used to predict values of y for given values of x but predictions should be made only for values of x within the range of the original data and can only be made if the linear model is the true picture. The starting point is to provisionally consider this to be the case, whereby the following first order mathematical model, pertaining to the population from which the sample x's and y's were derived, can be written:

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$

where β_0 and β_1 are the population values for the intercept and slope, which are estimated from the sample values, and e_i, usually called the error, is the difference between the observed value and predicted value of Y_i at the point X_i.

Underlying the method of linear regression are two assumptions. Firstly, for any value of the independent variable the dependent variable is a random variable with a normal distribution. Secondly the degree of scattering about the regression line is the same regardless of the value of the independent variable.

These assumptions can be examined by study of the residuals of the regression. These are the differences between values for the dependent variable predicted from the regression and the corresponding observed values and are estimates of the true errors e_i.

The characteristics of the residuals may be revealed by plotting the predicted values against the residuals where a broad band of apparently random points might indicate that the linear model was appropriate and that the assumption of equal variance was met.

The normality of the residuals may be examined by construction of a histogram of observed frequencies, also by the production of a normal probability plot. This can be done by either plotting the observed cumulative distribution of the residual values against the expected cumulative distribution (*i.e.* a P-P plot) or by plotting the values of the residuals against the expected values given a normal distribution. The expected values are based on the number of cases in the sample and their rank order. In either type of plot a straight line is produced if the residuals follow a normal distribution.

Norusis (1993) has stated that in regression analysis it is rare for assumptions not to be violated in some way but by careful examination of residuals it is possible to gauge the worth of the analysis and realistically interpret the results.

The above graphical methods of examining residuals are indications of the goodness of fit of the model. The coefficient of determination, R Square, (R^2) is also a measure of the goodness of fit of the model. This is the square of the correlation coefficient between the observed value of the dependent variable and its predicted value. If there is no linear relationship between x and y then $R^2=0$. If the relationship between the two variables is precisely linear then $R^2=1$.

R^2 may also be interpreted directly as describing the proportion of the variation in the dependent variable which is explained by the model.

As a measure of the goodness of fit of a model to the population, R^2 is optimistic as it fits the sample better than it does the population. The adjusted R Square statistic, (R_a^2) is given by

$$R_a^2 = R^2 - \frac{p(1-R^2)}{N-p-1}$$

where p is the number of independent variables in the equation. R_a^2 more closely reflects the goodness of fit of the model in the population.

The hypothesis that in the population there is no linear association between the dependent and independent variables, *i.e.* $R^2=0$, can be tested by analysis of variance between the variability due to the regression *i.e.* the sum of squares of the difference between the mean of the dependent variable and the predicted values, and the variability due to the residuals. The ratio of the mean square due to regression to the mean square due to the residuals gives the F statistic and the corresponding significance level may be obtained.

The regression can be used to predict either a mean value or an individual value for the dependent variable given a certain value of the independent variable. The predicted mean value and the predicted individual value will be the same, however the standard error and thus the prediction interval in each case will be different. This is because not only might the estimate of the population mean differ from the population mean, but as an additional source of error a predicted individual value may differ from the population mean.

8.3.2 Multiple linear regression

Simple linear regression with a dependent variable and a single independent or predictor variable can be extended by adding further independent variables to the model *i.e.*

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_p x_p + e$$

With this model it is assumed that there is a normal distribution of the dependent variable for every combination of the values of the independent variables and again the error terms are normally distributed with constant variance. The goodness of fit of the model can be examined in similar fashion to that in the simple model. The presence of any linear association between the dependent variable and all the independent variables can also be tested by analysis of variance with the F statistic.

It is possible to include categorical predictor variables *e.g.* sex or whether or not headgear was worn, as well as continuous predictor variables in a regression equation by the use of "dummy" variables. For example, in the analysis used in this study the variable indicating the sex of the individual was the dummy variable "female" and for females was given the value 1. In a regression equation the β coefficient applied to the predictor variable "female" would measure the average difference in the dependent variable between males and females if all other predictor variables were equal.

The regression coefficients can not be directly compared to identify the relative importance of independent variables in the regression equation because the units of measurement of the variables may differ. However the sign of the coefficient may indicate whether the value of that variable increases or decreases the predicted value of the dependent variable. Assuming all other elements of a regression equation do not alter, for a unit change in the value of a predictor variable, the dependent variable will change by the amount of the partial regression coefficient (β) of the predictor variable which has changed. Regression coefficients may be made more comparable by their expression as beta coefficients when all variables are presented in standardised form.

Another method of comparing the relative importance of predictor variables is to determine what proportion of the unexplained variation the change in R^2 constitutes as a new variable is entered into the equation. If most of the variation had been explained by the other variables already in the equation then only a small change in R^2 is possible for the new variable.

Whether it is worth entering a particular independent variable into the model may be judged by the significance of the change in R^2 as each variable is entered. This may be decided using a partial F test which tests the null hypothesis that the true population change in the value of R^2 is 0.

Increasing the number of variables in the regression equation to include all the variables under investigation will increase the goodness of fit of the model to the sample, but not necessarily to the population so that the model's usefulness for prediction may not increase. On the other hand potentially relevant independent variables should not be excluded.

A number of regression models could be constructed for the same set of variables and a number of methods are available for the selection of independent variables to include in the model. Draper and Smith (1987) declared the "stepwise" method to be a suitable procedure for model construction as it avoids the use of too many predictor variables yet improves the regression equation at every stage.

In stepwise selection of variables the first predictor variable entered into the equation is that which has the highest correlation with the dependent variable. An F test is performed and the F value compared to an established criterion for entry into the equation. The criterion used in this study was that the probability associated with the F statistic should be less than or equal to 0.05.

At each step the next variable for entry into the equation is selected and tested against the entry criterion whilst the variables already in the equation are examined for that with the smallest partial correlation which is then F tested for removal from the equation. In this study the criterion for removal is that the probability associated with the F statistic should be greater than 0.1.

This entry and removal procedure is carried out step by step until none of the variables already in the equation or yet to be entered pass the removal or entry criteria.

This regression model comprising the persistent variables might be regarded as the "best" however it should be remembered that different methods of selection may yield different models even with identical criteria. The decision as to which is the best predictive model is open to judgement and may depend on practical issues such as usefulness, reliability of variable measurement and economy of effort in acquiring data. How closely the assumptions of regression are adhered to will be a factor in assessing the model's worth and for this reason examination of the residuals of regression should be undertaken.

Chapter 9

Error study

9.1 Method

A sub-sample of twenty of the cases examined in this study was selected by computer generated random number listing.

Pre-treatment and post-treatment study casts of the sub-sample were re-examined by the author at least one month after data collection to assess intra-examiner reliability and were also examined by a second examiner to assess inter-examiner reliability.

The second examiner had been similarly calibrated in the use of PAR (Intra-class Coefficient of Reliability, R=0.94, lower 95% confidence limit=0.89, RMS=3.6 no bias demonstrated) and IOTN (Aesthetic Component - weighted Kappa=0.71, lower 95% confidence limit 0.59, no bias demonstrated; Dental Health Component - weighted Kappa=0.82, lower 95% confidence limit=0.71, no bias demonstrated).

Data from the repeat examinations was collected in a manner similar to that described above for the main study. The data included IOTN grade and descriptive letter, PAR scores for the individual PAR elements, overjet measured to the nearest 0.5 mm, and the presence or otherwise of an incisal crossbite.

9.2 Analysis

9.2.1 IOTN

IOTN data is considered ordinal and thus necessitated non-parametric analysis. The Kappa statistic, (Cohen, 1960) is a chance corrected measure of agreement with general formula:

$$\frac{Po - Pe}{1 - Pe}$$

where *Po* is the proportion of observed agreement and *Pe* is the proportion of agreement that could be expected by chance.

If agreement was only by chance then Kappa=0. If agreement was perfect then Kappa=1. For good agreement to be demonstrated a Kappa value should be significantly greater than 0 and not significantly less than 0.8.

Landis and Koch (1977) have suggested the following guidelines for interpretation of the Kappa statistic:

Kappa	Agreement
<0.00	Poor
0.00 - 0.20	Slight
0.21 - 0.40	Fair
0.41 - 0.60	Moderate
0.61 - 0.80	Substantial
0.81 - 1.00	Almost perfect

Unweighted Kappa does not take into consideration the amount of disagreement *i.e.* all discrepancies between two sets of scores are regarded as equally serious no matter how many grades difference there are between them. Weighted Kappa on the other hand differentiates between large and small margins of error. The restricted linear weights used in this analysis were those suggested by Ciccheti (1976). The weights for DHC are presented in Table 1e and those for AC are given in Table 2e.

In this study both weighted and unweighted Kappa were calculated using the Microsoft Excel v3.0[3] computer spreadsheet program.

9.2.2 PAR

Total PAR scores may be considered to be normally distributed interval data and therefore parametric analysis was employed to assess intra- and inter-examiner agreement. The Intraclass Correlation Coefficient of reliability provides a single figure which summarises reliability based on comparison between examiners and comparison between subjects, (Fleiss, 1986).

The variation in a series of measurements on different study casts comprises the variation amongst the study casts and also the variation amongst the examiners. The Intraclass Correlation Coefficient expresses the relative magnitude of these two components of the total variation. A figure close to 1 indicates that most of the variation is due to that between study casts rather than between examiners.

9.2.3 Other variables

Agreement in the assessment of the presence of anterior crossbite, multiple rotations and incisor and canine angulation was investigated with the use of Kappa and Fisher's Exact test for 2 x 2 contingency tables.

---

[3]Microsoft Corporation
One Microsoft Way, Redmond, Washington

9.3 Results

9.3.1 IOTN

1 Intra-examiner agreement

Table 1a and Table 2a show the observed frequencies of agreement in DHC and AC for Examiner 1 between examinations carried out during the main period of data collection and the corresponding repeat examinations.

The observed proportions of the total for each cell are shown in Tables 1b and 2b whilst the expected values, calculated by multiplying the corresponding column total by the row total for each cell and then diving by the total number of comparisons, are given in Tables 1c and 2c.

The weighted observed and weighted expected proportions for DHC and AC are shown in Tables 1f and 1g and 2f and 2g.

For DHC unweighted Kappa was 0.72 and therefore demonstrated substantial agreement between DHC gradings for the main study and the repeated measurements, whilst the weighted Kappa of 0.82 confirmed a high level of reproducibility.

For AC only moderate agreement was demonstrated by unweighted Kappa (0.43) but a weighted Kappa of 0.81 showed almost perfect agreement for AC between main study and repeat assessments.

Error study: Intra-examiner agreement - DHC

Table 1a Observed values

Grade		Repeat 1	2	3	4	5	
Main study	1	0	0	0	0	0	0
	2	0	7	2	1	0	10
	3	0	0	9	2	0	11
	4	0	0	0	14	0	14
	5	0	0	0	2	3	5
		0	7	11	19	3	40

Table 1b Observed proportions

0	0	0	0	0
0	0.175	0.05	0.025	0
0	0	0.225	0.05	0
0	0	0	0.35	0
0	0	0	0.05	0.075

Po = 0.825

Table 1c Expected values

0	0	0	0	0	0
0	1.75	2.75	4.75	0.75	10
0	1.925	3.025	5.225	0.825	11
0	2.45	3.85	6.65	1.05	14
0	0.875	1.375	2.375	0.375	5
0	7	11	19	3	40

Table 1d Expected proportions

0	0	0	0	0
0	0.04375	0.06875	0.11875	0.01875
0	0.048125	0.075625	0.130625	0.020625
0	0.06125	0.09625	0.16625	0.02625
0	0.021875	0.034375	0.059375	0.009375

Pe = 0.295

k=(Po - Pe)/(1 - Pe)

Unweighted Kappa = **0.75**

Error study: Intra-examiner agreement - DHC

Table 1e Restricted linear weights (Ciccheti, 1976)

1	0.6	0.2	0	0
0.6	1	0.8	0.4	0
0.2	0.8	1	0.8	0.4
0	0.4	0.8	1	0.8
0	0	0.4	0.8	1

Table 1f Weighted observed proportions

0	0	0	0	0
0	0.175	0.04	0.01	0
0	0	0.225	0.04	0
0	0	0	0.35	0
0	0	0	0.04	0.075

Weighted *Po =* 0.955

Table 1g Weighted expected proportions

0	0	0	0	0
0	0.04375	0.055	0.0475	0
0	0.0385	0.075625	0.1045	0.00825
0	0.0245	0.077	0.16625	0.021
0	0	0.01375	0.0475	0.009375

Weighted *Pe =* 0.7325

Weighted Kappa = **0.83**

Error study: Intra-examiner agreement - AC

Table 2a Observed values

		Repeat										
Grade		1	2	3	4	5	6	7	8	9	10	
Main study	1	1	1	0	0	0	0	0	0	0	0	2
	2	1	3	0	0	0	0	0	0	0	0	4
	3	0	6	4	1	0	0	0	0	0	0	11
	4	0	0	0	1	3	0	0	0	0	0	4
	5	0	0	0	2	3	1	0	0	0	0	6
	6	0	0	0	0	1	2	0	1	0	0	4
	7	0	0	0	0	0	1	2	1	0	0	4
	8	0	0	0	0	0	0	1	0	0	0	1
	9	0	0	0	0	0	0	0	0	2	0	2
	10	0	0	0	0	0	0	0	0	0	2	2
		2	10	4	4	7	4	3	2	2	2	40

Table 2b Observed proportions

0.025	0.025	0	0	0	0	0	0	0	0
0.025	0.075	0	0	0	0	0	0	0	0
0	0.15	0.1	0.025	0	0	0	0	0	0
0	0	0	0.025	0.075	0	0	0	0	0
0	0	0	0.05	0.075	0.025	0	0	0	0
0	0	0	0	0.025	0.05	0	0.025	0	0
0	0	0	0	0	0.025	0.05	0.025	0	0
0	0	0	0	0	0	0.025	0	0	0
0	0	0	0	0	0	0	0	0.05	0
0	0	0	0	0	0	0	0	0	0.05

Po = 0.5

Error study: Intra-examiner agreement - AC

Table 2c Expected values

0.1	0.5	0.2	0.2	0.35	0.2	0.15	0.1	0.1	0.1	2
0.2	1	0.4	0.4	0.7	0.4	0.3	0.2	0.2	0.2	4
0.55	2.75	1.1	1.1	1.925	1.1	0.825	0.55	0.55	0.55	11
0.2	1	0.4	0.4	0.7	0.4	0.3	0.2	0.2	0.2	4
0.3	1.5	0.6	0.6	1.05	0.6	0.45	0.3	0.3	0.3	6
0.2	1	0.4	0.4	0.7	0.4	0.3	0.2	0.2	0.2	4
0.2	1	0.4	0.4	0.7	0.4	0.3	0.2	0.2	0.2	4
0.05	0.25	0.1	0.1	0.175	0.1	0.075	0.05	0.05	0.05	1
0.1	0.5	0.2	0.2	0.35	0.2	0.15	0.1	0.1	0.1	2
0.1	0.5	0.2	0.2	0.35	0.2	0.15	0.1	0.1	0.1	2
2	10	4	4	7	4	3	2	2	2	40

Table 2d Expected proportions

0.0025	0.0125	0.005	0.005	0.00875	0.005	0.00375	0.0025	0.0025	0.0025
0.005	0.025	0.01	0.01	0.0175	0.01	0.0075	0.005	0.005	0.005
0.01375	0.06875	0.0275	0.0275	0.048125	0.0275	0.020625	0.01375	0.01375	0.01375
0.005	0.025	0.01	0.01	0.0175	0.01	0.0075	0.005	0.005	0.005
0.0075	0.0375	0.015	0.015	0.02625	0.015	0.01125	0.0075	0.0075	0.0075
0.005	0.025	0.01	0.01	0.0175	0.01	0.0075	0.005	0.005	0.005
0.005	0.025	0.01	0.01	0.0175	0.01	0.0075	0.005	0.005	0.005
0.00125	0.00625	0.0025	0.0025	0.004375	0.0025	0.001875	0.00125	0.00125	0.00125
0.0025	0.0125	0.005	0.005	0.00875	0.005	0.00375	0.0025	0.0025	0.0025
0.0025	0.0125	0.005	0.005	0.00875	0.005	0.00375	0.0025	0.0025	0.0025

Pe = 0.115

$k=(Po - Pe)/(1 - Pe)$

Unweighted Kappa = **0.44**

Error study: Intra-examiner agreement - AC

Table 2e Restricted linear weights (Ciccheti, 1976)

1	0.6	0.2	0	0	0	0	0	0	0
0.6	1	0.8	0.4	0	0	0	0	0	0
0.2	0.8	1	0.8	0.4	0	0	0	0	0
0	0.4	0.8	1	0.8	0.4	0	0	0	0
0	0	0.4	0.8	1	0.8	0.4	0	0	0
0	0	0	0.4	0.8	1	0.8	0.4	0	0
0	0	0	0	0.4	0.8	1	0.8	0.4	0
0	0	0	0	0	0.4	0.8	1	0.8	0.2
0	0	0	0	0	0	0.4	0.8	1	0.6
0	0	0	0	0	0	0	0.2	0.6	1

Table 2f Weighted observed proportions

0.025	0.015	0	0	0	0	0	0	0	0
0.015	0.075	0	0	0	0	0	0	0	0
0	0.12	0.1	0.02	0	0	0	0	0	0
0	0	0	0.025	0.06	0	0	0	0	0
0	0	0	0.04	0.075	0.02	0	0	0	0
0	0	0	0	0.02	0.05	0	0.01	0	0
0	0	0	0	0	0.02	0.05	0.02	0	0
0	0	0	0	0	0	0.02	0	0	0
0	0	0	0	0	0	0	0	0.05	0
0	0	0	0	0	0	0	0	0	0.05

Weighted *Po* = 0.88

Error study: Intra-examiner agreement - AC

Table 2g Weighted expected proportions

0.0025	0.0075	0.001	0	0	0	0	0	0	0
0.003	0.025	0.008	0.004	0	0	0	0	0	0
0.00275	0.055	0.0275	0.022	0.01925	0	0	0	0	0
0	0.01	0.008	0.01	0.014	0.004	0	0	0	0
0	0	0.006	0.012	0.02625	0.012	0.0045	0	0	0
0	0	0	0.004	0.014	0.01	0.006	0.002	0	0
0	0	0	0	0.007	0.008	0.0075	0.004	0.002	0
0	0	0	0	0	0.001	0.0015	0.00125	0.001	0.00025
0	0	0	0	0	0	0.0015	0.002	0.0025	0.0015
0	0	0	0	0	0	0	0.0005	0.0015	0.0025

Weighted *Pe* = 0.36575

Weighted Kappa = **0.81**

2 Inter-examiner agreement

The observed and expected frequencies of agreement in DHC and AC between Examiner 1 and Examiner 2 and the corresponding proportions and weighted expected proportions are shown in Tables 3a to 4g.

Agreement between the examiners was almost perfect for DHC as indicated by both unweighted Kappa (0.84) and weighted Kappa (0.87).

There was greater discordance between examiners in AC. Agreement as tested by unweighted Kappa (0.37) rated only fair, however weighted Kappa (0.67) suggested there was substantial agreement.

A high degree of reproducibility was shown both between examiners and between two separate assessments of the error study sample by the same examiner. The reproducibility for DHC was higher than that for AC, but this is not surprising given the more subjective nature of the Aesthetic Component and the larger number of possible gradings compared to the Dental Health Component.

Error study: Inter-examiner agreement - DHC

Table 3a Observed values

	Grade	Examiner 2: 1	2	3	4	5	
Examiner 1	1	0	0	0	0	0	0
	2	0	6	0	1	0	7
	3	0	0	10	1	0	11
	4	0	0	1	18	0	19
	5	0	0	0	1	2	3
		0	6	11	21	2	40

Table 3b Observed proportions

0	0	0	0	0
0	0.15	0	0.025	0
0	0	0.25	0.025	0
0	0	0.025	0.45	0
0	0	0	0.025	0.05

Po = 0.9

Table 3c Expected values

0	0	0	0	0	0
0	1.05	1.925	3.675	0.35	7
0	1.65	3.025	5.775	0.55	11
0	2.85	5.225	9.975	0.95	19
0	0.45	0.825	1.575	0.15	3
0	6	11	21	2	40

Table 3d Expected proportions

0	0	0	0	0
0	0.02625	0.048125	0.091875	0.00875
0	0.04125	0.075625	0.144375	0.01375
0	0.07125	0.130625	0.249375	0.02375
0	0.01125	0.020625	0.039375	0.00375

Pe = 0.355

k=(Po - Pe)/(1 - Pe)

Unweighted Kappa = **0.84**

Error study: Inter-examiner agreement - DHC

Table 3e Restricted linear weights (Ciccheti, 1976)

1	0.6	0.2	0	0
0.6	1	0.8	0.4	0
0.2	0.8	1	0.8	0.4
0	0.4	0.8	1	0.8
0	0	0.4	0.8	1

Table 3f Weighted observed proportions

0	0	0	0	0
0	0.15	0	0.01	0
0	0	0.25	0.02	0
0	0	0.02	0.45	0
0	0	0	0.02	0.05

Weighted *Po* = 0.97

Table 3g Weighted expected proportions

0	0	0	0	0
0	0.02625	0.0385	0.03675	0
0	0.033	0.075625	0.1155	0.0055
0	0.0285	0.1045	0.249375	0.019
0	0	0.00825	0.0315	0.00375

Weighted *Pe* = 0.776

Weighted Kappa = **0.87**

Error study: Inter-examiner agreement - AC

Table 4a Observed values

		Examiner 2										
	Grade	1	2	3	4	5	6	7	8	9	10	
Examiner 1	1	1	1	0	0	0	0	0	0	0	0	2
	2	2	7	1	0	0	0	0	0	0	0	10
	3	0	0	3	0	1	0	0	0	0	0	4
	4	0	0	0	1	2	1	0	0	0	0	4
	5	0	0	0	1	1	0	4	1	0	0	7
	6	0	0	0	0	2	1	1	0	0	0	4
	7	0	0	0	0	0	1	0	1	1	0	3
	8	0	0	0	0	0	0	0	1	1	0	2
	9	0	0	0	0	0	0	0	0	2	0	2
	10	0	0	0	0	0	0	0	0	1	1	2
		3	8	4	2	6	3	5	3	5	1	40

Table 4b Observed proportions

0.025	0.025	0	0	0	0	0	0	0	0
0.05	0.175	0.025	0	0	0	0	0	0	0
0	0	0.075	0	0.025	0	0	0	0	0
0	0	0	0.025	0.05	0.025	0	0	0	0
0	0	0	0.025	0.025	0	0.1	0.025	0	0
0	0	0	0	0.05	0.025	0.025	0	0	0
0	0	0	0	0	0.025	0	0.025	0.025	0
0	0	0	0	0	0	0	0.025	0.025	0
0	0	0	0	0	0	0	0	0.05	0
0	0	0	0	0	0	0	0	0.025	0.025

Po = 0.45

Error study: Inter-examiner agreement - AC

Table 4c Expected values

0.15	0.4	0.2	0.1	0.3	0.15	0.25	0.15	0.25	0.05	2
0.75	2	1	0.5	1.5	0.75	1.25	0.75	1.25	0.25	10
0.3	0.8	0.4	0.2	0.6	0.3	0.5	0.3	0.5	0.1	4
0.3	0.8	0.4	0.2	0.6	0.3	0.5	0.3	0.5	0.1	4
0.525	1.4	0.7	0.35	1.05	0.525	0.875	0.525	0.875	0.175	7
0.3	0.8	0.4	0.2	0.6	0.3	0.5	0.3	0.5	0.1	4
0.225	0.6	0.3	0.15	0.45	0.225	0.375	0.225	0.375	0.075	3
0.15	0.4	0.2	0.1	0.3	0.15	0.25	0.15	0.25	0.05	2
0.15	0.4	0.2	0.1	0.3	0.15	0.25	0.15	0.25	0.05	2
0.15	0.4	0.2	0.1	0.3	0.15	0.25	0.15	0.25	0.05	2
3	8	4	2	6	3	5	3	5	1	40

Table 4d Expected proportions

0.00375	0.01	0.005	0.0025	0.0075	0.00375	0.00625	0.00375	0.00625	0.00125
0.01875	0.05	0.025	0.0125	0.0375	0.01875	0.03125	0.01875	0.03125	0.00625
0.0075	0.02	0.01	0.005	0.015	0.0075	0.0125	0.0075	0.0125	0.0025
0.0075	0.02	0.01	0.005	0.015	0.0075	0.0125	0.0075	0.0125	0.0025
0.013125	0.035	0.0175	0.00875	0.02625	0.013125	0.021875	0.013125	0.021875	0.004375
0.0075	0.02	0.01	0.005	0.015	0.0075	0.0125	0.0075	0.0125	0.0025
0.005625	0.015	0.0075	0.00375	0.01125	0.005625	0.009375	0.005625	0.009375	0.001875
0.00375	0.01	0.005	0.0025	0.0075	0.00375	0.00625	0.00375	0.00625	0.00125
0.00375	0.01	0.005	0.0025	0.0075	0.00375	0.00625	0.00375	0.00625	0.00125
0.00375	0.01	0.005	0.0025	0.0075	0.00375	0.00625	0.00375	0.00625	0.00125

Pe = 0.123125

k=(Po - Pe)/(1 - Pe)

Unweighted Kappa = **0.37**

Error study: Inter-examiner agreement - AC

Table 4e Restricted linear weights (Ciccheti, 1976)

1	0.6	0.2	0	0	0	0	0	0	0
0.6	1	0.8	0.4	0	0	0	0	0	0
0.2	0.8	1	0.8	0.4	0	0	0	0	0
0	0.4	0.8	1	0.8	0.4	0	0	0	0
0	0	0.4	0.8	1	0.8	0.4	0	0	0
0	0	0	0.4	0.8	1	0.8	0.4	0	0
0	0	0	0	0.4	0.8	1	0.8	0.4	0
0	0	0	0	0	0.4	0.8	1	0.8	0.2
0	0	0	0	0	0	0.4	0.8	1	0.6
0	0	0	0	0	0	0	0.2	0.6	1

Table 4f Weighted observed proportions

0.025	0.015	0	0	0	0	0	0	0	0
0.03	0.175	0.02	0	0	0	0	0	0	0
0	0	0.075	0	0.01	0	0	0	0	0
0	0	0	0.025	0.04	0.01	0	0	0	0
0	0	0	0.02	0.025	0	0.04	0	0	0
0	0	0	0	0.04	0.025	0.02	0	0	0
0	0	0	0	0	0.02	0	0.02	0.01	0
0	0	0	0	0	0	0	0.025	0.02	0
0	0	0	0	0	0	0	0	0.05	0
0	0	0	0	0	0	0	0	0.015	0.025

Weighted *Po* = 0.78

Error study: Inter-examiner agreement - AC

Table 4g Weighted expected proportions

0.00375	0.006	0.001	0	0	0	0	0	0	0
0.01125	0.05	0.02	0.005	0	0	0	0	0	0
0.0015	0.016	0.01	0.004	0.006	0	0	0	0	0
0	0.008	0.008	0.005	0.012	0.003	0	0	0	0
0	0	0.007	0.007	0.02625	0.0105	0.00875	0	0	0
0	0	0	0.002	0.012	0.0075	0.01	0.003	0	0
0	0	0	0	0.0045	0.0045	0.009375	0.0045	0.00375	0
0	0	0	0	0	0.0015	0.005	0.00375	0.005	0.00025
0	0	0	0	0	0	0.0025	0.003	0.00625	0.00075
0	0	0	0	0	0	0	0.00075	0.00375	0.00125

Weighted *Pe* = 0.324875

Weighted Kappa = **0.67**

9.3.2 PAR

1 Intra-examiner agreement

The scatterplots (Figures 3-4) indicate a high level of correlation between main study and repeat measurements with a maximum difference between measurements of 8 PAR points and a mean difference of 0.125. The most frequently occurring difference was 1 PAR point but with no apparent bias (Figure 5). The lack of bias was supported by a paired *t*-test, (t=0.25, DF=39, not significant).

The Intraclass Correlation Coefficient for intra-examiner agreement was 0.95 which indicated a high degree of reliability.

It could be concluded that there had been no change in the examiner's interpretation of the PAR criteria over the time between the collection of the main study data and the repeat measurements.

Error study: Intra-examiner error in total PAR score

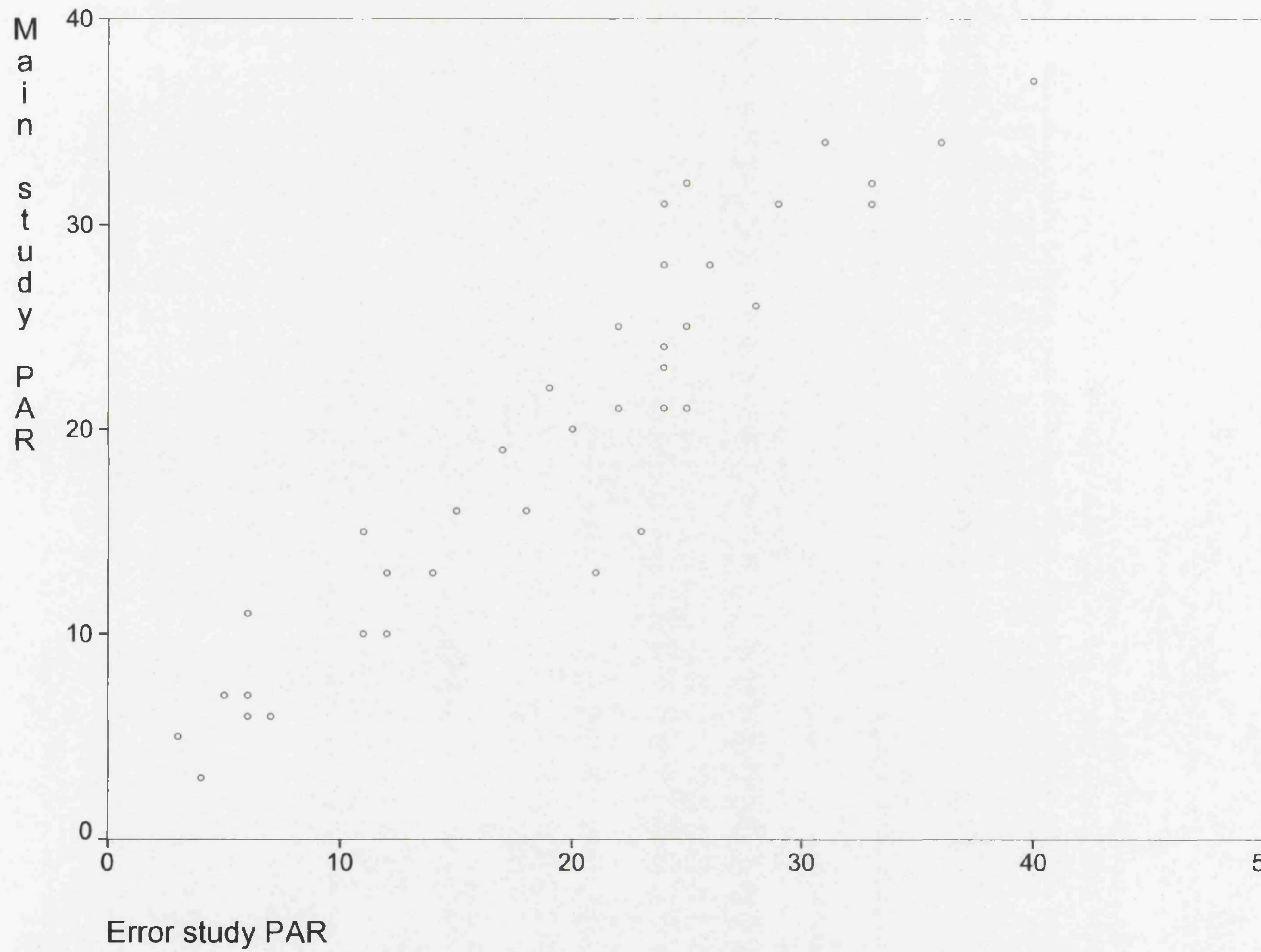


Figure 3

Error study: Intra-examiner differences in total PAR score

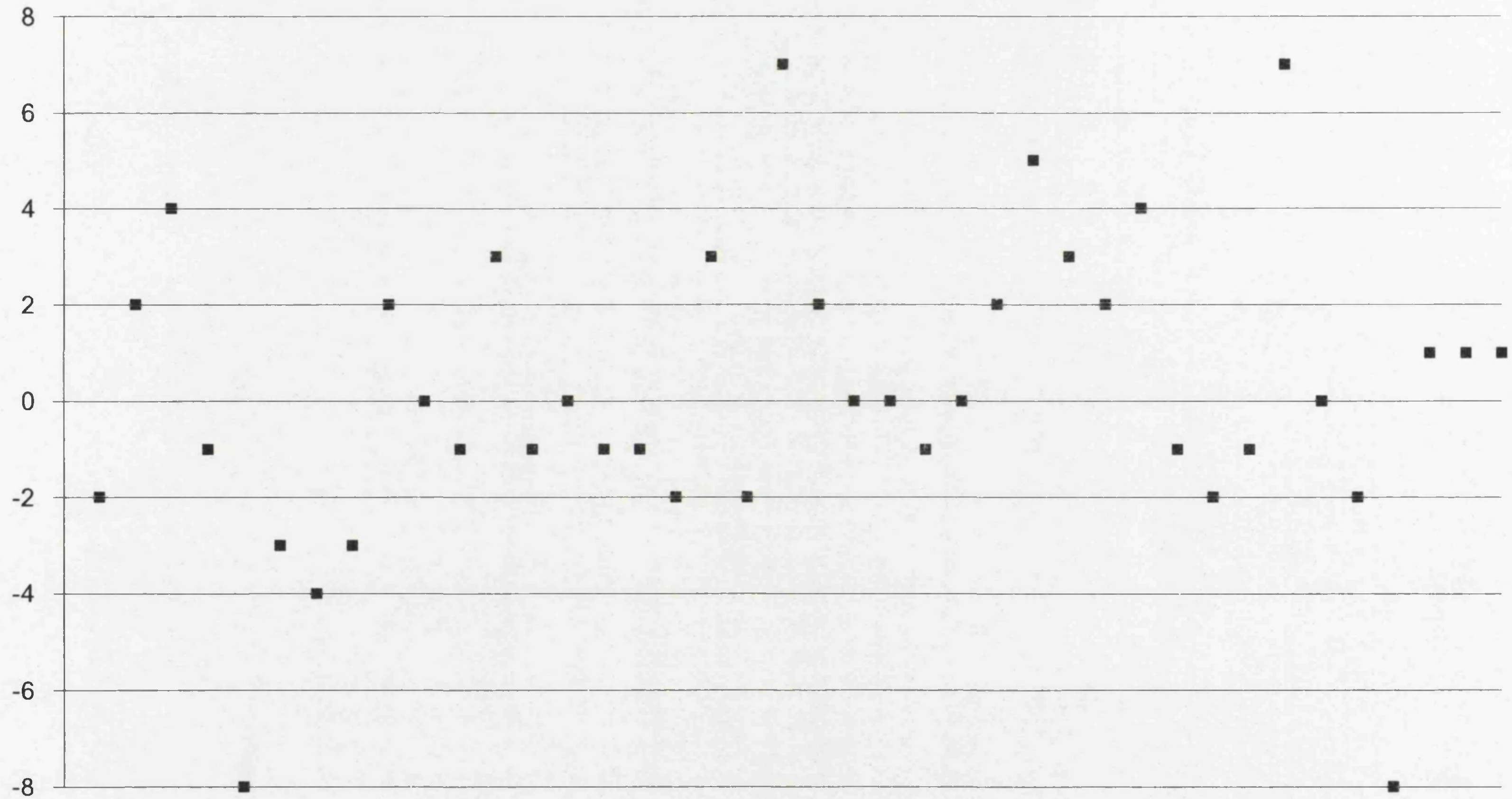


Figure 4

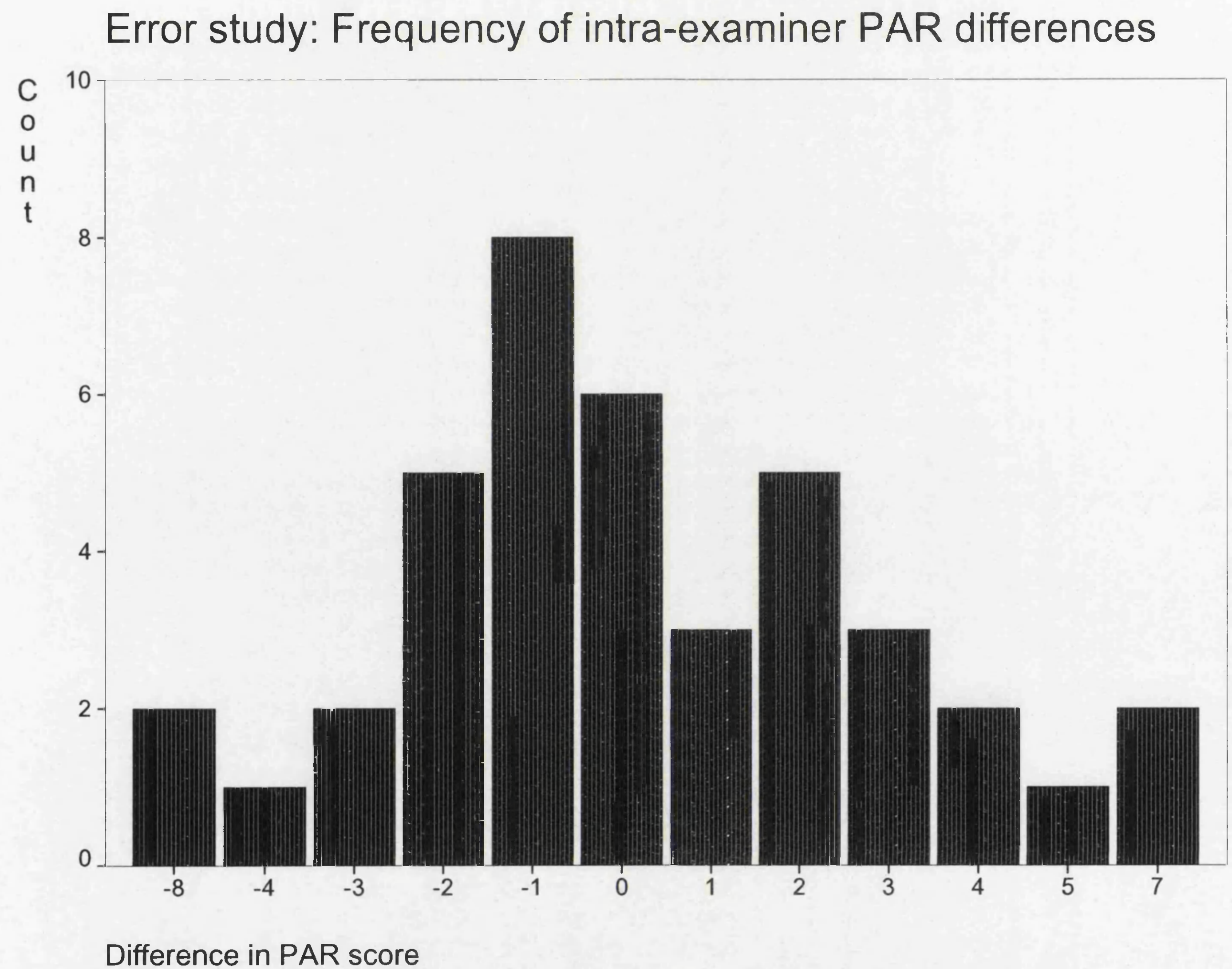

Error study: Frequency of intra-examiner PAR differences
Count
0
2
4
6
8
10
-8
-4
-3
-2
-1
0
1
2
3
4
5
7
Difference in PAR score

Figure 5

2 Inter-examiner agreement

The scatterplots for total PAR score differences between examiners 1 and 2 show a similar spread to that for intra-examiner scores (Figures 6-7). The maximum difference was 8 PAR points. However on examining the bar chart of difference frequency there appeared to be a tendency for Examiner 1 to score higher (Figure 8). The mean PAR score for Examiner 1 was 18.6 whilst that for Examiner 2 was 17.7. This difference in means proved just significant at the 95% level by paired t-testing, (t=2.04, DF=39, P=0.048) indicating the presence of bias.

The Intraclass Correlation Coefficient for inter-examiner agreement was 0.96 which indicated a high degree of reliability.

It could be concluded that although slight bias was demonstrated when comparing the two examiners' total PAR scores, PAR scoring in this study was consistent.

Error study: Total PAR score by examiner

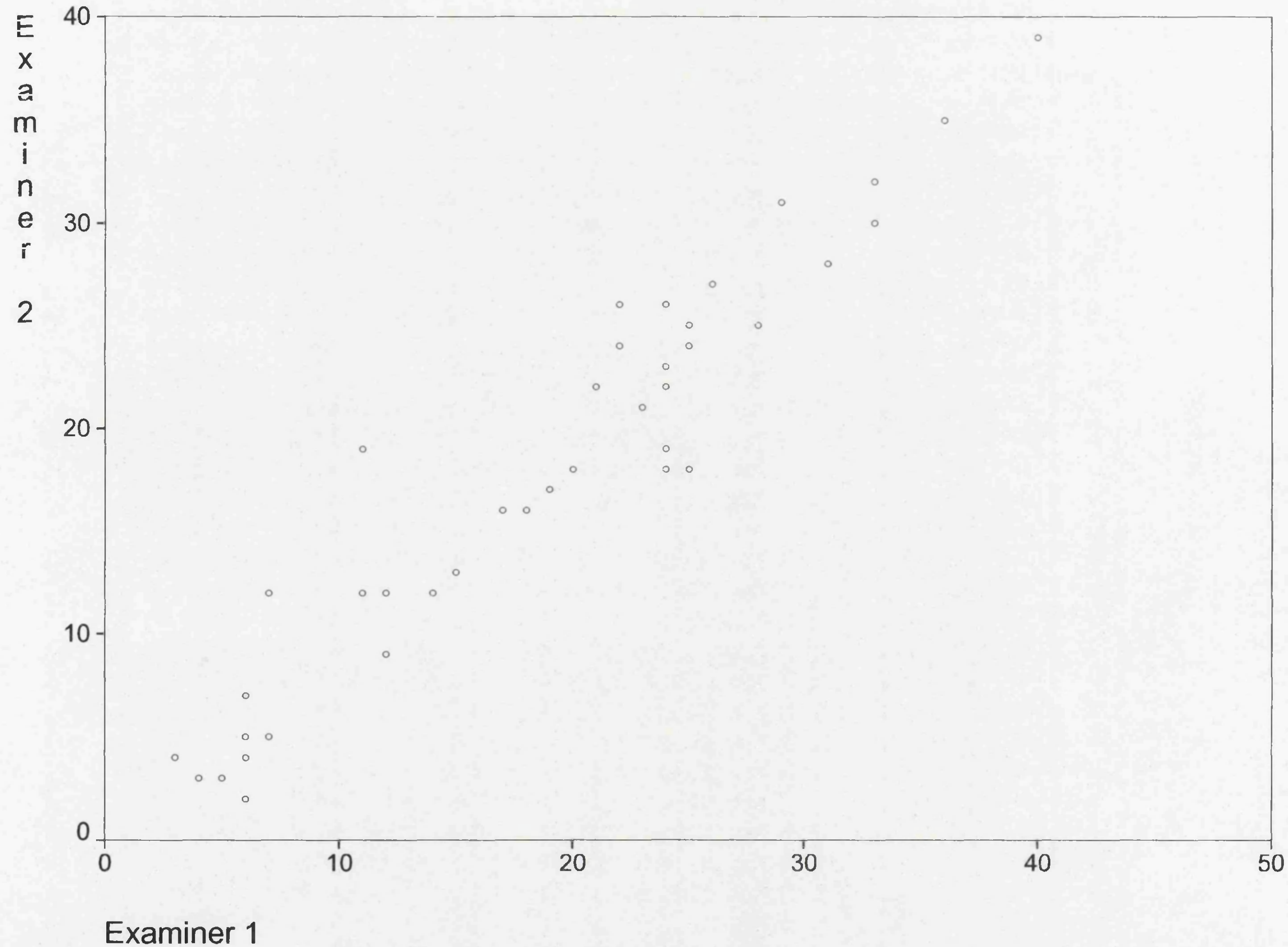


Figure 6

Error study: Inter-examiner differences in total PAR score

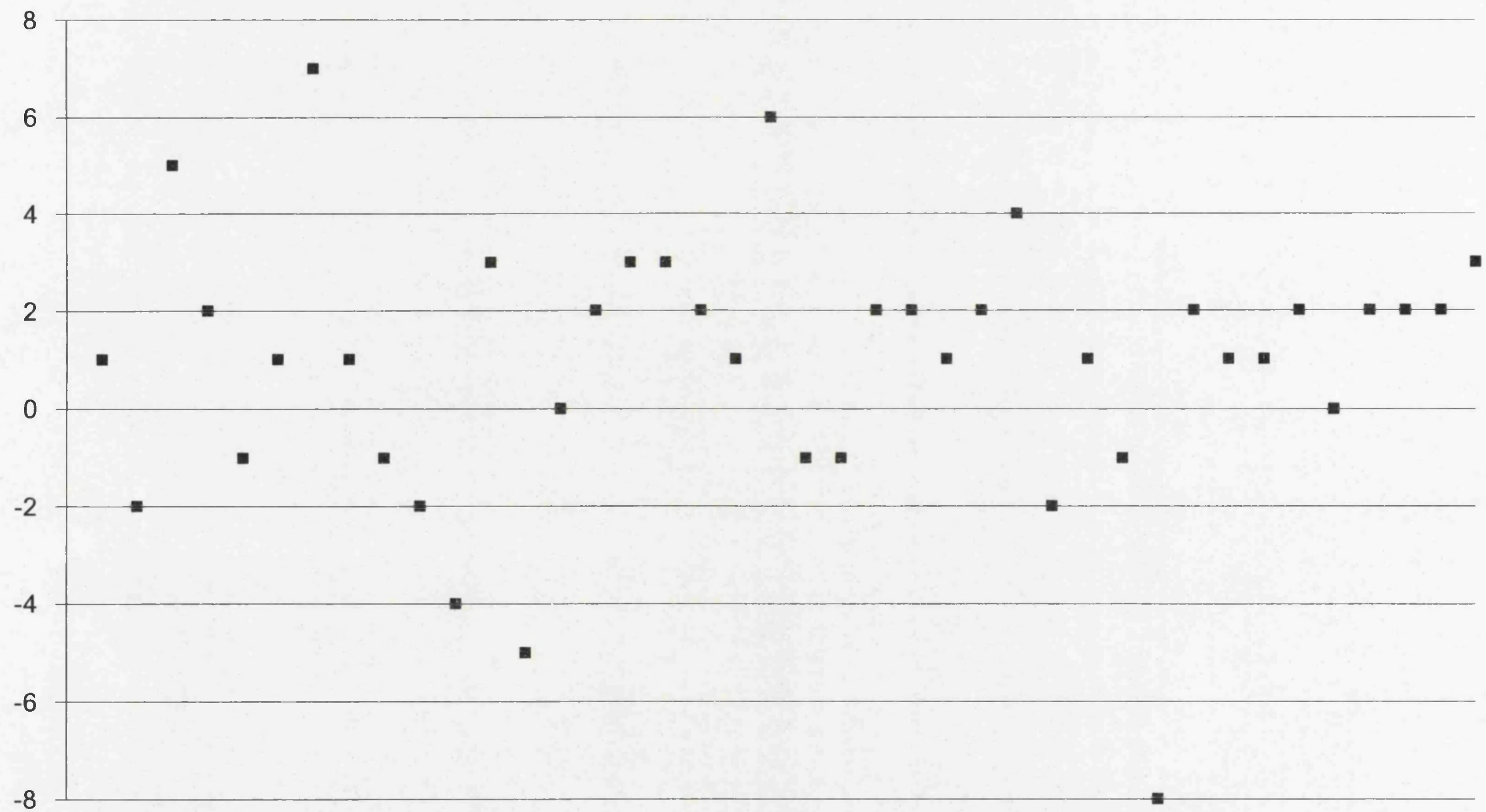


Figure 7

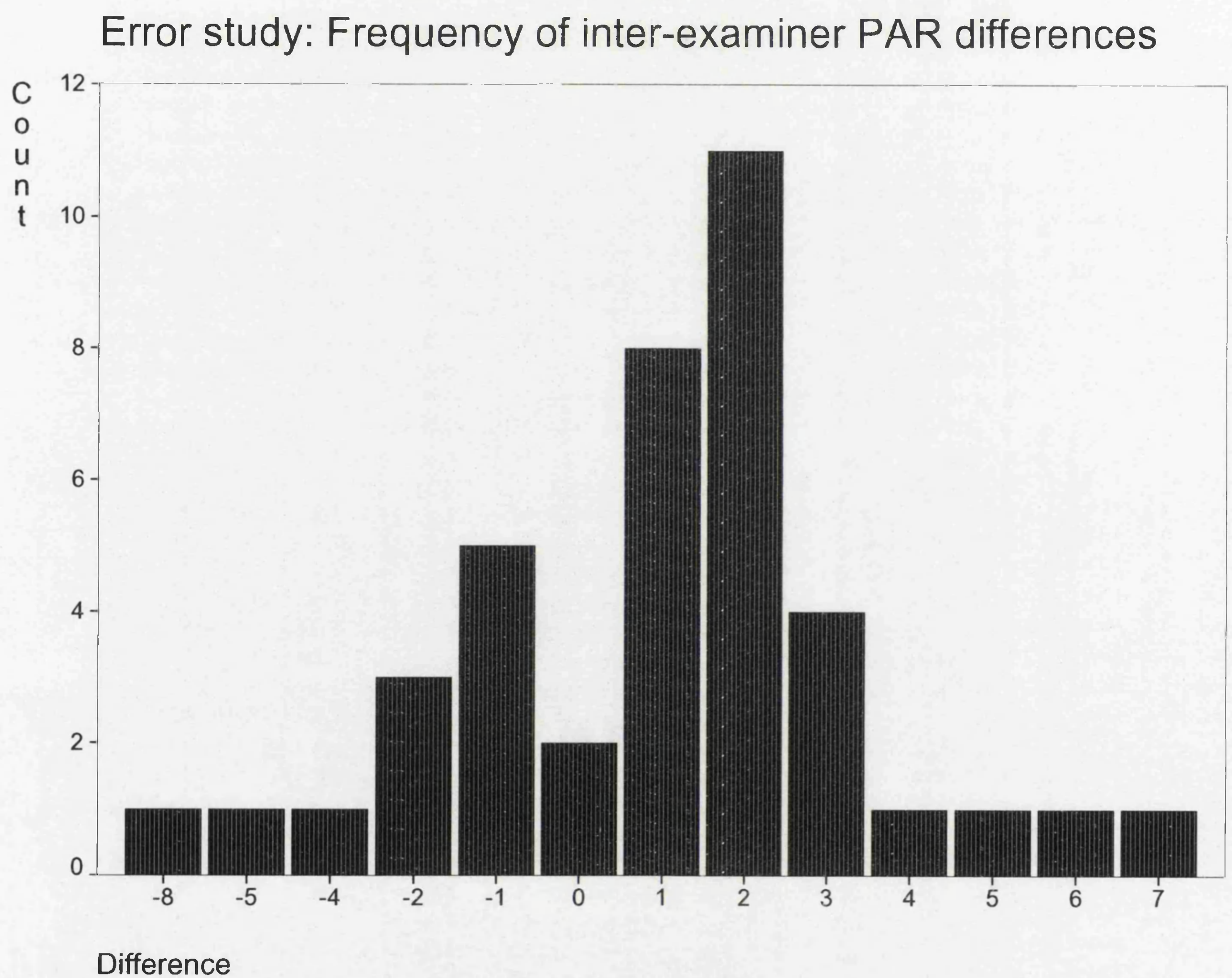

Error study: Frequency of inter-examiner PAR differences
Count
12
10
8
6
4
2
0
-8
-5
-4
-2
-1
0
1
2
3
4
5
6
7
Difference

Figure 8

9.3.3 Other variables

Anterior crossbite: Intra-examiner agreement for the presence of anterior crossbite was substantial (Kappa=0.74) whilst Fisher's Exact test was highly significant (P=0.003). Between examiners agreement was also very high (Kappa=0.87) and the association found between the two examiners assessments was unlikely to have been by chance (P<0.001).

Canine angulation: Kappa for intra-examiner agreement for canine angulation was very low at -0.07 indicating poor agreement. Fisher's Exact test demonstrated no association between the scores from the main study and the repeat assessments (P=1.0).

Incisor angulation: Intra-examiner agreement for incisor angulation as indicated by Kappa was substantial, however the association observed between initial and repeat assessments may have been due to chance (Fisher's Exact test, P=0.1).

Rotations: Similarly agreement for the presence of multiple rotations *i.e.* two or more, was only fair (Kappa=0.32). For Fisher's Exact test P=0.28, again indicating that there was no significant association between initial and repeat assessments.

The lack of significance demonstrated for multiple rotations, canine and incisor angulation as defined may have been attributed to the fact that only a very small proportion of the error study sample exhibited these features. As the reproducibility of the method for identifying these traits appeared in doubt it was decided to eliminate these factors from any further analysis.

A study of factors associated with orthodontic treatment outcome.

Section D

RESULTS AND DISCUSSION

In Chapter 10 the data for each variable is described with comparisons made between appliance groups. Chapter 11 sets out and discusses the results of regression procedures, firstly with all relevant variables and secondly with only information known prior to commencement of treatment, for each of the outcome variables.

Section contents

Chapter 10

Sample profile

10.1 Elimination of functional cases

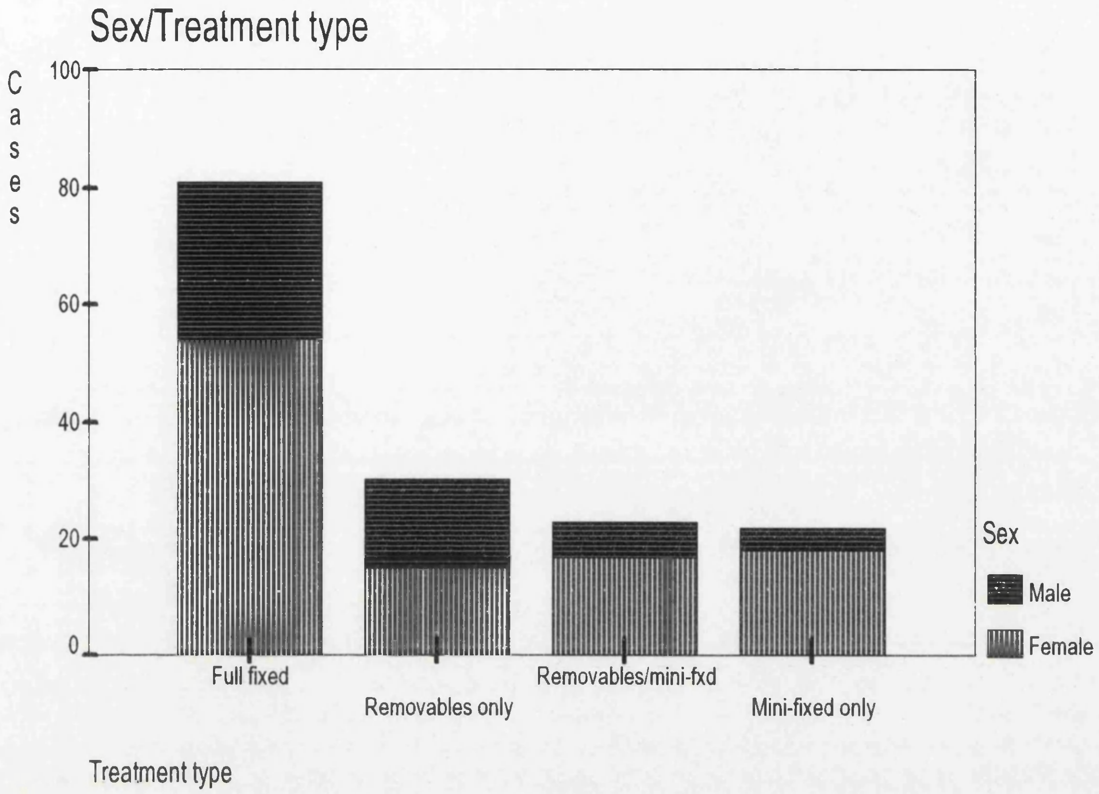
Of the original sample of 161 cases, sorting under appliance type revealed only five cases which had been almost entirely treated with functional appliance therapy. In view of the small number in this appliance type group and as this form of treatment tends to be restricted to use in children in the late mixed and early permanent dentition between the ages 10-15 it was decided to eliminate these cases from the study. The following description and subsequent analysis pertains only to the remaining 156 cases.

10.2 Descriptive analysis: Patient factors

10.2.1 Sex

The total working sample of 156 comprised two thirds females (104) and one third males (52).

Examining the sex distribution by treatment type (Figure 9) showed there to be a similar picture in the Full fixed appliance group to that in the overall sample, but in the Removables group there was an exactly even number of males and females. The Mini-fixed group contained a preponderance of females. Combining the non-full fixed appliance groups resulted in a sex distribution of 66.7% female to 33.3% male which is identical to that observed in the Full fixed appliances.


Sex/Treatment type
Cases
100
80
60
40
20
0
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type
Sex
Male
Female

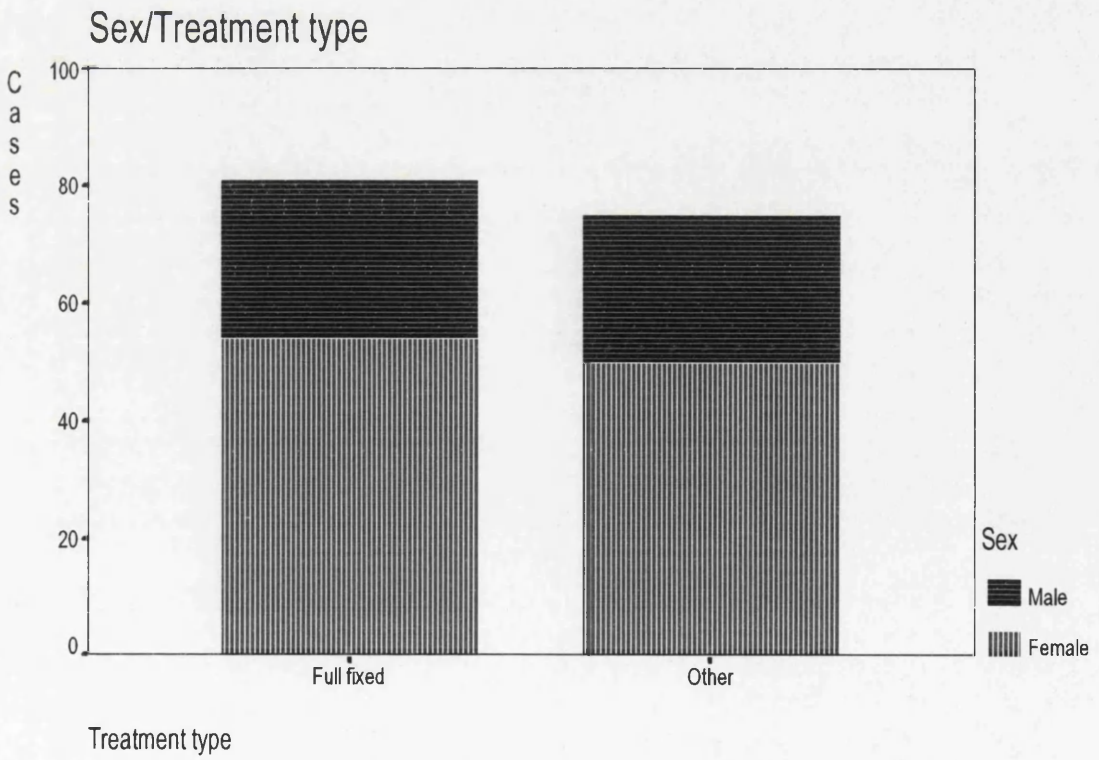

Sex/Treatment type
Cases
100
80
60
40
20
0
Full fixed
Other
Treatment type
Sex
Male
Female

Figure 9

Chi-square testing failed to show any significant association between the sex of the patient and the mode of treatment (Table 5).

Count Col%	**Full fixed**	**Removable only**	**Removable /mini-fixed**	**Mini-fixed only**	Row total
Female	54 66.7%	15 50.0%	17 73.9%	18 81.8%	104 66.7%
Male	27 33.3%	15 50.0%	6 26.1%	4 18.2%	52 33.3%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=6.57 DF=3 *P*=0.087

Count Col%	**Full fixed**	**Other**	Row total
Female	54 66.7%	50 66.7%	104 66.7%
Male	27 66.7%	25 33.3%	52 33.3%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=0 00 DF=1 *P*=1.0

Table 5

10.2.2 Age at start of treatment

Overall, the age at start of treatment ranged from 8.0 to 29.4 years with a mean of 13.8 years (S.D. 3.1).

Mean starting age in years for each appliance group was:

Overall	13.8	S.D. 3.1
Full fixed	14.0	S.D. 2.9
Other	13.6	S.D. 3.3
Removable only	12.4	S.D. 2.8
Removable/Mini-fixed	14.0	S.D. 3.6
Mini-fixed only	14.9	S.D. 3.0

One-way analysis of variance and a Scheffé multiple comparisons procedure indicated a significant difference in age at start of treatment between the Removable appliance and Mini-fixed appliance groups (P=0.02), (Table 6).

Source	DF	Sum of Squares	Mean Squares	F Ratio	F Prob
Between Groups	3	88.04	29.34	3.25	.02
Within Groups	152	1373.68	9.04		
Total	155	1461.72			

Table 6

The mean ages at start of treatment for males and females were practically identical at 13.8 (S.D. 3.1) and 13.9 years (S.D. 3.0) respectively.

10.2.3 Developmental stage

Of all the cases in the study only one did not begin treatment in either the late mixed or early permanent dentition therefore the two mixed dentition groups were combined.

Differences between treatment groups in respect of developmental stage at start of treatment were found to be significant when appliance types were considered both separately and combined (Table 7).

Count Col%	**Full fixed**	**Removable only**	**Removable/ mini-fixed**	**Mini-fixed only**	Row total
Mixed	16 19.8%	12 40.0%	11 47.8%	3 13.6%	42 26.9%
Perm.	65 80.2%	18 60.0%	12 52.2%	19 86.4%	114 73.1%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=11.81 DF=3 *P*=0.008

Count Col%	**Full fixed**	**Other**	Row total
Mixed	16 19.8%	26 34.7%	42 26.9%
Permanent	65 80.2%	49 65.3%	114 73.1%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=4.40 DF=1 *P*=0.04

Table 7

10.2.4 Incisor classification

The distribution of malocclusion as described by incisor classification among the various treatment groups is shown in Table 8.

No significant difference in terms of incisor classification was found between groups when the Full fixed appliance cases were compared to a combined group of all the other appliance types. A chi-square procedure was not performed for the contingency table showing all treatment groups as over one third of the expected frequencies were less than 5 and two were less than 2.

Count Col%	Full fixed	Removable only	Removable/mini-fixed	Mini-fixed only	Row total
Class I	29 35.8%	11 36.7%	12 52.2%	15 68.2%	67 42.9%
Class II/1	37 45.7%	12 40.0%	6 26.1%	4 18.2%	59 37.8%
Class II/2	10 12.3%	4 13.3%	3 13.0%	1 4.5%	18 11.5%
Class III	5 6.2%	3 10.0%	2 8.7%	2 9.1%	12 7.7%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Count Col%	Full fixed	Other	Row total
Class I	29 35.8%	38 50.7%	67 42.9%
Class II/1	37 45.7%	22 29.3%	59 37.8%
Class II/2	10 12.3%	8 10.7%	18 11.5%
Class III	5 6.2%	7 9.3%	12 7.7%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=5.35 DF=3 P=0.147

Table 8

10.2.5 Pre-treatment overjet

Pre-treatment overjet ranged from 14.5 mm to -3.0 mm. Both extreme figures were found in the Full fixed appliance group.

Group means in millimetres for pre-treatment overjet are given below (Table 9).

Group	Count	Mean	S.D.	Minimum	Maximum
Full fixed	81	4.9	3.7	-3.0	14.5
Removables	30	4.1	3.4	-2.0	12.0
Removables/mini-fixed	23	3.6	2.3	1.0	11.5
Mini-fixed	22	3.6	2.7	1.0	14.0
Other	*75*	*3.8*	*2.8*	*-2.0*	*14.0*
Total	156	4.6	3.3	-3.0	14.5

Table 9

An *F* test (Table 10) failed to show any significant difference between groups although the difference between Full fixed and combined non-full fixed types was significant.

ALL APPLIANCE GROUPS

Source	DF	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	54.34	18.11	1.63	.18
Within Groups	152	1684.16	11.08		
Total	155	1738.50			

FULL FIXED/OTHER

Source	DF	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	49.00	49.00	4.4	.04
Within Groups	154	1689.50	10.97		
Total	155	1738.50			

Table 10

10.2.6 Presence of anterior crossbite

34 cases out of the total sample exhibited the presence of an anterior crossbite. The proportion of cases with anterior crossbite in each group appeared to be similar (Table 11) and no significant difference was demonstrated between the separate treatment type groups or between Full fixed and Other groups in this respect.

Count Col%	Full fixed	Removable only	Removable/ mini-fixed	Mini-fixed only	Row total
No crossbite	65 80.2%	22 73.3%	19 82.6%	16 72.7%	122 78.2%
Crossbite	16 19.8%	8 26.7%	4 17.4	6 27.3%	34 21.8%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=1.26 DF=3 *P*=0.74

Count Col%	Full fixed	Other	Row total
No crossbite	65 80.2%	57 76.0%	122 78.2%
Crossbite	16 19.8%	18 24.0%	34 21.8%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=0.41 DF=1 *P*=0.52

Table 11

10.2.7 Presence of ectopic teeth

19 cases out of the total sample exhibited the presence of one or more ectopic teeth with the highest proportion being found in the Removables/mini-fixed group, where an ectopic tooth was present in 30.4% of cases. This difference appeared to be significant. There was no significant difference found when the combined Other group was compared to the Full fixed group (Table 12).

Count Col%	Full fixed	Removable only	Removable/ mini-fixed	Mini-fixed only	Row total
Not present	74 91.4%	28 93.3%	16 69.6%	19 86.4%	137 87.8%
Present	7 8.6%	2 6.7%	7 30.4%	3 13.6%	19 12.2%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=9.00 DF=3 *P*=0.03

Table 12a

Count Col%	**Full fixed**	**Other**	Row total
Not present	74 91.4%	63 84.0%	137 87.8%
Present	7 8.6%	12 16.0%	19 12.2%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=1.97 DF=1 *P*=0.16

Table 12b

10.2.8 Pre-treatment IOTN

The frequency of occurrence of DHC grades in each of the treatment type groups is given below (Table 13).

Count Col%	**Full fixed**	**Removable only**	**Removable/ mini-fixed**	**Mini-fixed only**	Row total
Grade 2				1 4.5%	1 0.7%
Grade 3	18 22.2%	2 6.7%	7 30.4%	5 22.7%	32 20.5%
Grade 4	40 49.4%	23 76.7%	10 43.5%	13 59.1%	86 55.1%
Grade 5	23 28.4%	5 16.7%	6 26.1%	2 13.6%	37 23.7%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Table 13

The number of cases falling into DHC grades 4 and 5 was approximately the same in the Full fixed appliance group and the other groups combined (Figure 10).

The distribution of DHC grades with respect to sex is shown in Table 14

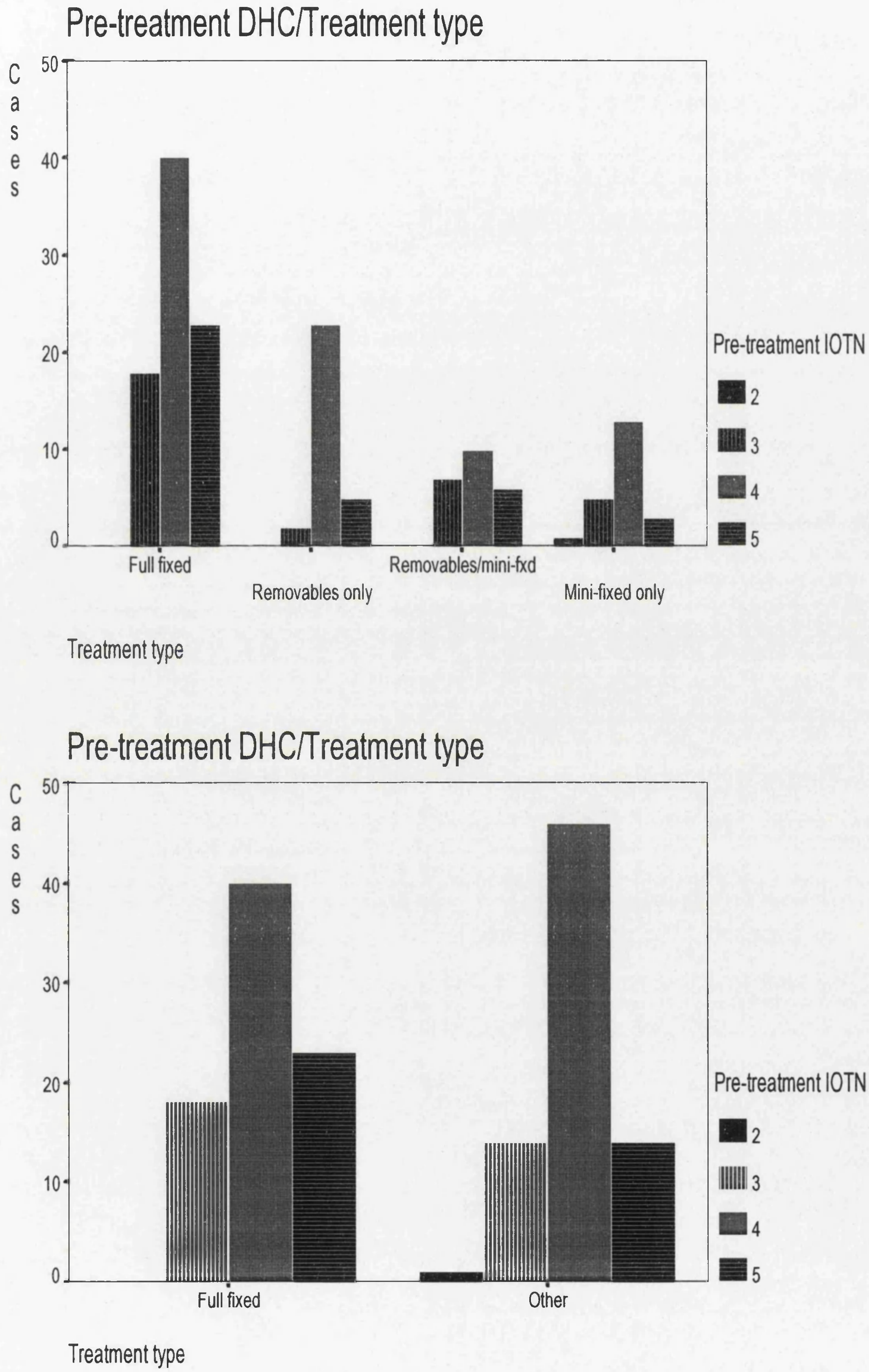

Pre-treatment DHC/Treatment type
Cases
50
40
30
20
10
0
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type
Pre-treatment IOTN
2
3
4
5
Pre-treatment DHC/Treatment type
Cases
50
40
30
20
10
0
Full fixed
Other
Treatment type
Pre-treatment IOTN
2
3
4
5

Figure 10

Count Col%	Grade 2	Grade 3	Grade 4	Grade 5	Row total
Female	1 100.0%	25 78.1%	57 66.3%	21 56.8%	104 66.7%
Male	0 0%	7 21.9%	29 33.7%	16 43.2%	52 33.3%
Column total	1 0.6%	32 20.5%	86 55.1%	37 23.7%	156 100.0%

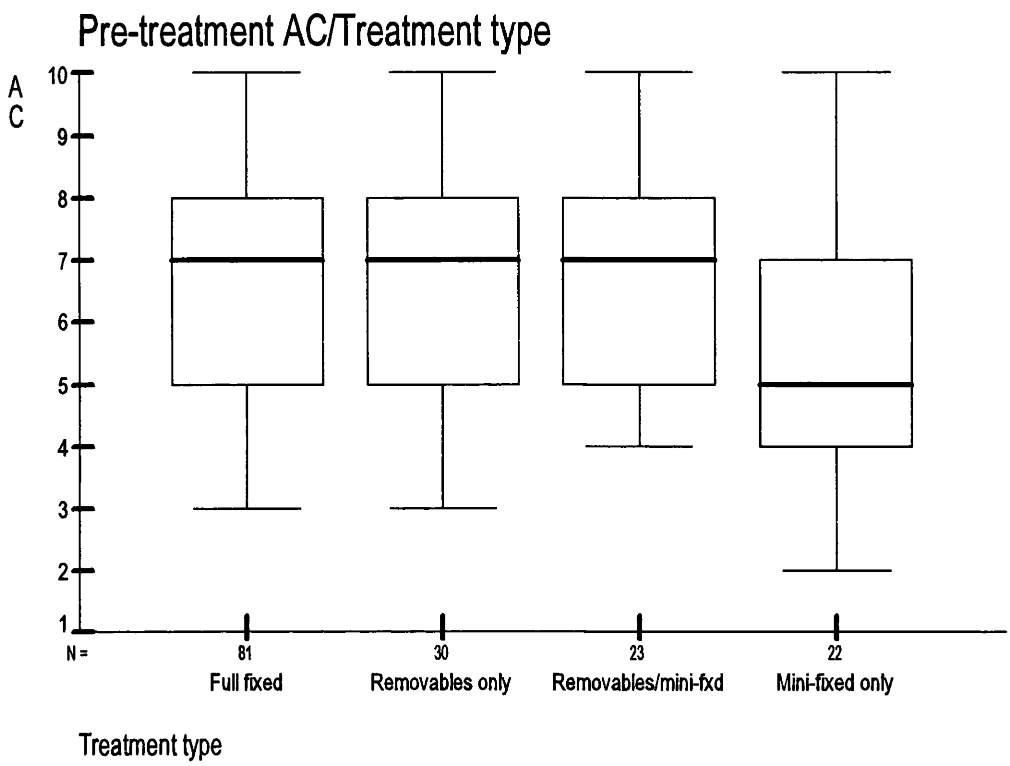
Table 14

Regarding the DHC letter categorisation for description of main feature in need of treatment, only six cases did not fall into categories A(overjet), C(crossbite), D(contact point displacement), or I(impaction). Category A was most common with 36.5% of cases, followed by C (31.4%), D (15.4%) and I (12.8%). The proportions in each treatment type group are shown in Table 14

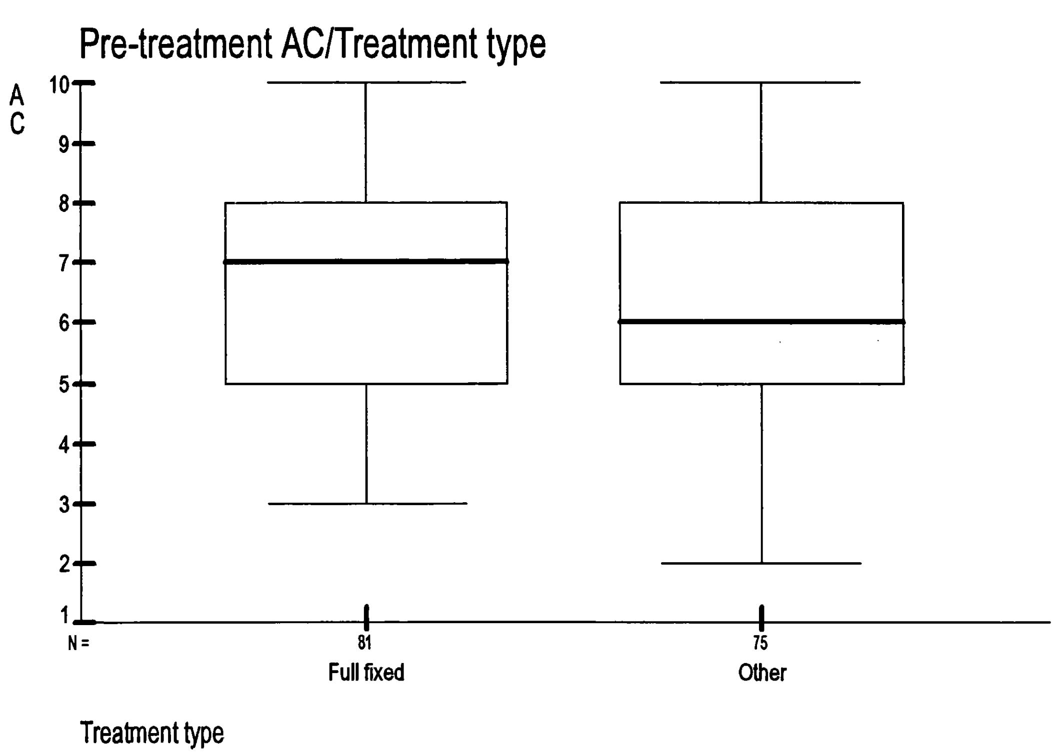
Count Col%	Full fixed	Removable only	Removable/ mini-fixed	Mini-fixed only
A overjet	43.2%	30.0%	30.4%	27.3%
C crossbite	25.9%	46.7%	34.8%	27.3%
D displacement	14.8%	13.3%	8.7%	27.3%
I impaction	13.6%	10.0%	17.4%	9.1%

Table 14

There was little difference between groups regarding the distribution of pre-treatment AC grades, although that for the Mini-fixed cases, 60% of which were grade 5 or less, may have been generally slightly lower. The median AC grade overall was 7, which was the median grade for all the treatment sub-groups except the Mini-fixed cases, the median of which was 5 (Figure 11).



Pre-treatment AC/Treatment type
A
C
10
9
8
7
6
5
4
3
2
1
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type



Pre-treatment AC/Treatment type
A
C
10
9
8
7
6
5
4
3
2
1
N =
81
75
Full fixed
Other
Treatment type

Figure 11

10.2.9 Pre-treatment PAR

Mean pre-treatment PAR scores for each appliance group are shown below (Table 16).

Group	Count	Mean	S.D.	Minimum	Maximum
Full fixed	81	30.74	9.07	12.00	53.00
Removables	30	28.73	9.39	13.00	51.00
Removables/mini-fixed	23	26.43	9.10	13.00	42.00
Mini-fixed	22	21.09	9.77	6.00	43.00
Other	*75*	*25.79*	*9.82*	*6.00*	*51.00*
Total	156	28.36	9.73		

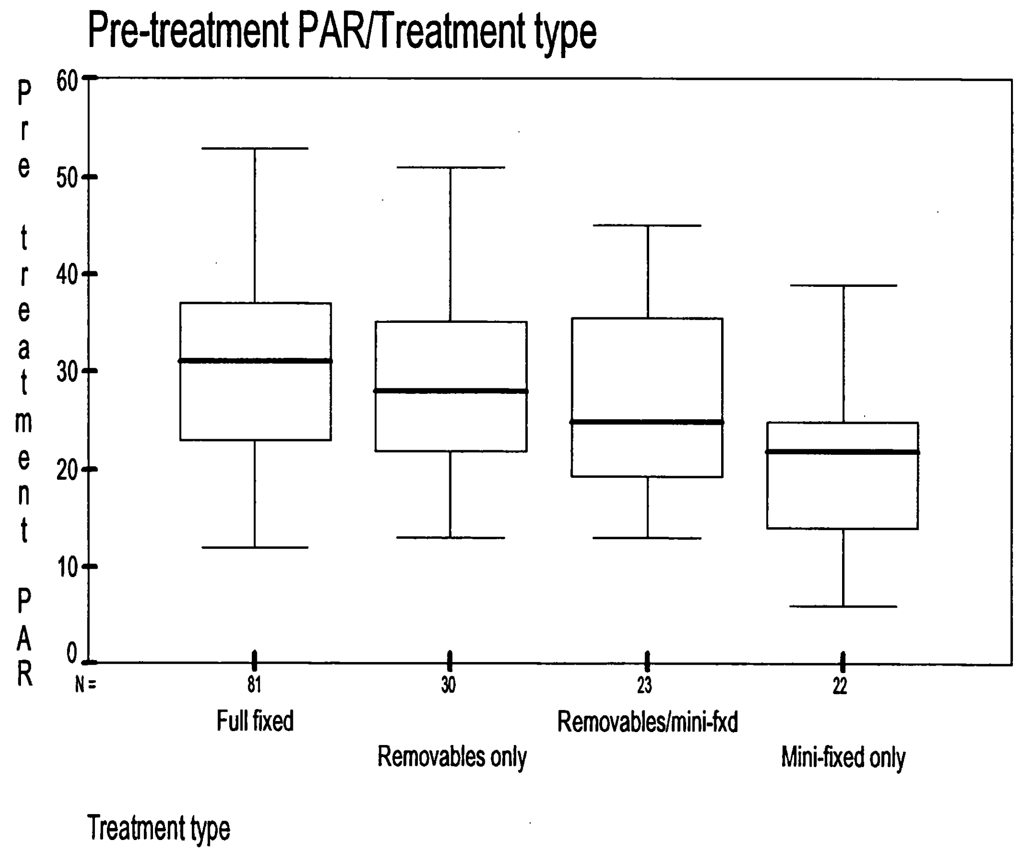
Table 16

Examination of box-and-whisker plots suggested that the Removables/mini-fixed group contained the smallest spread of scores and also that the scores for the Mini-fixed group were generally lower than for the other groups (Figure 12). The combined Other group appeared to have a similar distribution to the Full fixed group but again the scores for this group were generally lower in comparison.

Analysis of variance demonstrated significant differences in mean initial PAR score (P<0.005) between Mini-fixed and Full fixed, and also between Mini-fixed and Removable appliance groups (Table 17a). A significant difference in respect of pre-treatment PAR was also found between the Full fixed and Other groups (P=0.001, Table 17b).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	1711.00	3	570.33	6.69	.000
ALL TREATMENT TYPES	1711.00	3	570.33	6.69	.000
Explained	1711.00	3	570.33	6.69	.000
Residual	12964.89	152	85.29		
Total	14675.90	155	94.68		

Table 17a


Pre-treatment PAR/Treatment type
Pretreatment PAR
60
50
40
30
20
10
0
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type

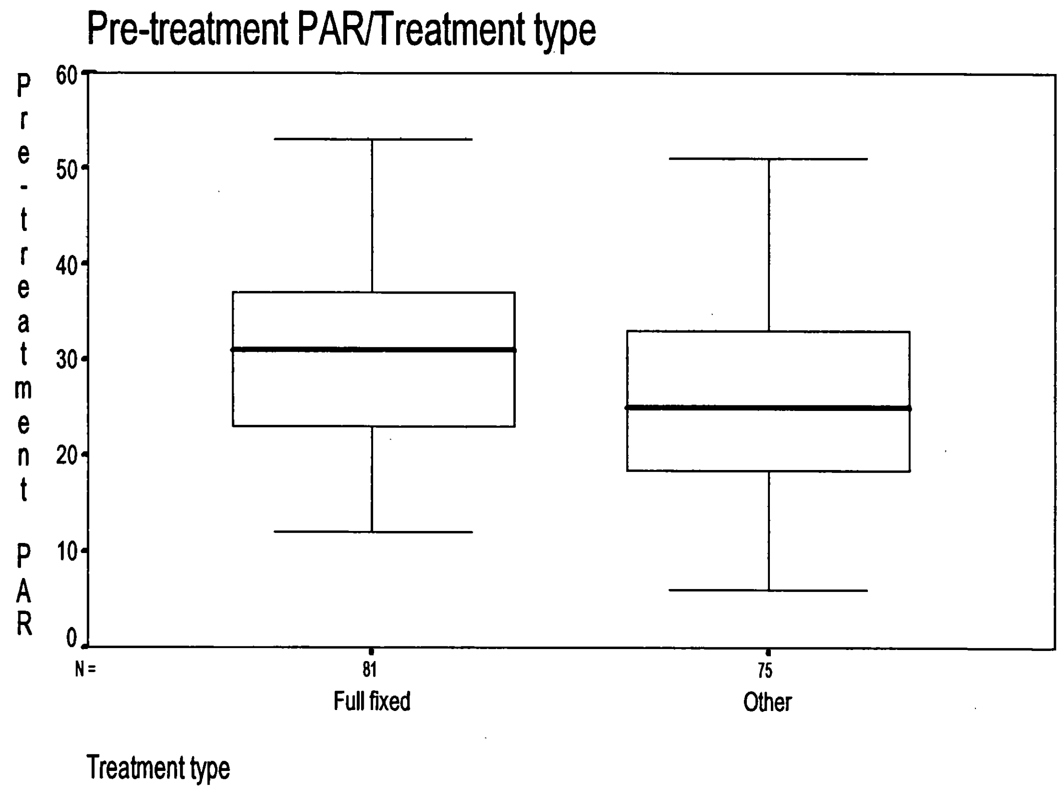

Pre-treatment PAR/Treatment type
Pre-treatment PAR
60
50
40
30
20
10
0
N =
81
75
Full fixed
Other
Treatment type

Figure 12

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	955.75	1	955.75	10.73	.001
FULL FIXED/OTHER	955.75	1	955.75	10.73	.001
Explained	955.75	1	955.75	10.73	.001
Residual	13720.14	154	89.09		
Total	14675.90	155	94.68		

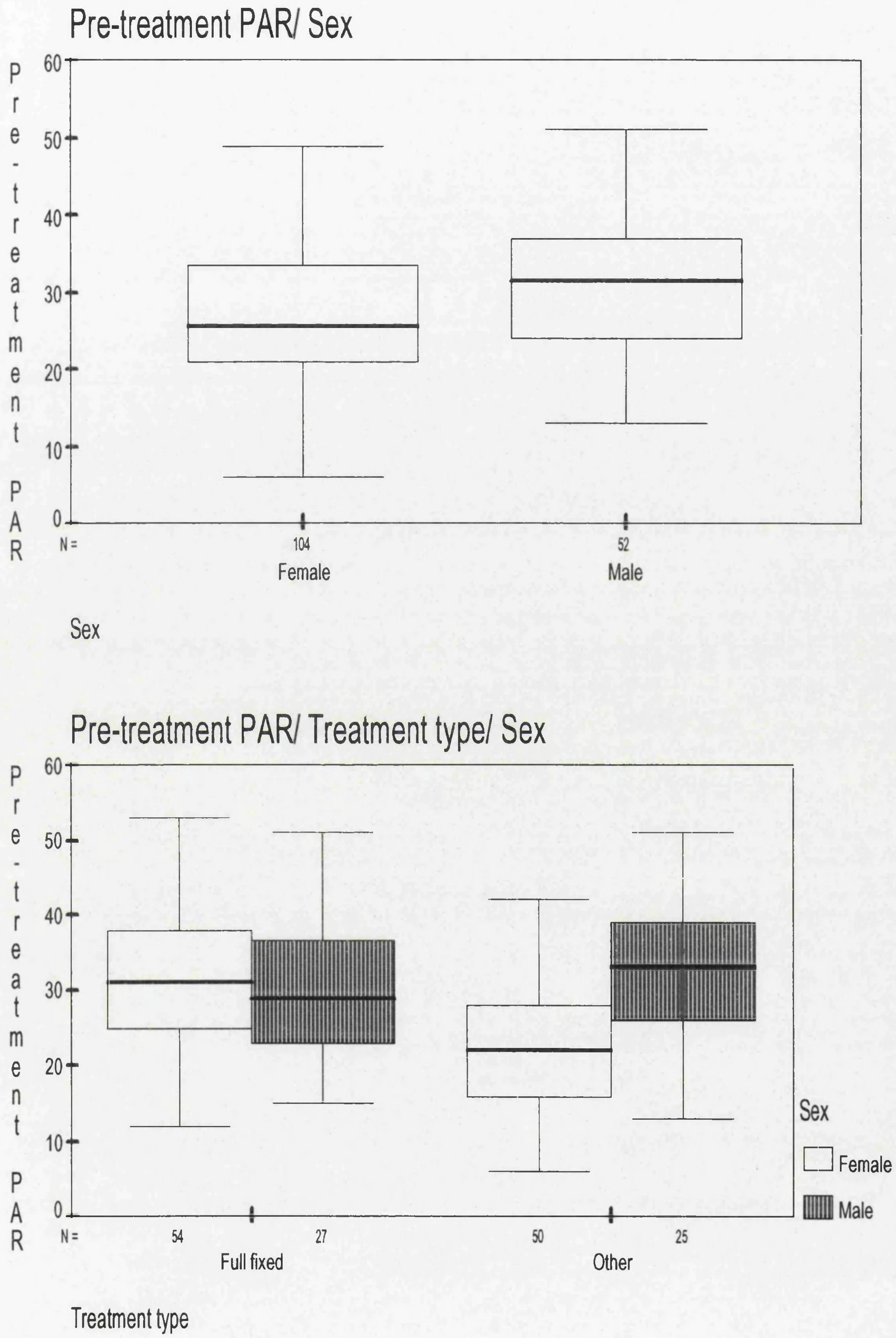
Table 17b

The distribution of pre-treatment PAR scores grouped for sex is shown in Figure 13. The mean values for males and females were 31 (S.D. 8.66) and 27 (S.D. 9.99) respectively which were found to be significantly different at the 1% level (Table 18).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	568.08	1	568.08	6.20	.01
SEX	568.08	1	568.08	6.20	.01
Explained	568.08	1	568.08	6.20	.01
Residual	14107.81	154	91.61		
Total	14675.89	155	94.68		

Table 18

The pre-treatment PAR distribution grouped for treatment type (Full fixed/Other) and sub-divided by sex is also shown in Figure 13. Values for females in the Other group appeared to be markedly lower than for males in the same group. The mean pre-treatment PAR for females not treated with full fixed appliances was 22.4 (S.D. 8.55) compared to 32.6 (S.D. 8.74) for males.



Pre-treatment PAR/ Sex
Pre-treatment PAR
60
50
40
30
20
10
0
N =
104
Female
52
Male
Sex
Pre-treatment PAR/ Treatment type/ Sex
Pre-treatment PAR
60
50
40
30
20
10
0
N =
54
27
Full fixed
50
25
Other
Treatment type
Sex
Female
Male

Figure 13

Significant interaction between sex and appliance type was found (Table 19).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	953.89	2	476.94	6.06	.003
FULL FIXED/OTHER	309.92	1	309.92	3.94	.049
SEX	632.45	1	632.45	8.04	.005
2-Way Interactions	1199.07	1	1199.07	15.25	.000
FF/OTHER SEX	1199.07	1	1199.07	15.25	.000
Explained	2722.90	3	907.63	11.54	.000
Residual	11952.99	152	78.64		
Total	14675.90	155	94.68		

Table 19

10.2.10 Oral hygiene

The numbers in each category of oral hygiene status for Full fixed and Other appliance groups are given in Table 20.

	Treatment type	
	Full fixed	Other
	Count	Count
Oral hygiene		
Good	19	26
Fair	34	20
Poor	12	11
No indication	16	18

Table 20

The difference between groups was not found to be significant (P=0.2).

A significant difference in standard of oral hygiene (P=0.05) was found between males and females, the females being reported as having better oral hygiene on average than males (Table 21)

Count Row%	Good	Fair	Poor	No ind.	Row total
Female	34 32.7%	35 33.7%	10 9.6%	25 24.0%	104 66.7%
Male	11 21.2%	19 36.5%	13 25.0%	9 17.3%	52 33.3%
Column total	45 28.8%	54 34.6%	23 14.7%	34 21.8%	156 100.0%

Chi-square=7.97 DF=3 P=0.05

Table 21

10.2.11 Number of broken appointments

The mean number of broken appointments per case ranged between 1.3 (Mini-fixed) and 1.8 (Full fixed). No significant difference was found between groups (P=0.7). 38% of cases had no record of any broken appointments and 75% had broken two or fewer. The maximum number of broken appointments in any single case was ten.

10.2.12 Number of removable appliances broken or lost

The maximum number of removable appliance breakages in a single case was four. Only two cases had more than two occasions where an appliance had been broken or lost. The frequencies of appliance breakage were:

Broken removable appliances

Number	Frequency	Percent
0	45	60.0
1	21	28.0
2	7	9.3
3	1	1.3
4	1	1.3
Total	75	100.0

Considering the number of removables appliances lost or broken with respect to sex, within the Removable only group males showed significantly more breakages than females (P=0.02, Table 22). The mean number of breakages for males was 1.1 (S.D. 1.09) and for females was 0.2 (S.D. 0.59).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	4.80	1	4.80	6.15	.02
SEX	4.80	1	4.80	6.15	.02
Explained	4.80	1	4.80	6.15	.02
Residual	21.86	28	.78		
Total	26.66	29	.92		

Table 22

10.2.13 Number of fixed appliance breakages

Fixed attachments were placed in a total of 126 cases. 73% of these had at least one bracket or band dislodged or archwire fracture and over 50% had two or more breakages (Table 23).

Number of breakages	Frequency	Percent	Cumulative Percent
0	34	27.0	27.0
1	27	21.4	48.4
2	12	9.5	57.9
3	8	6.3	64.3
4	8	6.3	70.6
5	9	7.1	77.8
6	10	7.9	85.7
7	3	2.4	88.1
8	1	.8	88.9
9	1	.8	89.7
10	3	2.4	92.1
11	2	1.6	93.7
12	2	1.6	95.2
15	2	1.6	96.8
17	1	.8	97.6
18	1	.8	98.4
22	1	.8	99.2
24	1	.8	100.0
Total	126	100.0	

Table 23

There was no significant difference between males and females in the number of episodes of damage (Table 24).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	50.00	1	50.00	1.912	.171
SEX	50.00	1	50.00	1.912	.171
Explained	50.00	1	50.00	1.912	.171
Residual	2065.55	79	26.14		
Total	2115.55	80	26.44		

Table 24

10.2.14 Altered treatment plan

The original treatment plan was altered in 16.7% of the total sample. Similar proportions were found when appliance groups were considered both separately and when combined (Table 25). No significant difference was demonstrated.

Count Column %	**Full fixed**	**Removable only**	**Removable /mini-fixed**	**Mini-fixed only**	**Row total**
Treatment plan maintained	69 85.2%	27 90.0%	17 73.9%	17 73.9%	130 83.8%
Treatment plan altered	12 14.8%	3 10.0%	6 26.1%	5 22.7%	26 16.7%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=3.21 DF=3 *P*=0.36

Count Column %	**Full fixed**	**Other**	Row total
Treatment plan maintained	69 85.2%	61 81.3%	130 83.8%
Treatment plan altered	12 14.8%	14 18.7%	26 16.7%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=0.42 DF=1 *P*=0.52

Table 25

When sex was considered in conjunction with alteration in treatment plan and treatment type, no association was found between sex and whether or not the treatment plan was altered (Table 26).

	OVERALL		FULL FIXED		OTHER	
Count Column%	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Maintained	84 80.8%	46 88.5%	46 85.2%	23 85.2%	38 76.0%	23 92%
Altered	20 19.2%	6 11.5%	8 14.8%	4 14.8%	12 24.0%	2 8.0%
Chi-square	1.47692		1.00000		2.81030	
Significance	0.224		1.000		0.094	

Table 26

10.3 Descriptive analysis: Treatment factors

10.3.1 Operator type

Table 27 details the distribution of operators within the various appliance groups. The distribution is illustrated by Figure 14.

Count Col%	**Full fixed**	**Removable only**	**Removable/ mini-fixed**	**Mini-fixed only**	Row total
Consultant	21 15.9%	4 13.3%	8 34.8%	10 45.5%	43 27.6%
S.R.	23 28.4%	4 13.3%	5 21.7%	8 36.4%	40 25.6%
Postgrad.	37 45.7%	2 6.7%	6 26.1%	4 18.2%	49 31.4%
Undergrad.		20 66.7%	4 17.4%		24 15.4%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Count Col%	**Full fixed**	**Other**	Row total
Consultant	21 15.9%	22 29.3%	43 27.6%
S.R.	23 28.4%	17 23.7%	40 25.6%
Postgrad.	37 45.7%	12 16.0%	49 31.4%
Undergrad.		24 32.0%	24 15.4%
Column total	81 51.9%	75 48.1%	156 100.0%

Table 27

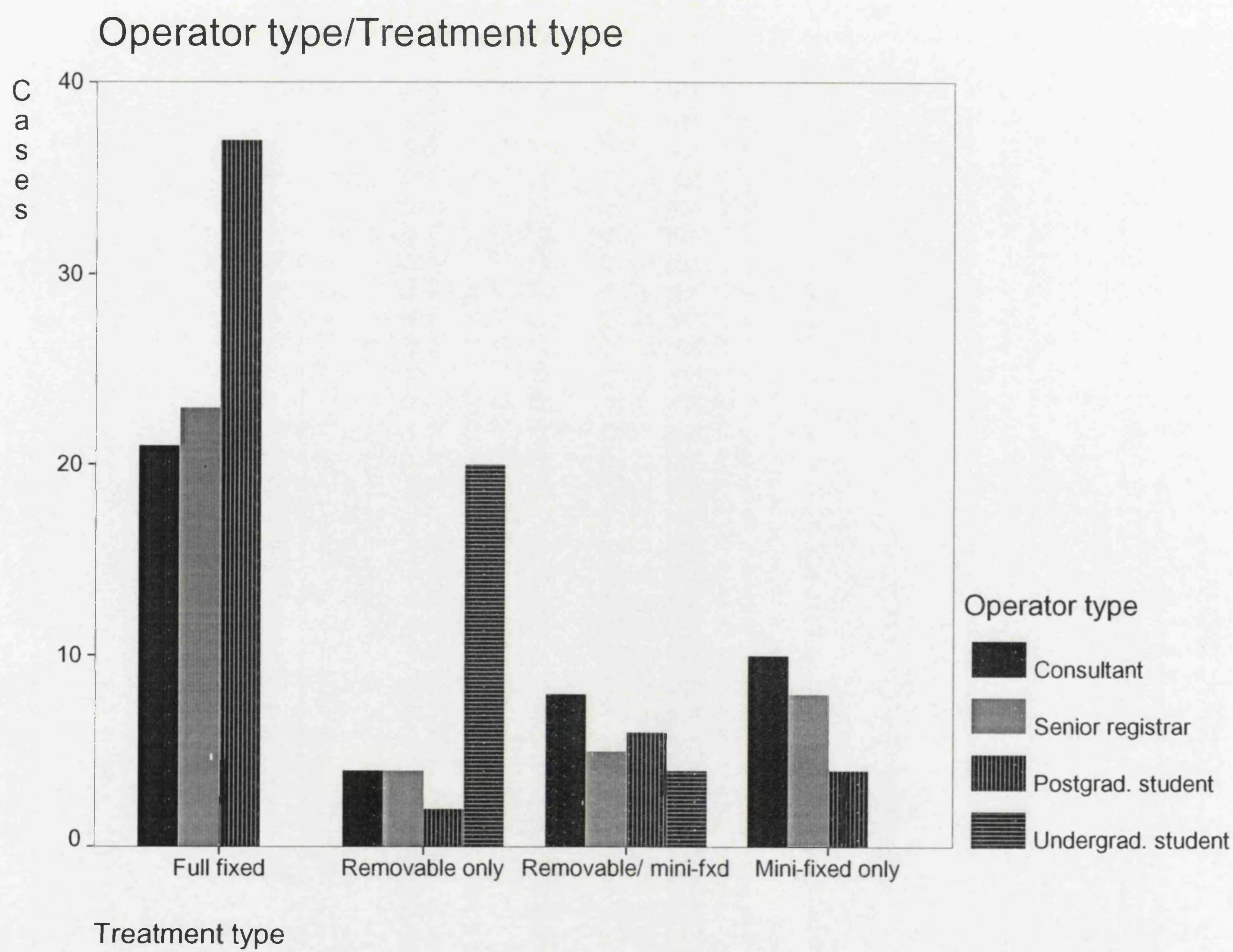

Operator type/Treatment type
C
a
s
e
s
40
30
20
10
0
Full fixed
Removable only
Removable/ mini-fxd
Mini-fixed only
Treatment type
Operator type
Consultant
Senior registrar
Postgrad. student
Undergrad. student

Figure 14

10.3.2 Extraction pattern

Dental extraction of some sort was carried out in 63.1% of cases when considering the total study sample. Of all those cases where teeth were removed as part of treatment, 37% involved the extraction of one premolar in each quadrant, 13% included the removal of at least one first permanent molar, and some other combination of teeth was extracted in the remaining 49% (Table 28).

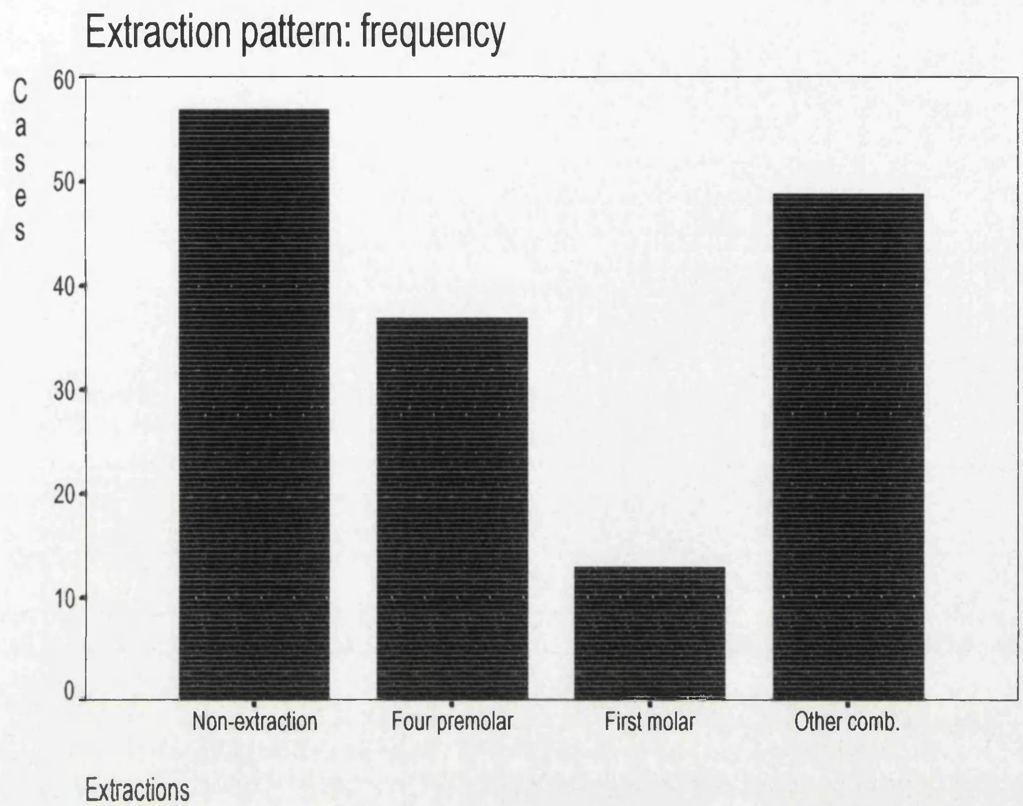
Count Col%	**Full fixed**	**Removable only**	**Removable/ mini-fixed**	**Mini-fixed only**	Row total
Non-extraction	16 19.8%	13 43.3%	13 56.5%	15 68.2%	57 36.9%
Four premolars	33 40.7%%	2 6.7%	1 4.3%	1 2.7%	37 23.7%
First molar included	9 11.1%	1 3.3%	2 8.7%	1 4.5%	13 8.3%
Other combination	23 28.4%	14 46.7%	8 34.8%	5 22.7%	49 31.4%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Count Col%	**Full fixed**	**Other**	Row total
Non-extraction	16 19.8%	41 54.7%	57 36.9%
Four premolars	33 40.7%	4 5.3%	37 23.7%
First molar included	9 11.1%	4 5.3%	13 8.3%
Other combination	23 28.4%	26 34.7%	49 31.4%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=35.62 DF=3 $P<0.001$

Table 28

Considerable differences between treatment groups were readily apparent (Figure 15).


Extraction pattern: frequency
Cases
60
50
40
30
20
10
0
Non-extraction
Four premolar
First molar
Other comb.
Extractions

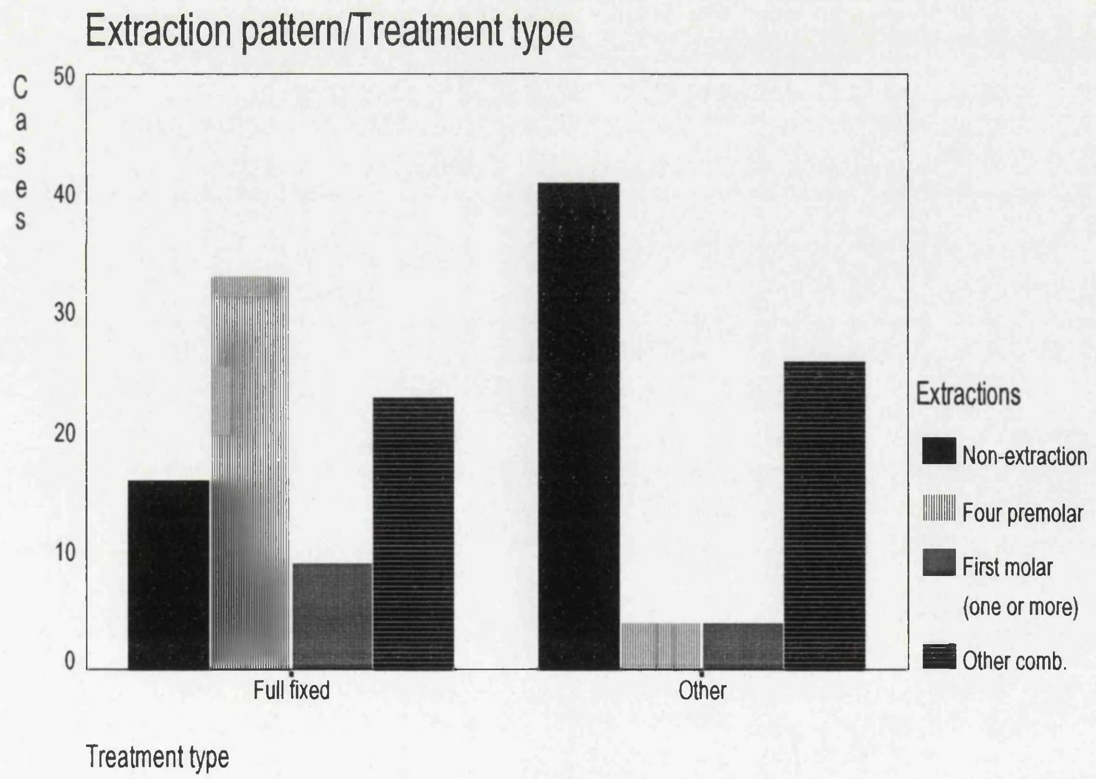

Extraction pattern/Treatment type
Cases
50
40
30
20
10
0
Full fixed
Other
Treatment type
Extractions
Non-extraction
Four premolar
First molar
(one or more)
Other comb.

Figure 15

10.3.3 Number of archwires used

The figures given here are restricted to the Full fixed appliance group only.

Number of archwires used	Frequency	Percent
3	1	1.2
4	3	3.7
5	8	9.9
6	3	3.7
7	5	6.2
8	6	7.4
9	11	13.6
10	7	8.6
11	7	8.6
12	6	7.4
13	11	13.6
14	6	7.4
15	6	7.4
21	1	1.2
Total	81	100.0

The number of archwires used during treatment ranged from 3 to 21 with approximately 50% of cases requiring between 9 and 13 archwires. The mean value for archwires used was approximately 10 with a standard deviation of 3.5.

It may be appropriate to detail the proportions of wide bracket and narrow bracket appliances used in the Full fixed group. Narrow bracket appliances might be expected to require fewer archwires during alignment because of longer inter-bracket spans and auxiliary springs may be used more frequently in the later stages rather than progressively increasing sizes of rectangular wire.

79% of full fixed appliances were of a wide bracket type, the remaining 21% used either Begg or Beddtiot brackets. The average number of archwires used per case in the narrow bracket group was 6.53 (S.D. 3.10) with a maximum of 12, compared to an average of 10.99 (S.D. 2.99) for the wide bracket group. One-way analysis of variance indicated a significant difference in mean number of archwires used between the wide and narrow bracket appliances (Table 29).

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	262.86	262.86	28.88	.0000
Within Groups	79	719.09	9.10		
Total	80	981.95			

Group	Count	Mean	S.D.	Minimum	Maximum
Wide	64	10.95	2.99	5	21
Narrow	17	6.53	3.10	3	12
Total	81	10.02	3.50	3	21

Table 29

10.3.4 Number of removable appliances

The table below (Table 30) refers to all treatments from the overall sample where at least one removable appliance was provided, excluding those treatments which fall into the Full fixed appliance group.

Number of appliances used	Frequency	Percent	Cum Percent
1	20	37.7	37.7
2	20	37.7	75.5
3	11	20.8	96.2
4	1	1.9	98.1
5	1	1.9	100.0
Total	53	100.0	

Mean=1.92 S.D.=0.92

Table 30

For all cases within the Other group, mean durations of active treatment according to number of appliances used were:

One appliance	12.5 months	(S.D. 6.86)
Two appliances	16.3 months	(S.D. 6.73)
Three appliances	18.8 months	(S.D. 7.43)

The mean durations in the removables only group were:

One appliance	7.7 months (S.D. 4.14)
Two appliances	14.6 months (S.D. 3.60)
Three appliances	17.9 months (S.D. 7.12)

10.3.5 Use of headgear

Of the total sample only 29 cases involved the use of headgear at some stage during the treatment.

The distribution of headgear use among the treatment type groups appeared remarkably random with no association being demonstrated when non-full fixed groups were considered either separately (*P*=0.74) or combined (*P*=0.40).

Count Col%	**Full fixed**	**Removable only**	**Removable/ mini-fixed**	**Mini-fixed only**	Row total
No headgear	68 84.0%	24 80.0%	17 73.9%	18 81.8%	127 81.4%
Headgear	13 16.0%	6 20.0%	6 26.1%	4 18.2%	29 18.6%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Chi-square=0.72 DF=3 *P*=0.40

Count Col%	**Full fixed**	**Other**	Row total
No headgear	68 84.0%	59 78.7%	127 81.4%
Headgear	13 16.0%	16 21.3%	29 18.6%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=1.24 DF=1 *P*=0.74

Table 31

10.3.6 Number of appointments during active treatment

The overall mean for the sample was 16.6 appointments with a standard deviation of 6.69 and a range from 2 to 36. A difference between separate treatment type groups and between full fixed and non-full fixed groups was apparent (Figure 16). Generally speaking the full fixed appliance cases required the greatest number of appointments.

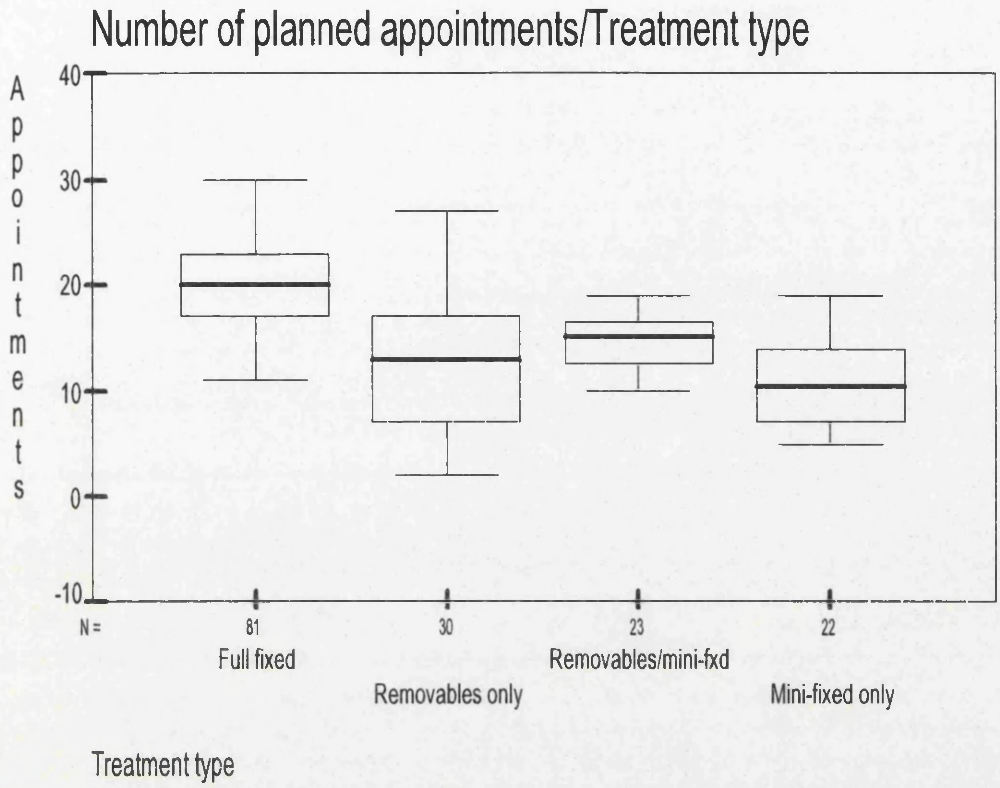
The average number of appointments for the individual treatment types was:

Full fixed	20.1	S.D. 5.63
Removable only	12.3	S.D. 6.43
Removable/mini fixed	15.7	S.D. 4.95
Mini-fixed only	10.9	S.D. 4.29
Other	*13.0*	*S.D. 5.70*

One-way ANOVA using Scheffé's multiple comparison procedure indicated significant differences between the Full fixed appliance group and all the other groups, also between the Mini-fixed and Removable/mini-fixed appliance groups (Table 32).

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	2279.97	759.99	24.81	.00
Within Groups	152	4655.34	30.63		
Total	155	6935.30			

Table 32


Number of planned appointments/Treatment type
Appointments
40
30
20
10
0
-10
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type

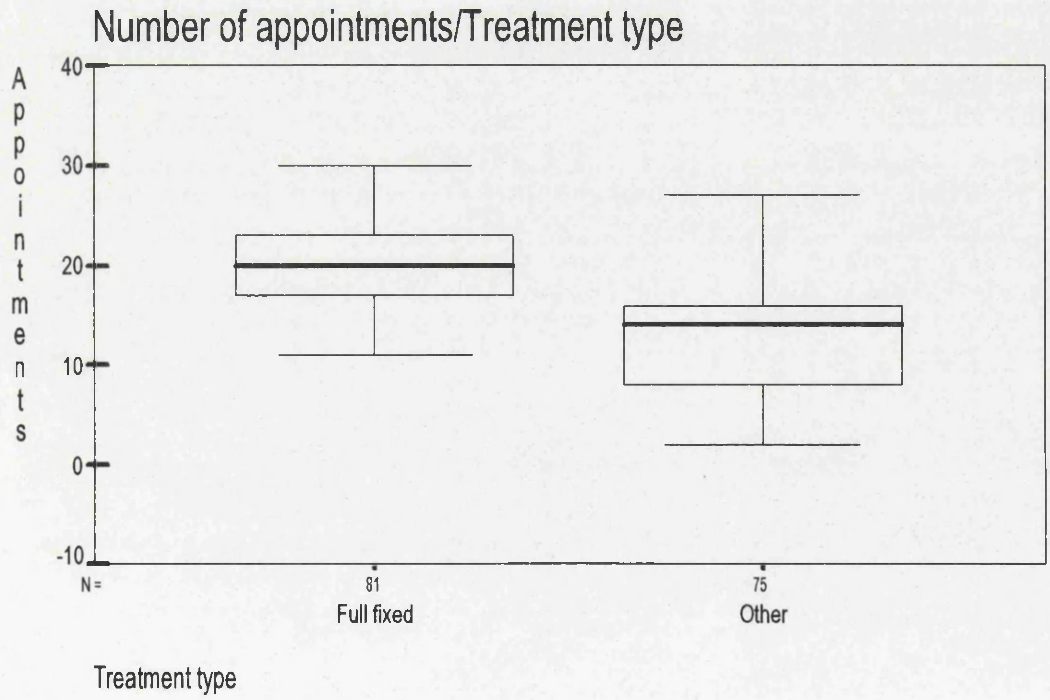

Number of appointments/Treatment type
Appointments
40
30
20
10
0
-10
N =
81
75
Full fixed
Other
Treatment type

Figure 16

10.4 <u>Descriptive analysis: Outcomes</u>

10.4.1 Post-treatment PAR score

The mean post-treatment PAR score for the complete sample was 11.3 (S.D. 5.78) with a maximum of 32 and a minimum of 2. The distribution of post-treatment PAR scores among treatment groups is illustrated in Figure 17.

Mean post-treatment PAR values for each appliance group were:

Full fixed	9.9	S.D. 5.28
Removable only	13.7	S.D. 5.55
Removable/mini-fixed	13.4	S.D. 6.72
Mini-fixed only	11.1	S.D. 5.43
Other	*12.6*	*S.D. 5.93*

ANOVA identified a significant difference between groups (*P*=.004) whilst use of the Scheffé multiple comparison procedure located a significant difference at the 0.05 level between the Full fixed appliance group and the Removables only group (Table 33).

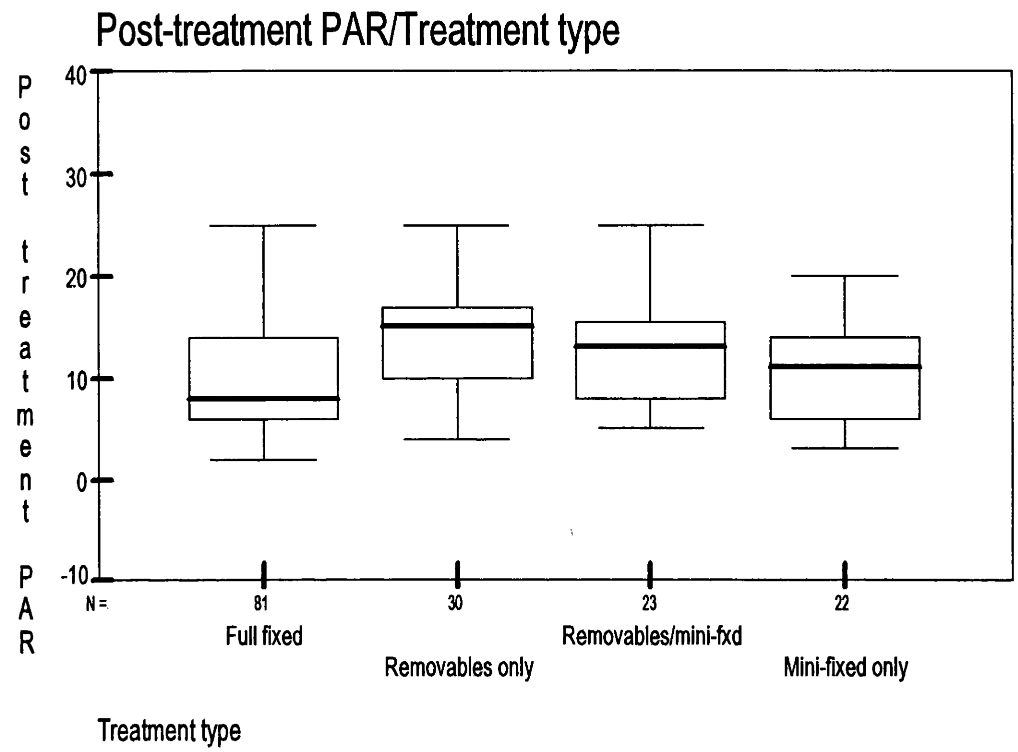
Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob
Between Groups	3	438.93	146.31	4.69	.004
Within Groups		152	4737.73	31.16	
Total		155	5176.67		

Table 33

The Full fixed appliances also differed significantly from the combined Other treatment groups (*P*=0.001, Table 34).

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	339.60	339.60	10.81	.001
Within Groups	154	4837.06	31.41		
Total	155	5176.67			

Table 34


Post-treatment PAR/Treatment type
Post treatment PAR
40
30
20
10
0
-10
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type

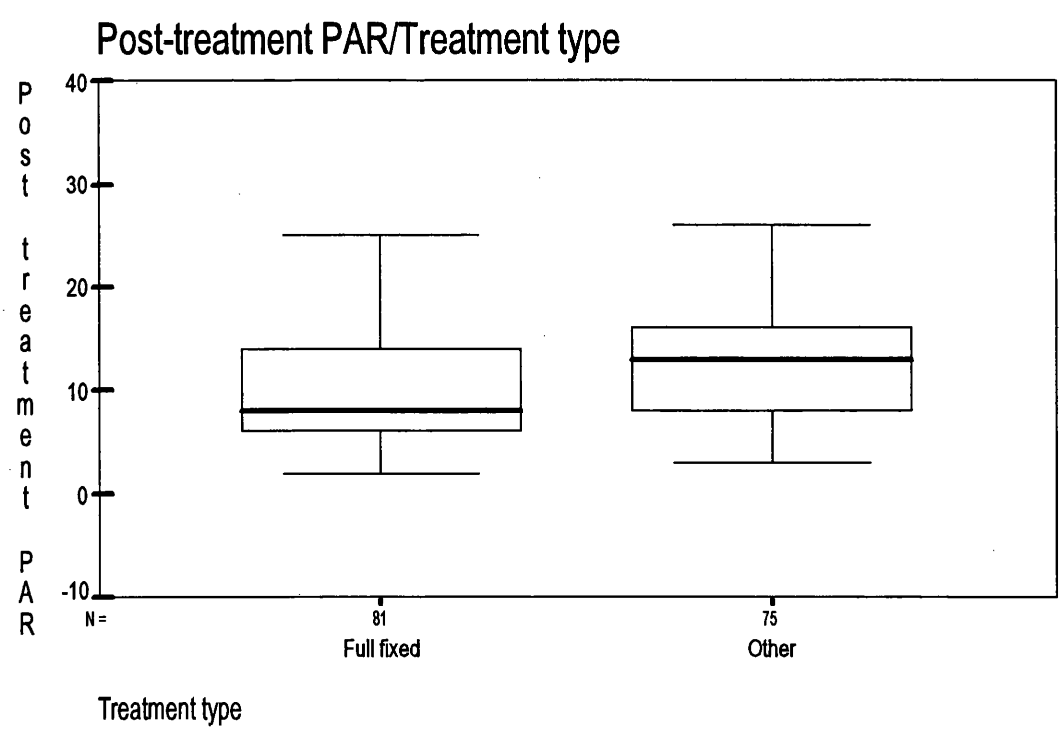

Post-treatment PAR/Treatment type
Post treatment PAR
40
30
20
10
0
-10
N =
81
75
Full fixed
Other
Treatment type

Figure 17

Relating post-treatment PAR scores to sex, for the total sample there appeared to be very little difference between the two sexes (Figure 18), although the mean value for males was 11.8 (S.D. 5.70), being higher than that for females which was 11.1 (S.D. 5.87).

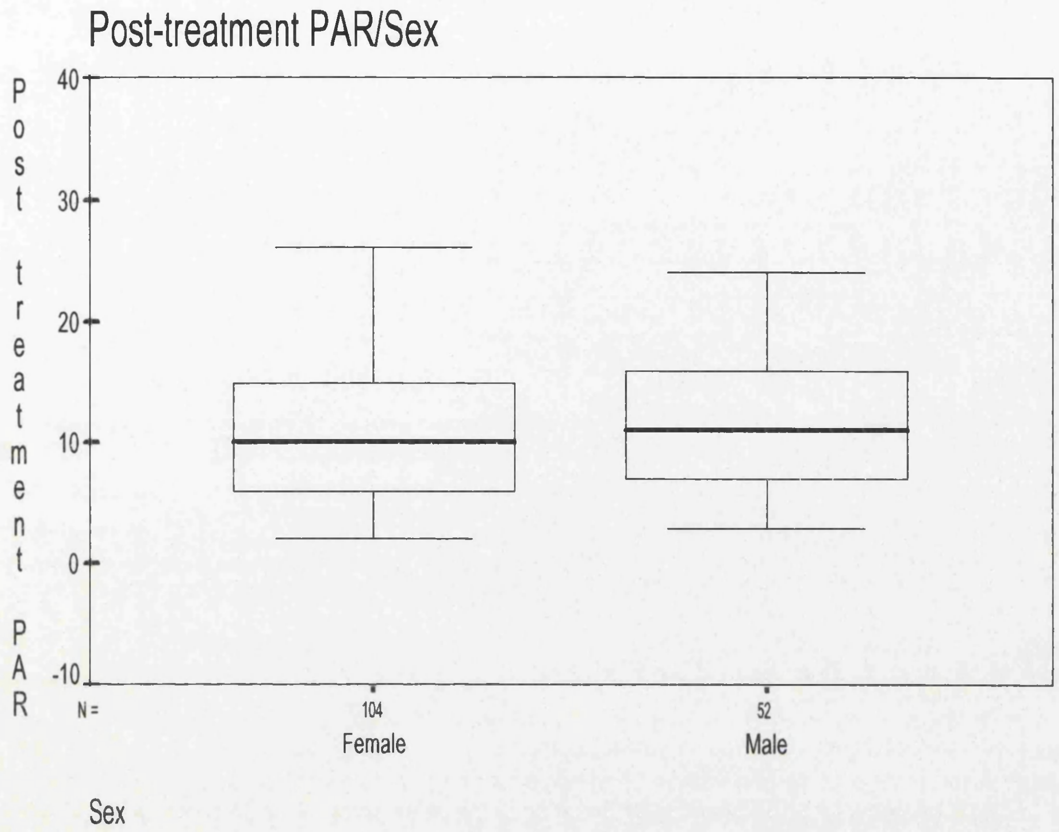
Comparison of post-treatment PAR scores for male and females separately in the Full fixed and Other appliance groups, where the proportions of each sex were close to identical (Figure 18), showed that males had slightly worse final results in the Other category but very slightly better results in the Full fixed category. By factorial ANOVA these tendencies were shown not to be significant (Table 35).

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	358.39	2	179.19	5.66	.004
SEX	15.87	1	15.87	.50	.480
FULL FIXED/OTHER	344.36	1	344.36	10.88	.001
2-Way Interactions	12.58	1	12.58	.40	.529
SEX FF/OTHER	12.58	1	12.58	.40	.529
Explained	367.01	3	122.34	3.87	.011
Residual	4809.67	152	31.64		
Total	5176.67	155	33.40		

Table 35

The difference between the Full fixed and Other groups was highly significant (P=0.001).

Overall correlation between pre-treatment PAR and post-treatment PAR was low (r=0.13). Plotting post-treatment PAR scores against pre-treatment scores resulted in an overall distribution of points without any distinct pattern. The majority of Full fixed points were concentrated into a horizontal band below the PAR score of 10 level (Figure 19).


Post-treatment PAR/Sex
Post treatment PAR
40
30
20
10
0
-10
N =
104
52
Female
Male
Sex

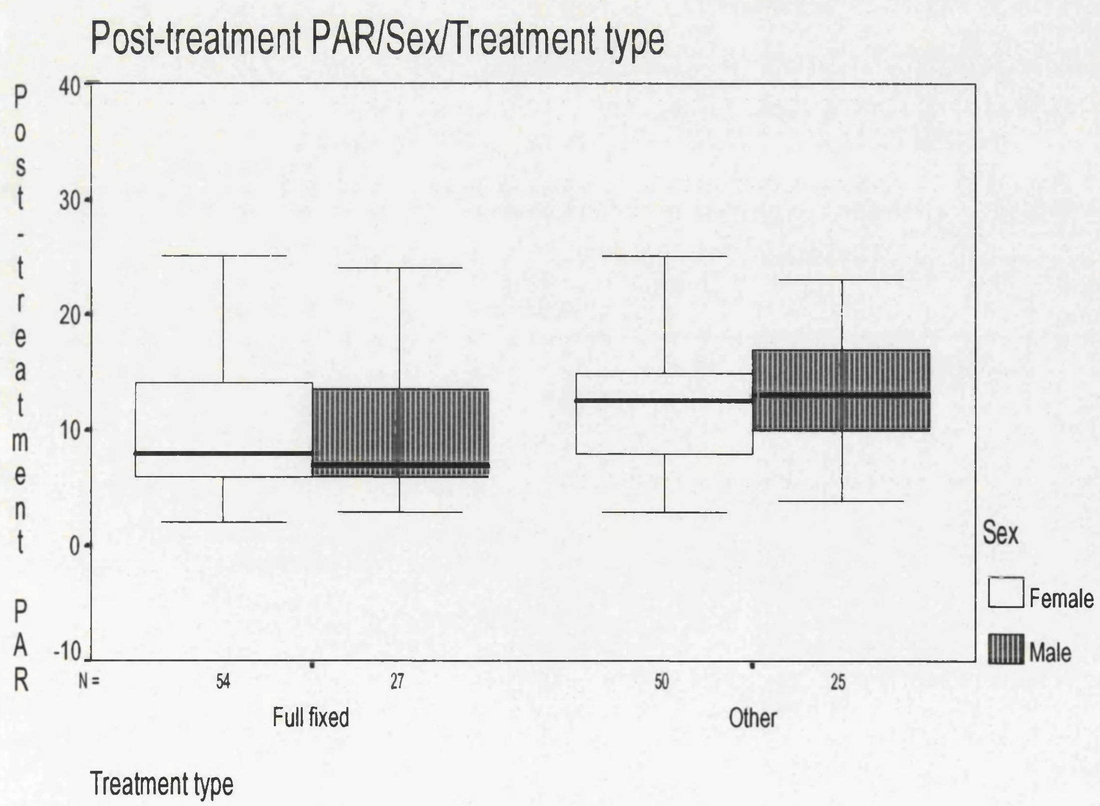

Post-treatment PAR/Sex/Treatment type
Post-treatment PAR
40
30
20
10
0
-10
N =
54
27
50
25
Full fixed
Other
Treatment type
Sex
Female
Male

Figure 18

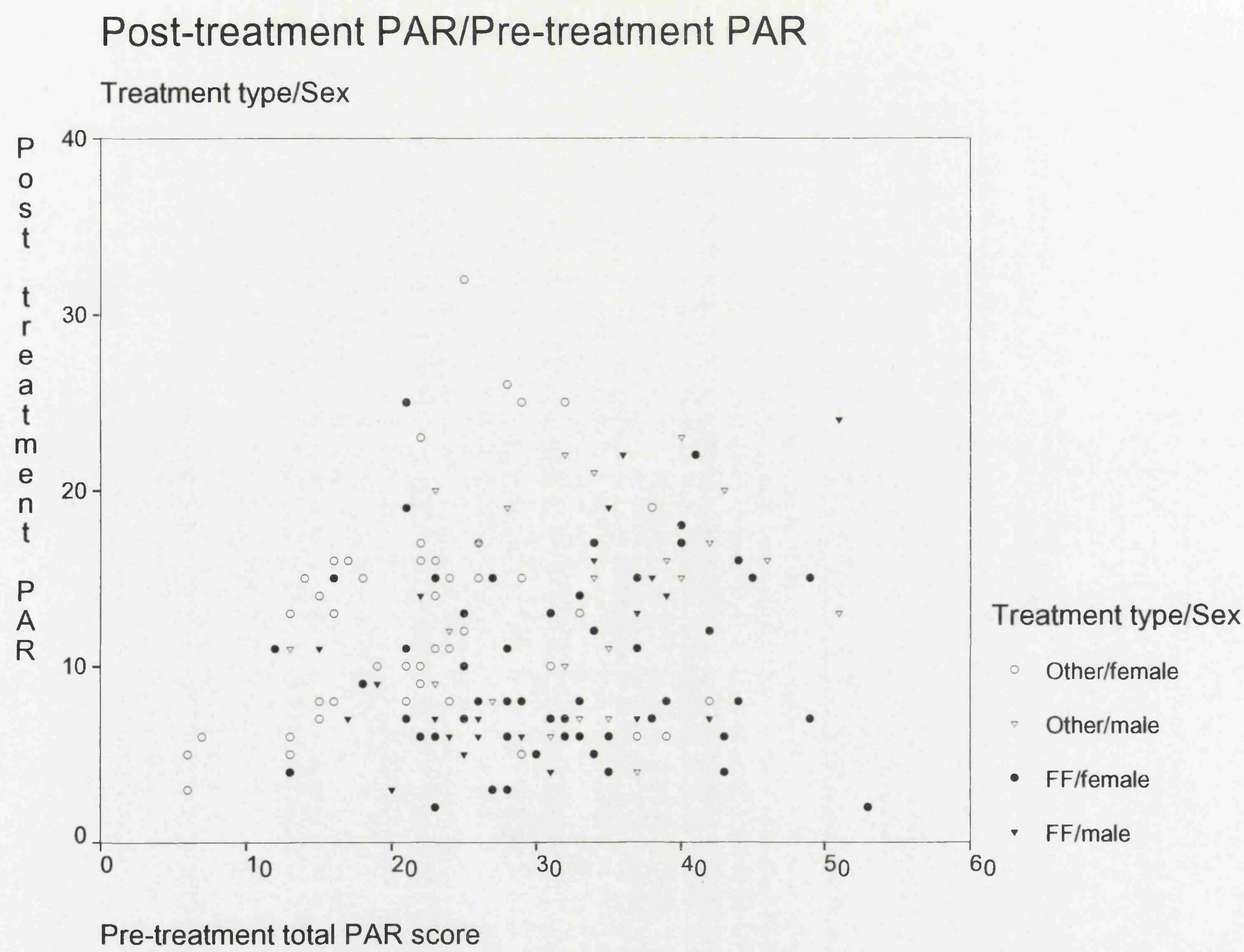

Post-treatment PAR/Pre-treatment PAR
Treatment type/Sex
Post treatment PAR
40
30
20
10
0
0
10
20
30
40
50
60
Pre-treatment total PAR score
Treatment type/Sex
Other/female
Other/male
FF/female
FF/male

Figure 19

The points for the Other cases did not demonstrate this horizontal pattern but seemed more randomly scattered. However, when the plotted points for males and females within the Other group were compared, the females were observed to have generally lower pre- and post-treatment PAR scores.

10.4.2 Change in total PAR score

The overall mean value for difference between pre- and post- PAR scores was 17.0 PAR points (S.D. 10.63) with a maximum of 51 and a minimum of -7.

The appliance group means for change in PAR points were:

Full fixed	20.8	S.D. 9.54
Removable only	15.0	S.D. 9.85
Removable/mini-fixed	13.0	S.D. 11.90
Mini-fixed only	10.0	S.D. 8.61
Other	*12.9*	*S.D. 10.27*

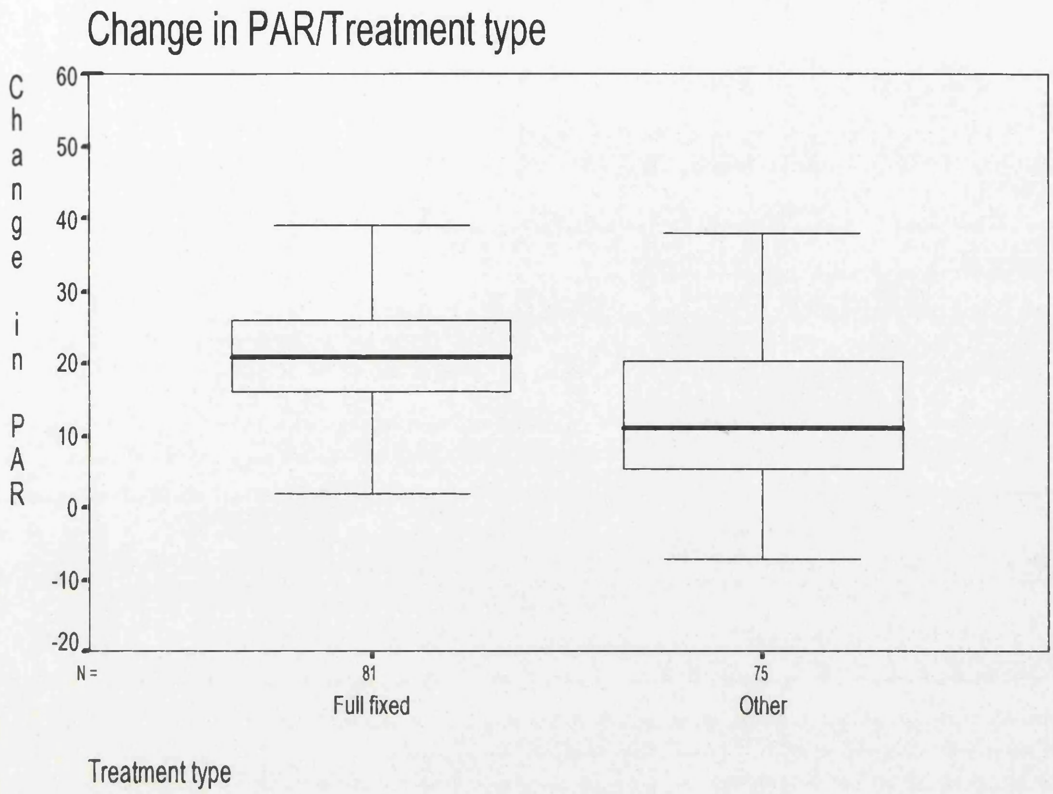
ANOVA identified a significant difference between groups (*P*=0.001) whilst use of a Scheffé multiple comparisons procedure located significant differences at the 0.05 level between the Full fixed appliance group and the Mini-fixed and Removables/mini-fixed groups (Table 36, Figure 20)

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob
Between Groups	3	2752.32	917.44	9.44	.001
Within Groups	152	14773.58	97.19		
Total	155	17525.90			

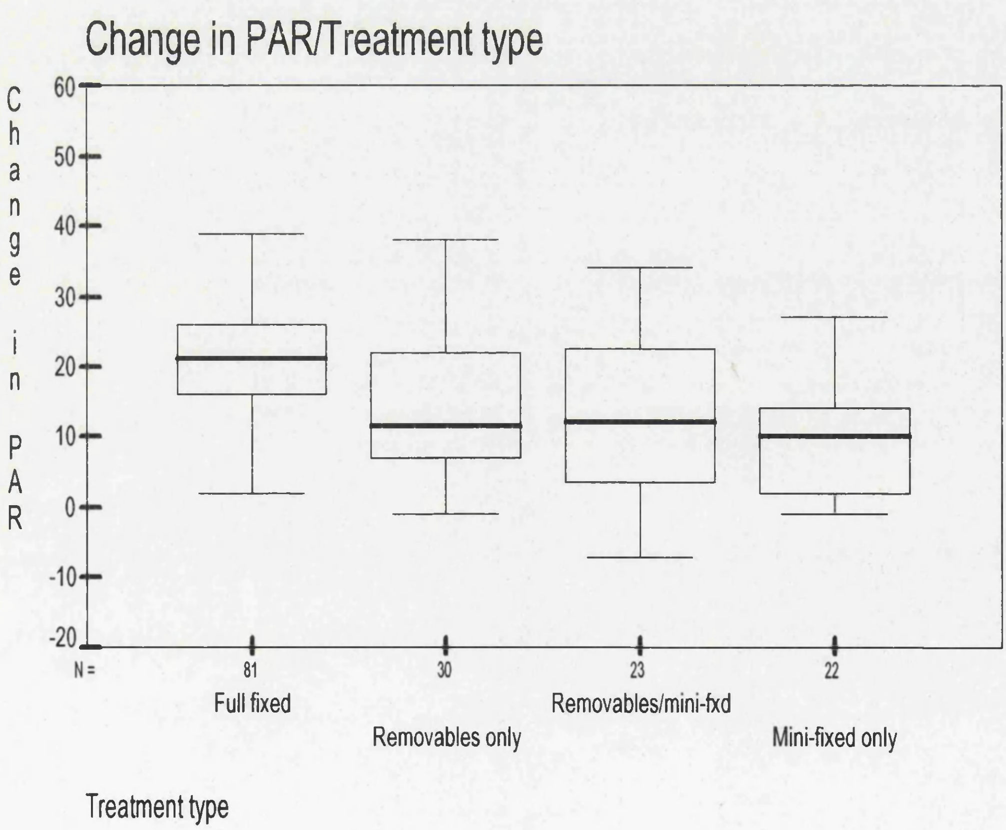
Table 36

In relating change in PAR to sex for the overall sample, males had a higher mean PAR reduction than females, the figure for males being 19.3 (S.D. 8.11) and that for females was 15.9 (S.D. 11.56). This difference was close to significance at the 5% level (*P*=0.06).

Within the Full fixed group the mean change in PAR values for male and female were 19.7 (S.D. 7.05) and 21.4 (S.D. 10.58) respectively. By one-way ANOVA this difference in means was found not to be significant (*P*=0.457).



Change in PAR/Treatment type
Change in PAR
60
50
40
30
20
10
0
-10
-20
N =
81
75
Full fixed
Other
Treatment type



Change in PAR/Treatment type
Change in PAR
60
50
40
30
20
10
0
-10
-20
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type

Figure 20

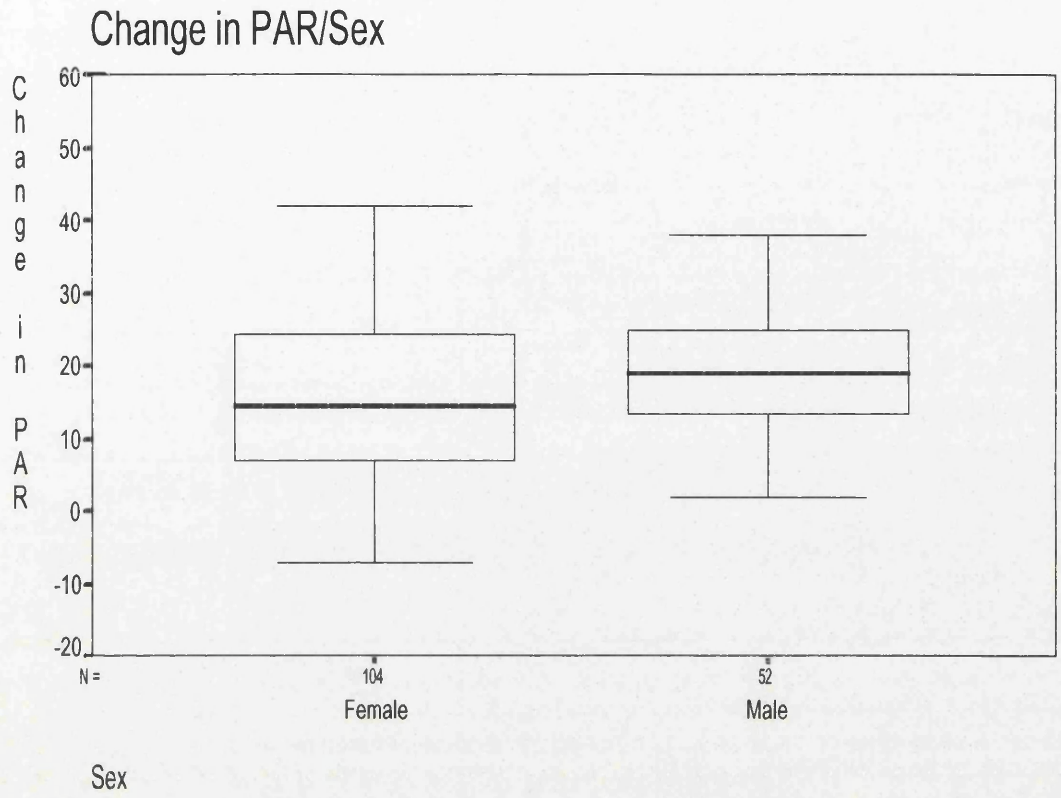
However in the non-full fixed appliance group there was clear sexual dimorphism with regard to change in PAR (Figure 21). Within this group the mean for males was 18.8 (S.D. 9.24) which was found to be significantly different ($P<0.001$) to that for females at 10.0 (S.D. 9.52).

Factorial ANOVA indicated a significant interaction between appliance type and sex (Table 37) showing that change in PAR was jointly affected by sex and appliance type.

Source	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	1775.55	2	887.77	9.83	.000
FF/OTHER	1307.66	1	1307.66	14.48	.000
SEX	447.97	1	447.97	4.96	.037
2-Way Interactions	965.97	1	965.97	10.70	.001
FF/OTHER SEX	965.97	1	965.97	10.70	.001
Explained	3800.15	3	1266.72	14.03	.000
Residual	13725.74	152	90.30		
Total	17525.90	155	113.07		

Table 37

Change in PAR was plotted against pre-treatment PAR (Figure 22). A strong correlation ($r=0.84$) was noted. The resultant scatterplot could be divided into two areas along the line of association. Points representing cases with pre-treatment PAR scores greater than 26 and change in PAR scores greater than 15 consisted mainly of points formed by Full fixed appliance cases. Cases with pre-treatment PAR scores and change in PAR scores less than 26 and 15 respectively were predominantly Other appliance types.


Change in PAR/Sex
Change in PAR
60
50
40
30
20
10
0
-10
-20
N =
104
52
Female
Male
Sex

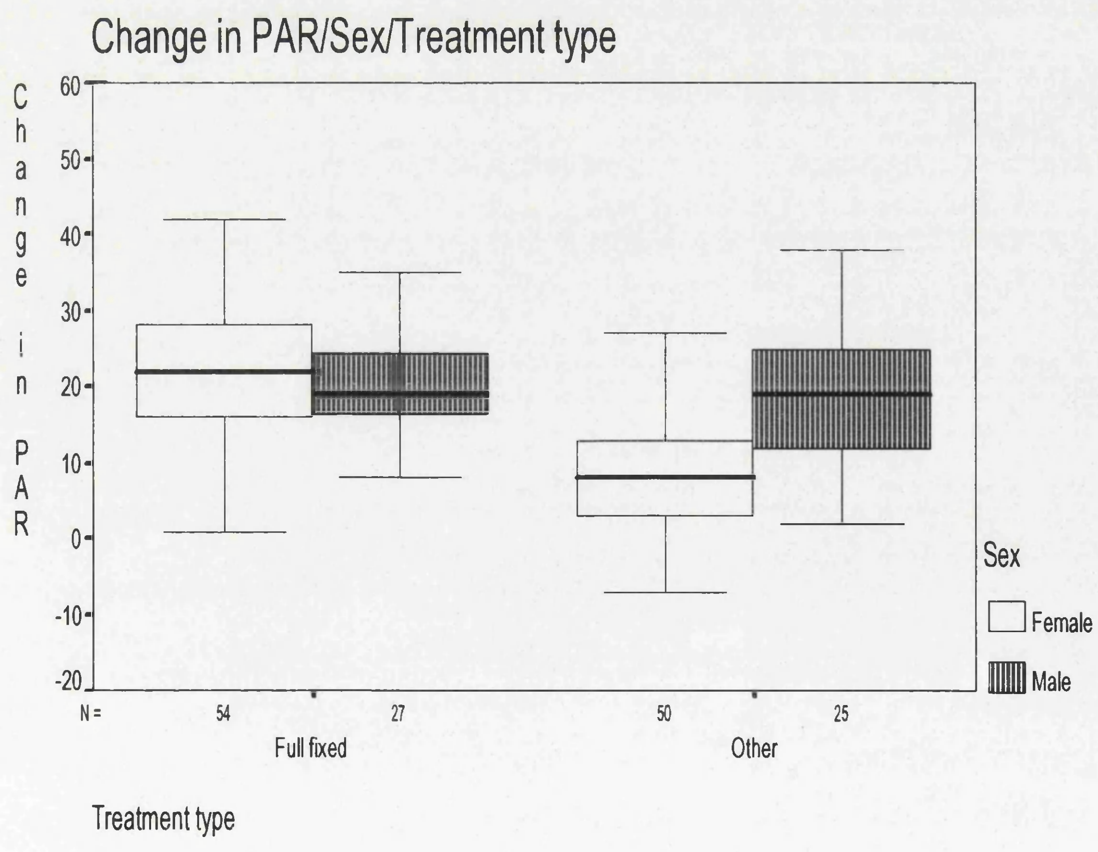

Change in PAR/Sex/Treatment type
Change in PAR
60
50
40
30
20
10
0
-10
-20
N =
54
27
50
25
Full fixed
Other
Treatment type
Sex
Female
Male

Figure 21

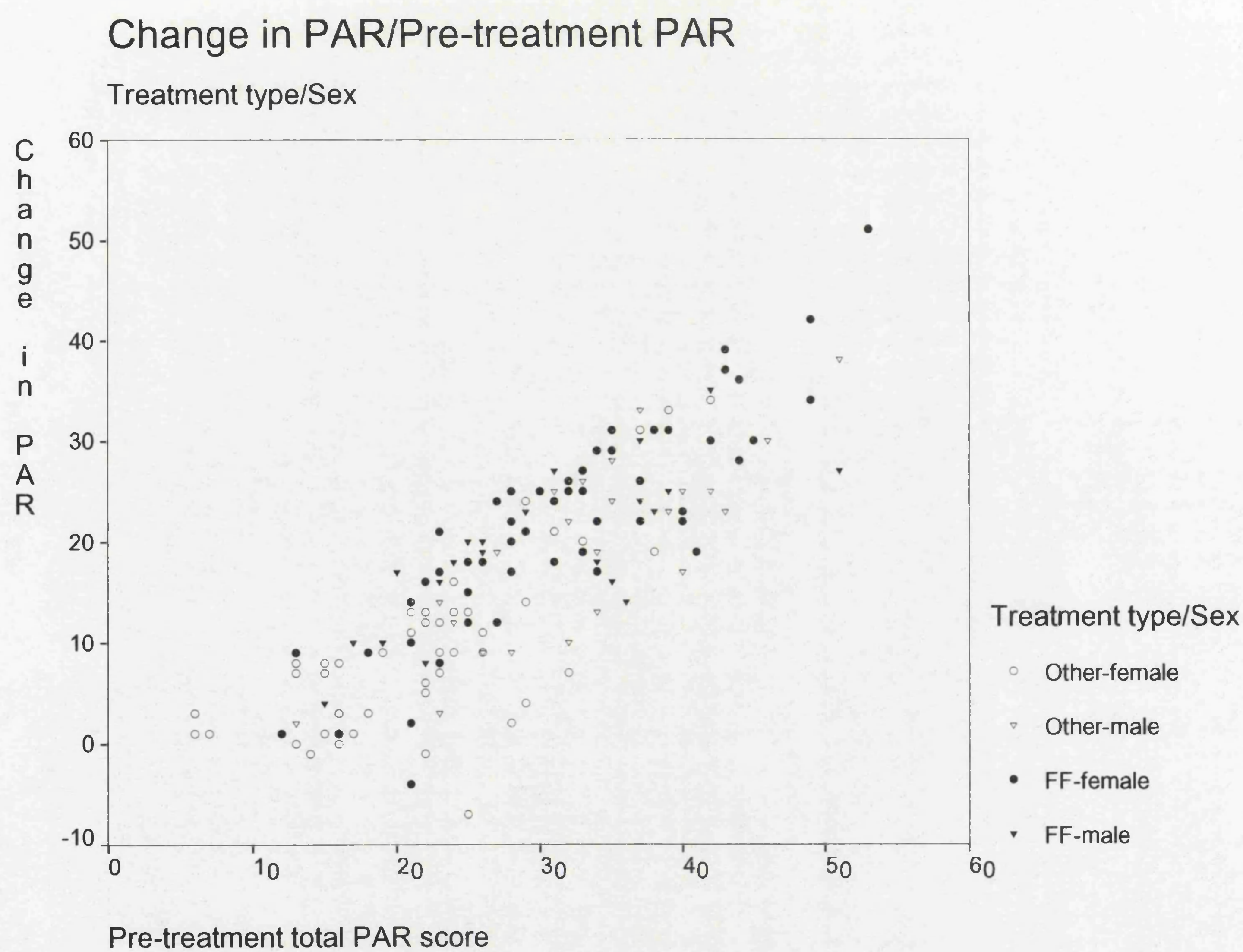

Change in PAR/Pre-treatment PAR
Treatment type/Sex
Change in PAR
60
50
40
30
20
10
0
-10
0
10
20
30
40
50
60
Pre-treatment total PAR score
Treatment type/Sex
Other-female
Other-male
FF-female
FF-male

Figure 22

The numbers and proportions of cases falling into each of the three nomogram categories following treatment are given in Table 38.

Count Col. %	**Full fixed**	**Removable only**	**Removable /mini-fxd**	**Mini-fixed only**	Row total
Worse/No different	4 4.9%	5 16.7%	7 30.4%	8 36.4%	24 15.4%
Improved	39 48.1%	18 60.0%	10 43.5%	10 45.5%	77 49.4%
Greatly improved	38 46.9%	7 23.3%	6 26.1%	4 18.2%	55 35.3%
Column total	81 51.9%	30 19.2%	23 14.7%	22 14.1%	156 100.0%

Count Col. %	**Full fixed**	**Other**	Row total
Worse/No different	4 4.9%	20 26.7%	24 15.4%
Improved	39 48.2%	38 50.7%	77 49.4%
Greatly improved	38 46.9%	17 22.6%	55 35.2%
Column total	81 51.9%	75 48.1%	156 100.0%

Chi-square=18.49 DF=2 $P<0.001$

Table 38

The proportion of "Improved" cases in the Full fixed and Other groups was similar, however there was a significant difference ($P<0.001$) in the proportion of "Worse/No different" and "Greatly improved" cases found between groups. Less than five percent of Full fixed cases were "Worse/No different" compared with approximately 27 percent of non-full fixed cases. Conversely 47 percent of the Full fixed group were found to be "Greatly improved" compared to approximately 23 percent of the Other group.

The overall mean percentage change in PAR for the sample was 55.4 percent (S.D. 25.83). Higher mean percentage change in PAR was found in the Full fixed group compared to other groups and this difference was found to be significant (*P*<0.001):

Full fixed	65.4%	S.D. 21.00
Removable only	48.7%	S.D. 22.44
Removable/mini-fixed	42.7%	S.D. 31.74
Mini-fixed only	41.3%	S.D. 25.70
Other	*44.7%*	*S.D. 26.37*

10.4.3 Duration of active treatment

The overall mean duration of active treatment was 18.3 months (S.D. 7.86) with a maximum of 49 months and a minimum of 1 month.

The possibility of this very low minimum figure being due to an error during data collection or entry was recognised. The minimum case was identified and was found to be a simple correction of a single tooth in anterior crossbite with an upper removable appliance. The duration of active treatment for this case was confirmed.

The mean durations and standard deviations in months for the different appliance groups were:

Full fixed	21.4	S.D. 6.47
Removable only	14.3	S.D. 8.66
Removable/mini-fixed	18.3	S.D. 7.15
Mini-fixed only	12.4	S.D. 6.48
Other	*14.9*	*S.D. 7.88*

One-way ANOVA indicated that a significant difference between group means existed ($P<0.001$). A Scheffé multiple comparisons procedure demonstrated significant differences in means between the Full fixed and Removables only groups, and the Full fixed and Mini-fixed groups (Table 39)

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	2056.15	685.38	13.83	.0001
Within Groups	152	7529.83	49.54		
Total	155	9585.97			

Table 39

As might be anticipated a significant difference in means was also demonstrated between the Full fixed and combined non-full fixed groups (Table 40).

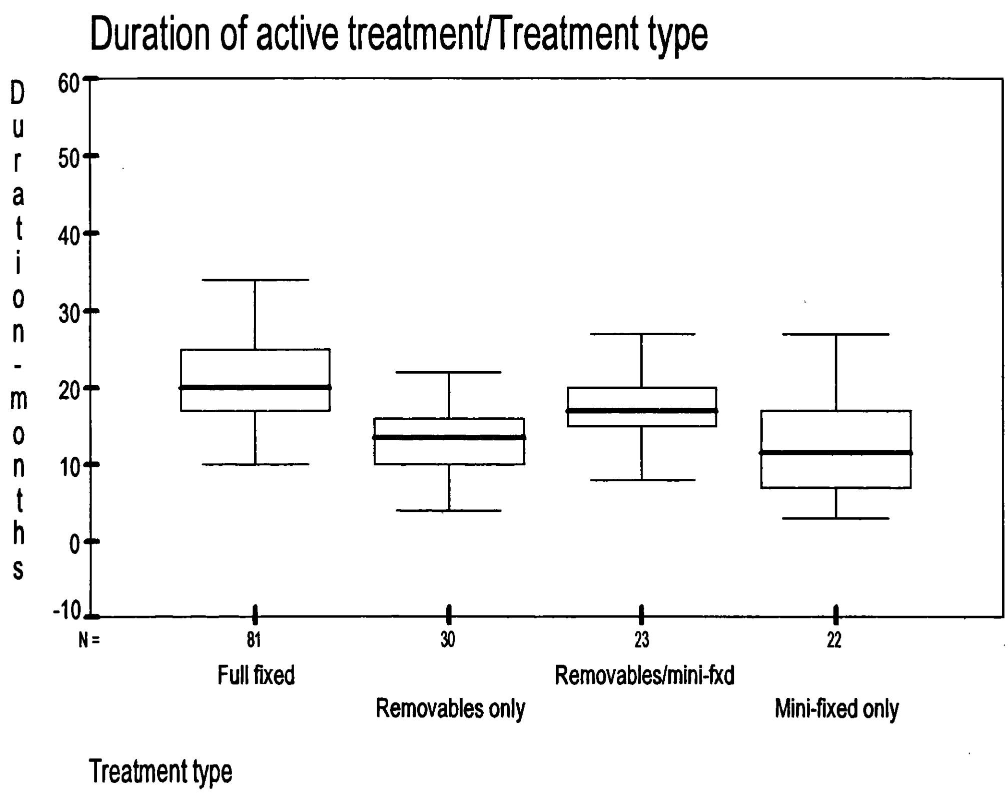
Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	1797.55	1797.55	33.34	.0001
Within Groups	154	8297.34	53.88		
Total	155	10094.90			

Table 40

The distribution of cases within each group is illustrated by box-and-whisker plots in Figure 23. The distributions of the Full fixed and Other groups appeared to be similar but with the Other group having a larger spread and generally shorter durations.

Duration of active treatment was plotted against pre-treatment PAR for Full fixed and Other groups (Figure 24). Overall correlation was noted to be weak (r=0.37).

The number of planned appointments during treatment and the duration of treatment were closely correlated (r=0.86), however there was a poorer correlation between number of appointments and change in PAR (r=0.38).


Duration of active treatment/Treatment type
Duration - months
60
50
40
30
20
10
0
-10
N =
81
30
23
22
Full fixed
Removables only
Removables/mini-fxd
Mini-fixed only
Treatment type

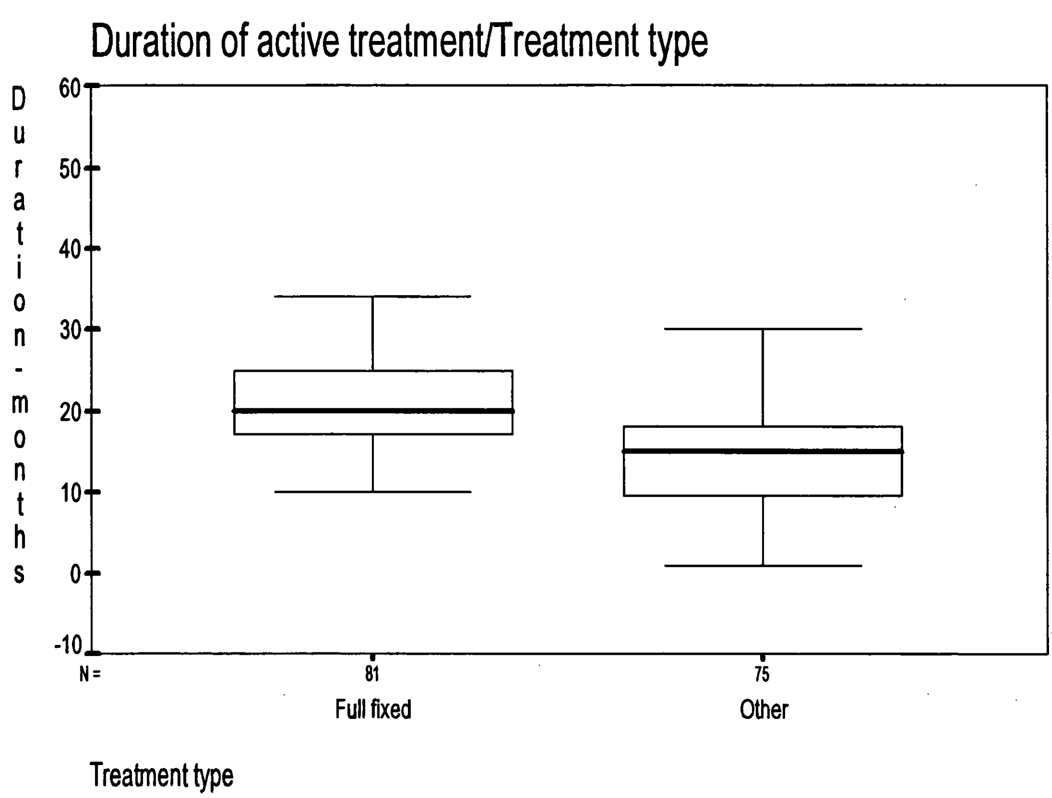
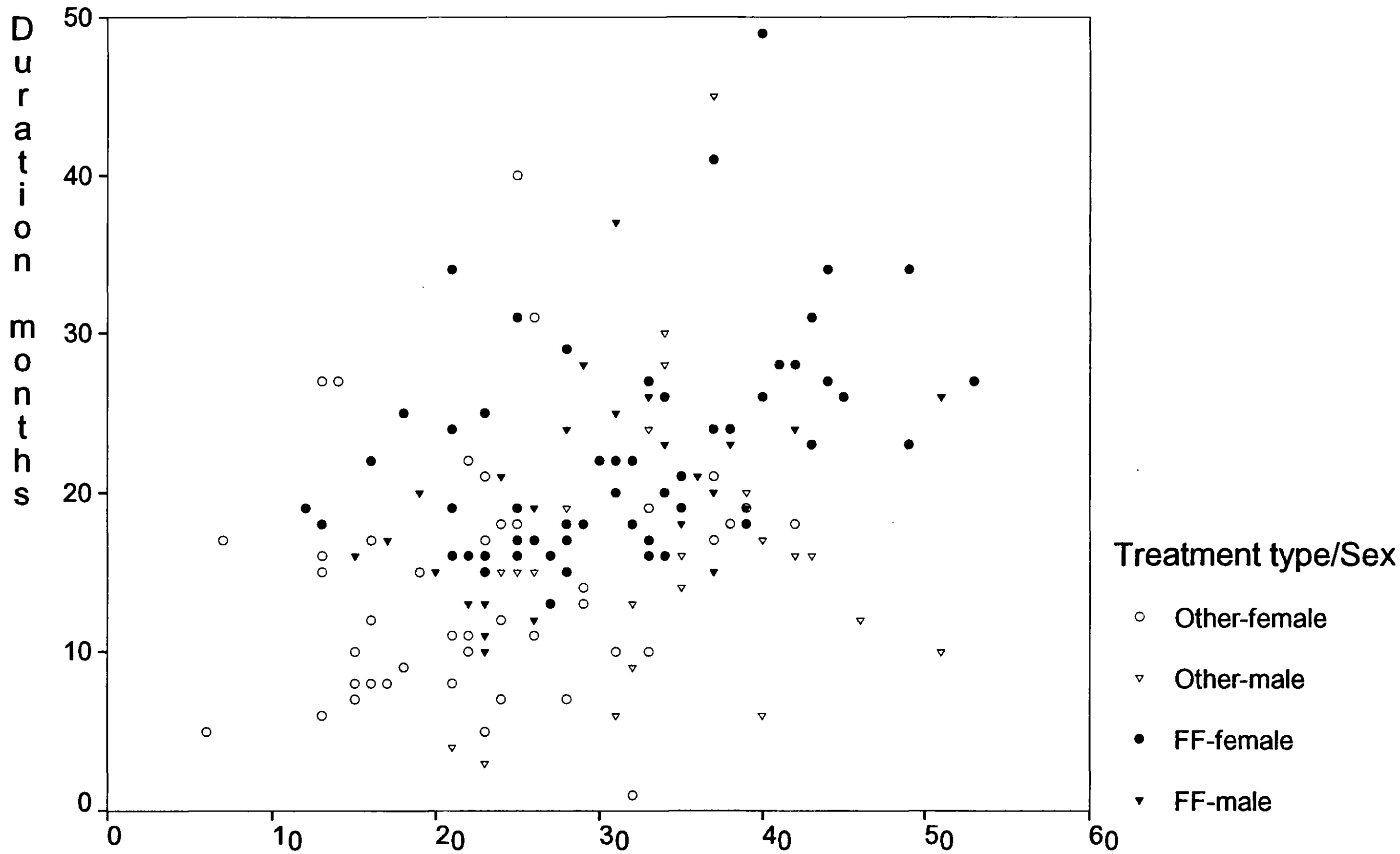

Duration of active treatment/Treatment type
Duration - months
60
50
40
30
20
10
0
-10
N =
81
75
Full fixed
Other
Treatment type

Figure 23

Duration of treatment/Pre-treatment PAR

Treatment type/Sex



Pre-treatment total PAR score

Figure 24

10.5 Descriptive analysis: Discussion

The cases for this retrospective study were taken at random and so comprised a wide range of unique malocclusions in a variety of individuals. It is to be expected that this sample would contain cases which might be regarded as atypical, yet because any set of randomly selected cases from the spectrum of orthodontic patients would contain such cases, these cases have been included with the other data for the descriptive investigation.

Although in the following discussion differences between various groups are highlighted, and the probabilities that these differences might not have been due to chance have been calculated, no inference as to cause and effect is made as there is great potential for confounding due to interactions between variables.

Differences were found between treatment type groups when considering personal factors and treatment variables.

10.5.1 Personal factors

Proportionally more males than females were provided with removable appliances solely, despite males showing a greater degree of initial occlusal irregularity. Also, within the Other group females had considerably lower pre-treatment PAR scores than males. This suggests that males with severe malocclusions might be given either full fixed appliances or removable appliances whereas females with such malocclusions would be less likely to receive non-full fixed appliances.

This may be due to full fixed appliances perhaps being more acceptable to female patients, different expectations of quality of result being held for males and females, or differing suitability for full fixed appliances because of co-operation considerations such as oral hygiene. On average, males in this study were reported as having poorer oral hygiene than females but once again subjective impression may have been influenced by preconceived ideas *e.g.* girls take more care of themselves than boys therefore classifying their oral hygiene as good is more likely to be "correct".

The wearing of fixed appliances demands a high standard of oral hygiene and inadequate oral hygiene might be put forward as a reason, albeit a dubious one, for selecting a removable appliance over a fixed appliance. It is interesting to note therefore that oral hygiene was less commonly rated as "Good" by the operator in the Full fixed appliance group than in the non-full fixed group. However this anomaly may be attributed to operator type as a substantial proportion of the removable and combination treatments were carried out by undergraduate students with a view to providing removable appliances.

Full fixed appliance treatments were usually carried out by postgraduate students or more senior staff intending to set complex mechanics perhaps more sensitive to patient negligence. A level of oral hygiene that might have been regarded as "Good" for a removable appliance may only have rated "Fair" where a fixed appliance was concerned.

Proportionally more females were given appliances of the Mini-fixed type than males. It could be conjectured that more females with minor localised anomalies such as single tooth rotations or midline diastema of the type suitable for treatment with Mini-fixed appliances would present for treatment than males with similar malocclusions. This assumes greater awareness or concern about visible dental irregularity on the part of females or their parents.

Patients in the Removables only group tended to begin treatment at a earlier age than in other appliance groups and significantly so when compared to Mini-fixed appliance patients. This is unsurprising as removable appliances are more likely to be used in mixed dentition procedures, as confirmed by the significantly higher proportion of mixed dentition starts in treatment groups involving removable appliances than in purely fixed appliance treatments.

Within the Other group the effects of the Removable and the Mini-fixed groups appeared to balance each other out and brought the mean starting age close to that of the Full fixed group. However the number of mixed dentition starts remained significantly higher in the Other group compared to that in the Full fixed group.

10.5.2 Occlusal factors

Differences between treatment groups were noted in respect of incisor classification and pre-treatment overjet although only the difference between the Full fixed and Other groups for pre-treatment overjet was shown to be significant.

It appeared that whereas Class II division 1 treatments were more numerous than Class I treatments in the Full fixed and Removable only groups, in groups where there was limited used of fixed components this position was reversed, perhaps indicating the suitability of the Mini-fixed approach to local, intra-arch anomalies or posterior transverse discrepancies rather than inter-maxillary, antero-posterior discrepancies.

This is supported by the significant differences observed in pre-treatment PAR score between the Mini-fixed and Full fixed and Mini-fixed and Removable appliances.

It is clear that in this study a significant difference existed in initial dental irregularity between the Full fixed and Other groups as judged by pre-treatment PAR. Any consideration of the comparative "efficiency" of appliance systems using PAR as a measure would have to take this difference into account.

19 cases out of the total sample exhibited the presence of one or more ectopic teeth with the highest proportion being found in the Removables/mini-fixed group, where an ectopic tooth was present in 30.4 percent of cases. Although this was found to be significant, as three out of the eight cells in the cross-table had an expected frequency of less than 5, this result might not be reliable.

16 out of the 19 cases involved maxillary canines. A difficulty in interpreting this result may lie in the definition of ectopia and how clinicians distinguish between displacement due to crowding and that due to developmental malposition. In the case of the misplaced maxillary canine the issue may be more clear cut than for other poorly positioned teeth and the diagnosis of ectopia perhaps easier to make.

Regarding IOTN, Grade four was the predominant Dental Health Component category in every group but particularly among the Removable group where over three quarters of cases fell into this category. The proportion of "borderline" need cases was lowest in this group. It might be suggested that "borderline" cases in terms of need of treatment *i.e.* DHC grade 3, might be more likely to be treated with systems involving fixed components than purely removable appliances. Any difference between groups in this respect could not be shown as significant.

Similarly no significant difference in distribution of DHC grades between cases grouped by sex could be demonstrated.

From the data recorded using the IOTN letter classification of occlusal traits, it could be seen that increased overjet and the presence of a crossbite accounted for most of the cases in all treatment groups although contact point displacement was a frequently recorded trait in the Mini-fixed group. There was a higher percentage of crossbites in the Removable only group. This may be due to the fact that anterior crossbites are readily apparent, tending to be recognised early and treated with simple removable appliances in the mixed dentition.

It should be remembered that the application of IOTN to study casts assumes that all crossbites are associated with a mandibular displacement of greater than 2 mm and so the number of crossbites rated as grade 4 may be artificially high.

10.5.3 Co-operation factors

The loss or breakage of a removable appliance was a frequent event. Of all the cases from the total sample where at least one removable appliance had been fitted, 40 percent had one or more recorded episodes of loss or breakage. However, having more than one loss or breakage was considerably less common at approximately 12 percent whereas more than two episodes was comparatively rare at approximately 3 percent. Of the 30 patients with removable appliance breakage, 70 percent experienced breakage on only one occasion. This agrees exactly with the findings of Kerr (1984) who reported that 307 out of 438 cases where breakage had occurred had just one episode of breakage.

The traditional view that males are generally less careful with their appliances than females was upheld with regard to removable appliances, a significant difference between male and female groups having been found, although with regard to the other co-operation measures used in this study little difference was found between the sexes.

Concerning fixed appliances significant differences in number of breakages were found between the Full fixed and Mini-fixed and Full fixed and Removable/mini-fixed groups ($P<0.05$). This is to be expected as it seems reasonable that the likelihood of breakage would increase with increasing number of components placed. Substantially more components would be present in Full fixed mechanics than in Mini-fixed appliances.

It is also possible that the simultaneous wearing of an upper removable appliance, perhaps incorporating a bite plane, with a lower fixed appliance affords the fixed components a degree of protection.

10.5.4 Treatment factors

Clear differences between treatment type groups existed in respect of all the designated treatment factors apart from the use of headgear which appeared remarkably random.

No undergraduate student carried out Full fixed or purely Mini-fixed treatments. Removable appliance treatments were largely carried out by undergraduate students whilst Full fixed treatments were most often performed by postgraduate students. A more even occurrence of operator type was found in the Removable/mini-fixed group. This distribution was dictated by the undergraduate and postgraduate educational requirements of a teaching dental hospital.

Regarding extraction pattern, the most pronounced difference was the relative proportion of four premolar extraction cases and non-extraction treatments within the Full fixed and Other appliance groups. The Full fixed appliances had a much higher percentage of four premolar cases whilst the percentages of non-extraction cases were considerably higher in the non-full fixed groups.

The low proportion of four premolar extraction cases in the non-full fixed groups should not be surprising given that frequently only one arch is tackled in this type of treatment and that interceptive removable appliance therapy may leave lower arch crowding untreated. A treatment involving extraction of a premolar in each of the upper quadrants was categorised as an "other combination". 7 out of 28 removable cases required the bilateral loss of an upper premolar without extraction in the lower arch.

A clear difference was also found between treatment groups in the number of appointments attended during active treatment in that Full fixed treatments required considerably more appointments compared to all the non-full fixed treatments. Remembering that the use of full fixed appliances generally involves more chairside time for each appointment, and noting that the average number of archwire changes was approximately 10, the increased number of appointments in full fixed treatments must represent a considerably greater commitment of clinical resources than alternative therapies. This fact, rightly or wrongly, may be another factor which influences a comparison of efficacy between various treatment modalities.

With reference to removable appliances, three-quarters of the Removable and Removable/mini-fixed appliance groups involved two or fewer removable appliances. The use of more than three appliances was a rare event.

The extra treatment time which each additional appliance represented was very similar to the rough guideline figures proposed by Kerr *et al.* (1993) of 6 months, 14 months and 20 months for 1, 2 and 3 appliance treatments respectively. In that study the numbers of cases in each group were considerably larger at 43, 31 and 22 than in the equivalent sub-groups of the present study (10, 10 and 9).

10.5.5 Outcomes

As measured by PAR, the Full fixed appliance cases tended to have consistently less residual irregularity than other appliance groups at the end of treatment. This was illustrated by the positively skewed distribution with median of less than ten observed among the Full fixed appliances. This is similar to the observation made by Fox in his personal audit of 100 cases (Fox, 1993).

However when the non-full fixed groups were considered separately, only the difference between the Full fixed and the Removable treatments was found to be significant. This lack of significance for other differences may have been a function of the smaller number of cases in each of the non-full fixed groups.

The sex difference in pre-treatment PAR within the Other appliances was accompanied by a similar difference in post-treatment PAR, the females having generally lower scores than males.

The improvement due to treatment in this study was measured by change in total PAR score. Clearly there was greater improvement in the Full fixed cases compared to the non-full fixed cases but the other differences between the two groups need to be taken into consideration. Perhaps most notable is the difference in the initial PAR score. Although the Full fixed group attained greater PAR reductions than the Other appliance group the Other group tended to have lower initial PAR scores.

The correlation between pre-treatment PAR and change in PAR was high, confirming the close association found by Kerr *et al.*(1993) in their study of removable appliance treatments. Kerr *et al.* also found a similar distribution of nomogram improvements through removable appliance therapy as was found in this study, with the majority of cases falling into the improved category but with 10 percent of cases being "Worse/No different". The mean percentage change for Removable only cases was close to that observed in this study at 52 percent.

Within the Other group regarding change in PAR a similar effect due to sex was found to that seen when initial PAR score had been considered, significantly larger reductions being observed in the males. The sexual dimorphism found for change in PAR is likely to be related to the difference in mean pre-treatment PAR scores for males and females given the high correlation between pre- and post-treatment PAR score.

The PAR change in removable appliances was almost identical to that found by Fox (1993) although in that study the mean pre-treatment PAR scores for removable appliances was marginally greater. The change in PAR scores found by Fox for full fixed appliances was higher at 24 than in this study (21) whilst the initial PAR scores were similar.

It is interesting that in the overall sample males showed a greater mean reduction in PAR score than females. The difference in means for change in PAR was 4 PAR points which was almost identical to the difference in means for pre-treatment PAR (3.6 PAR points). The greater reduction in PAR for males is perhaps at variance with the traditionally held view that females might be more co-operative and committed to their orthodontic treatment however it lends support to the theory that on average a boy's malocclusion has to be more severe before treatment is requested or offered.

In this study durations of treatment for Full fixed and Removable groups were close to those found by Tang and Wei (1990) (20.2 and 13.4 respectively) and but with larger standard deviations. This might indicate a greater number of adult orthodontic treatments and possibly a tendency to refer for orthodontic assessment at an earlier age.

Although the overall correlation between pre-treatment PAR and treatment duration was noted to be weak (r=0.37), examination of pre-treatment scores plotted against duration suggested a tendency for longer treatment durations to be associated with higher initial PAR scores.

Chapter 11

Regression analysis: Results

11.1 Post treatment PAR score

11.1.1 Post-treatment PAR: Full fixed appliances

a) All information (Post-treatment PAR/Full fixed)

Stepwise regression of all relevant independent variables was performed with post-treatment PAR score as the dependent variable.

The following variables were entered in the regression equation:

Whether or not the treatment plan had been altered during treatment

The duration of active treatment

Whether or not an anterior crossbite was present

R^2 for this model was low at 0.27 (R_a^2=0.24). Examination of the residuals indicated that the assumptions of normality and equal variance were acceptable.

Lilliefors test is based on a modification of the Kolmogorov-Smirnov test and provides a check on normality where population means and variances are not known but must be estimated from the data (Norusis, 1993). In this case Lilliefors test (0.07, P>0.20) did not indicate that the distribution of residuals differed significantly from normal.

A detrended normal plot suggested the existence of two cases which appeared atypical (Figure 25). These cases were identified and investigated.

Detrended Normal Plot of Residual

Full fixed appliances: Post-treatment PAR

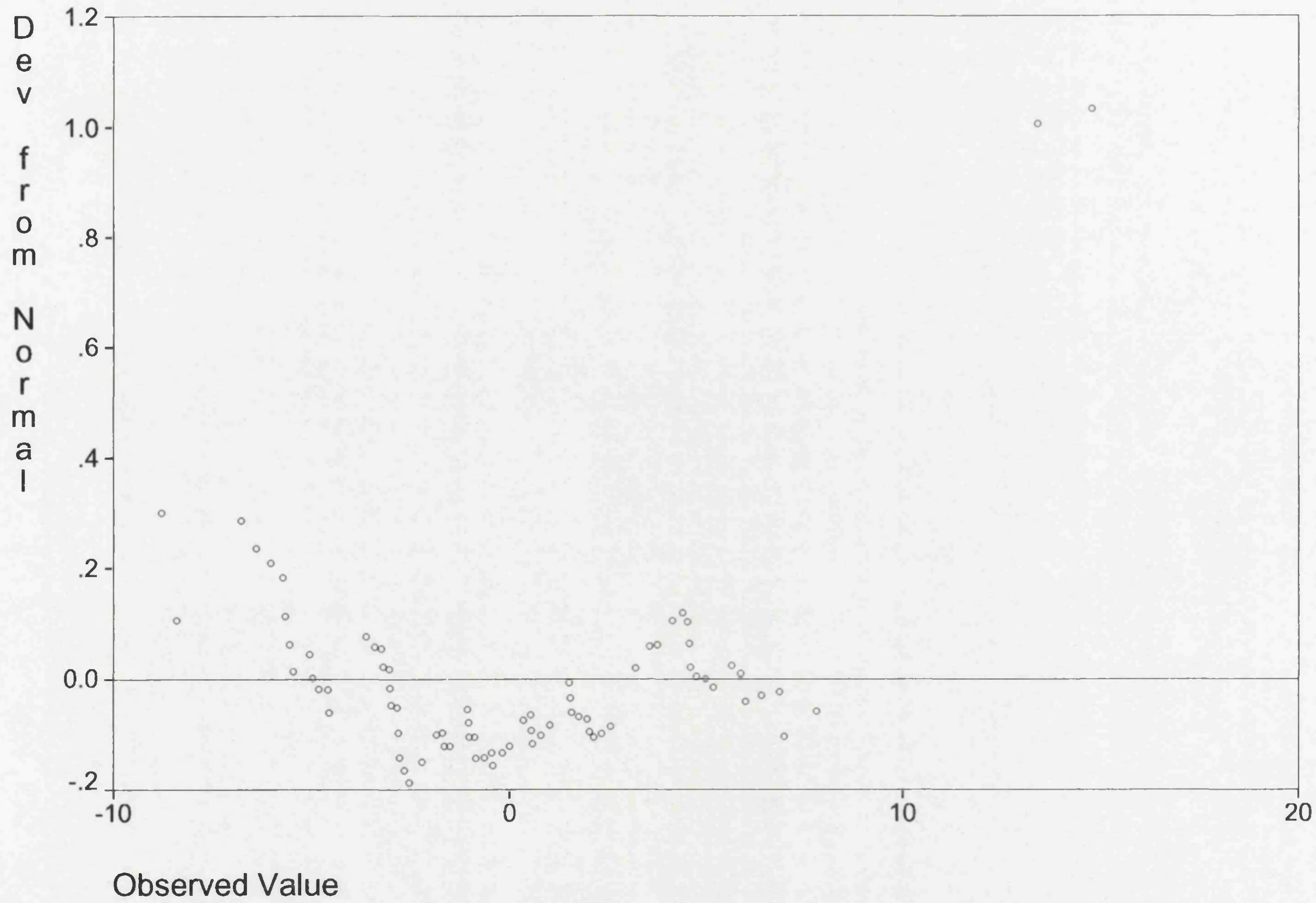


Figure 25

Both of the outlying cases were found to have had treatments dependant on the co-operation of the patient in the use of intra-arch elastics and that co-operation had not been forthcoming. No alteration in treatment plan had been recorded. The use of elastics and failure to comply in their use was not included in data collection for this study.

It was decided to repeat the regression having first eliminated the outlying cases.

The variables entered into this regression were:

Whether or not the treatment plan had been altered

Whether or not a first permanent molar had been extracted

The pre-treatment Aesthetic Component score

Whether or not an anterior crossbite was present

R^2 for this model was 0.40 (R_a^2=0.37) and the assumptions of regression did not appear to be substantially violated (Figures 26-27).

The constructed equation was:

POST-TREATMENT PAR =
4.65 + 6.32(T/P ALTERED) + 3.77(FIRST MOLAR XTN) +
0.60(AESTHETIC COMPONENT) + (- 2.75(ANTERIOR XB)) + ε

The Beta values suggesting the relative importance of each variable in the equation were:

Variable	Beta
ALTERED	.48
FIRST MOLAR EXTRACTED	.24
AESTHETIC COMPONENT	.24
ANTERIOR XB.	-.23

Normal(P-P) plot of standardised residual

Full fixed appliances: Post-treatment PAR

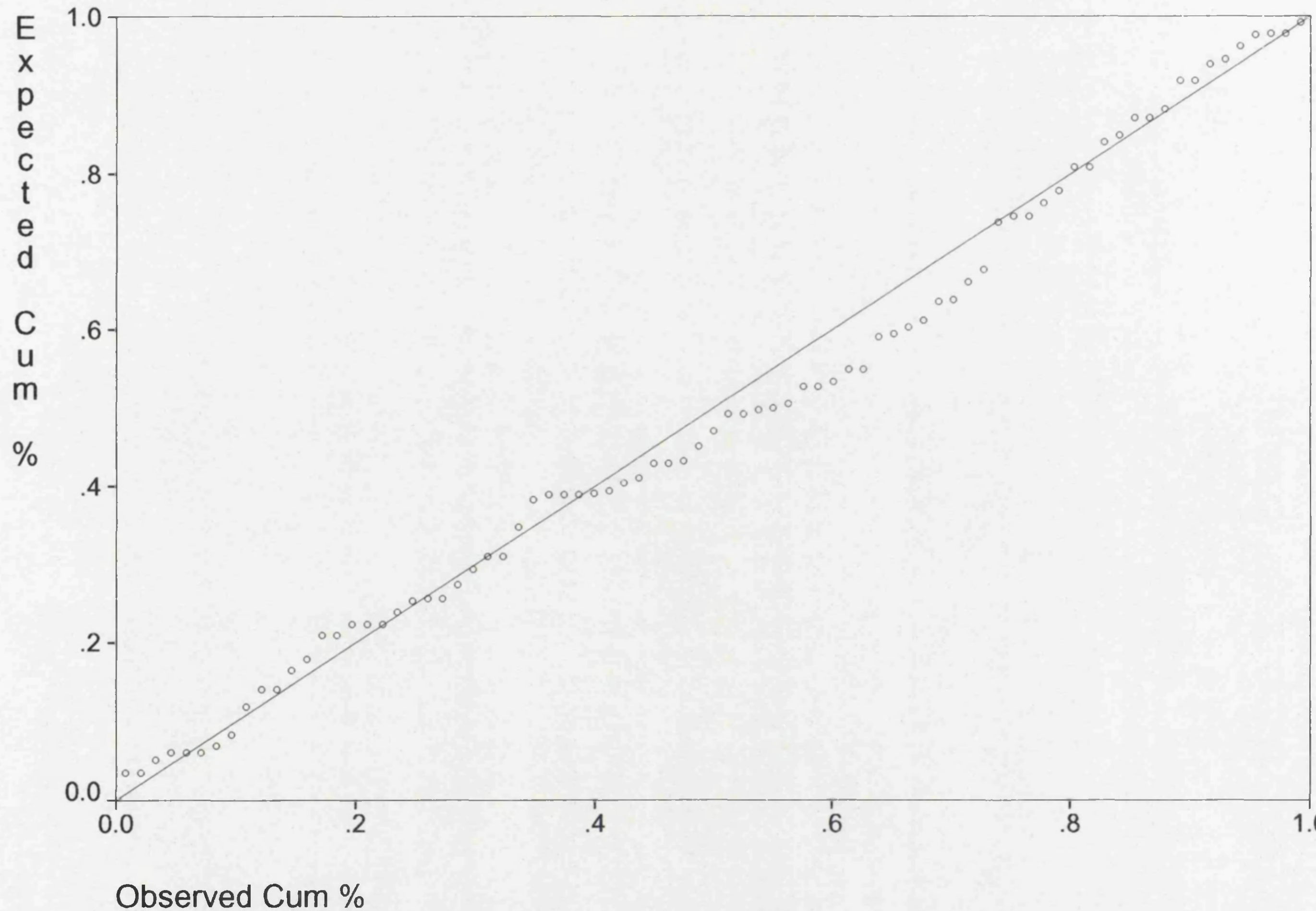


Figure 26

Histogram: Residual distribution

Full fixed appliances: Post-treatment PAR

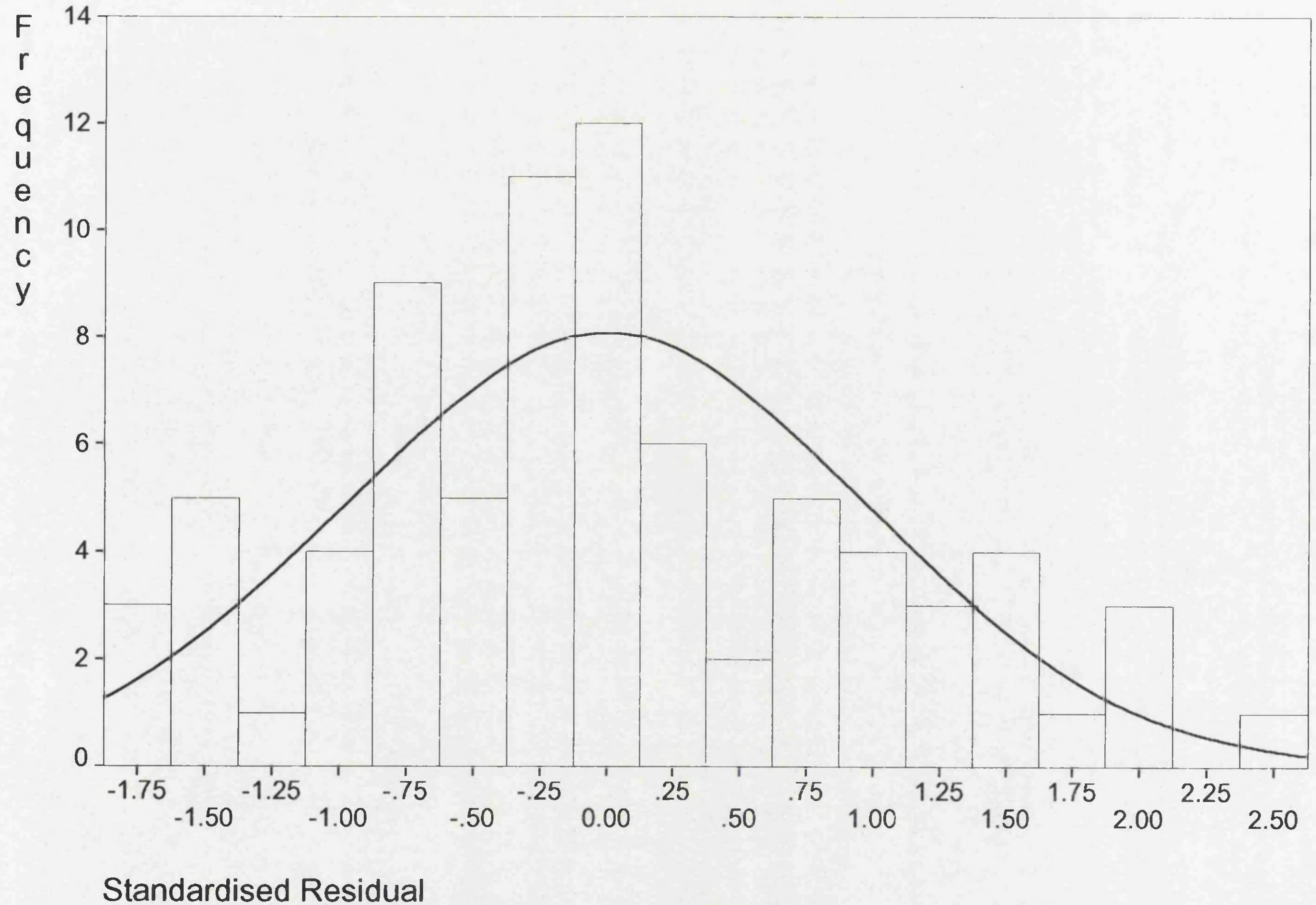


Figure 27

95% prediction intervals for post-treatment PAR score were calculated for various level of pre-treatment AC using pre-treatment AC and whether or not the treatment plan had been altered as the predictors (Table 41).

Pre-treatment AC		5	7	10
T/P altered				
	Mean	11.6 - 16.4	12.9 - 17.5	14.4 - 19.7
	Individual	6.0 - 21.9	7.3 - 23.1	9.0 - 25.0
T/P maintained				
	Mean	6.4 - 9.0	7.8 - 9.9	8.8 - 12.5
	Individual	0.0 - 15.4	1.2 - 16.5	2.9 - 18.5

Table 41

A further regression procedure was performed with pre-treatment Aesthetic Component excluded as a possible explanatory variable. The equation thus obtained was:

POST-TREATMENT PAR =
4.83 + 6.34(T/P ALTERED) + 4.32(FIRST MOLAR XTN) + (-3.29(ANTERIOR XB)) + 0.13(PRE-TREATMENT PAR) + ε

R^2 for this model was 0.40 (R_a^2=0.37) with the following beta values:

Variable	Beta
ALTERED	.48
FIRST MOLAR EXTRACTED	.27
PRE-TREATMENT PAR	.24
ANTERIOR XB.	-.23

Corresponding 95% prediction intervals were (Table 42):

Pre-treatment PAR		22	35	44
T/P altered				
	Mean	11.6 - 16.5	13.4 - 17.0	14.3 - 19.5
	Individual	6.1 - 22.0	7.8 - 23.6	8.8 - 25.0
T/P maintained				
	Mean	6.4 - 9.0	8.2 - 10.6	8.8 - 12.3
	Individual	0.0 - 15.4	1.7 - 17.1	2.7 - 18.3

Table 42

95% prediction intervals at different levels of pre-treatment PAR with whether or not a first permanent molar had been extracted as an explanatory variable were (Table 43):

Pre-treatment PAR	21	29	43
T/P maintained, no 1st molar xtn			
Mean	6.2 - 8.9	7.5 - 9.7	8.7 - 12.1
Individual	0.0 - 15.3	0.9 - 16.3	2.6 - 18.2
T/P maintained, + 1st molar xtn			
Mean	9.1 - 14.7	10.5 - 15.6	8.0 - 14.9
Individual	3.8 - 20.0	4.9 - 21.0	3.1 - 19.8

Table 43

The assumptions of normality and equal variance were acceptable (Figures 28-29).

Normal(P-P) plot of standardised residual

Full fixed appliances: Post-treatment PAR

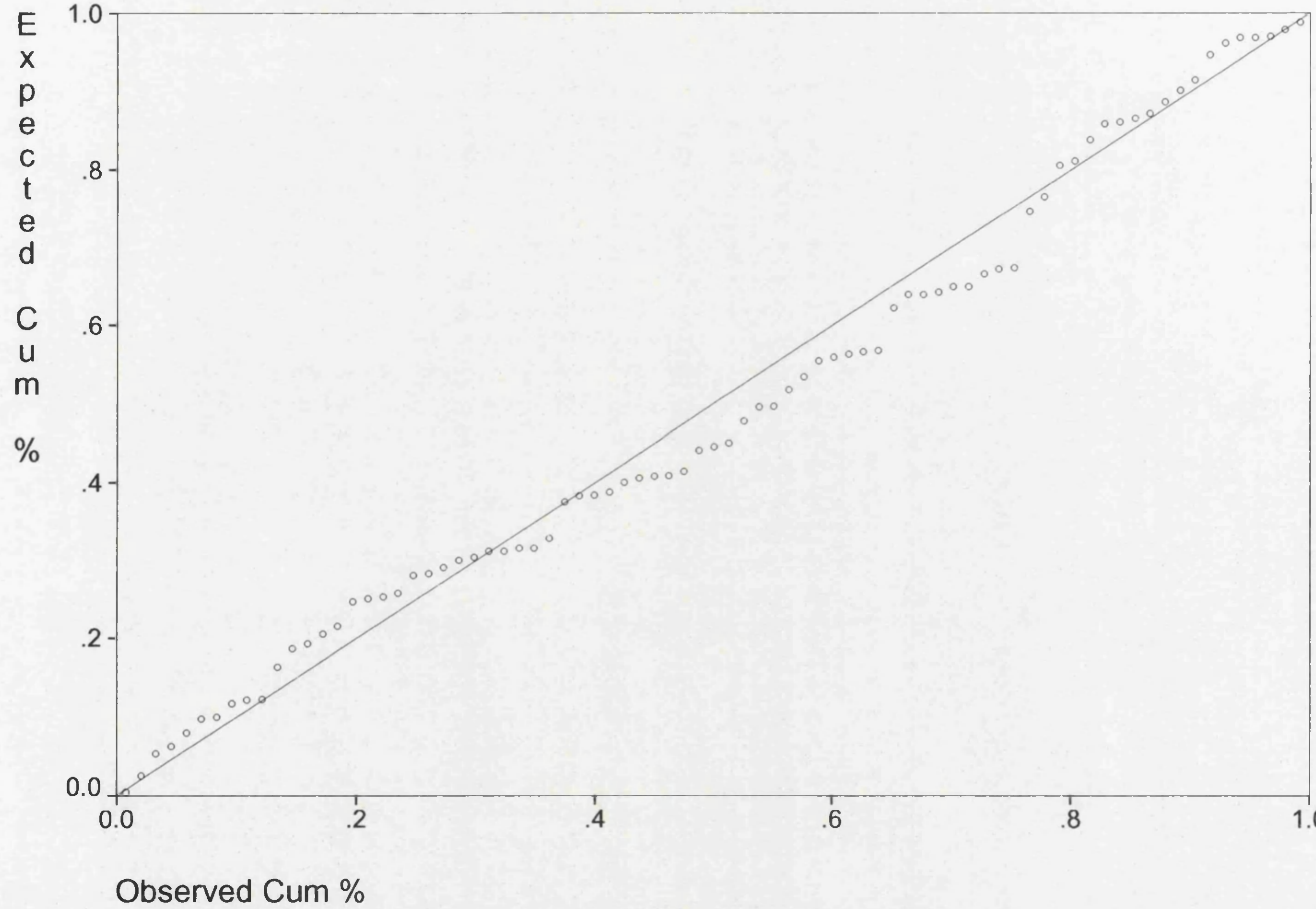


Figure 28

Histogram: Residual distribution

Full fixed appliances: Post-treatment PAR

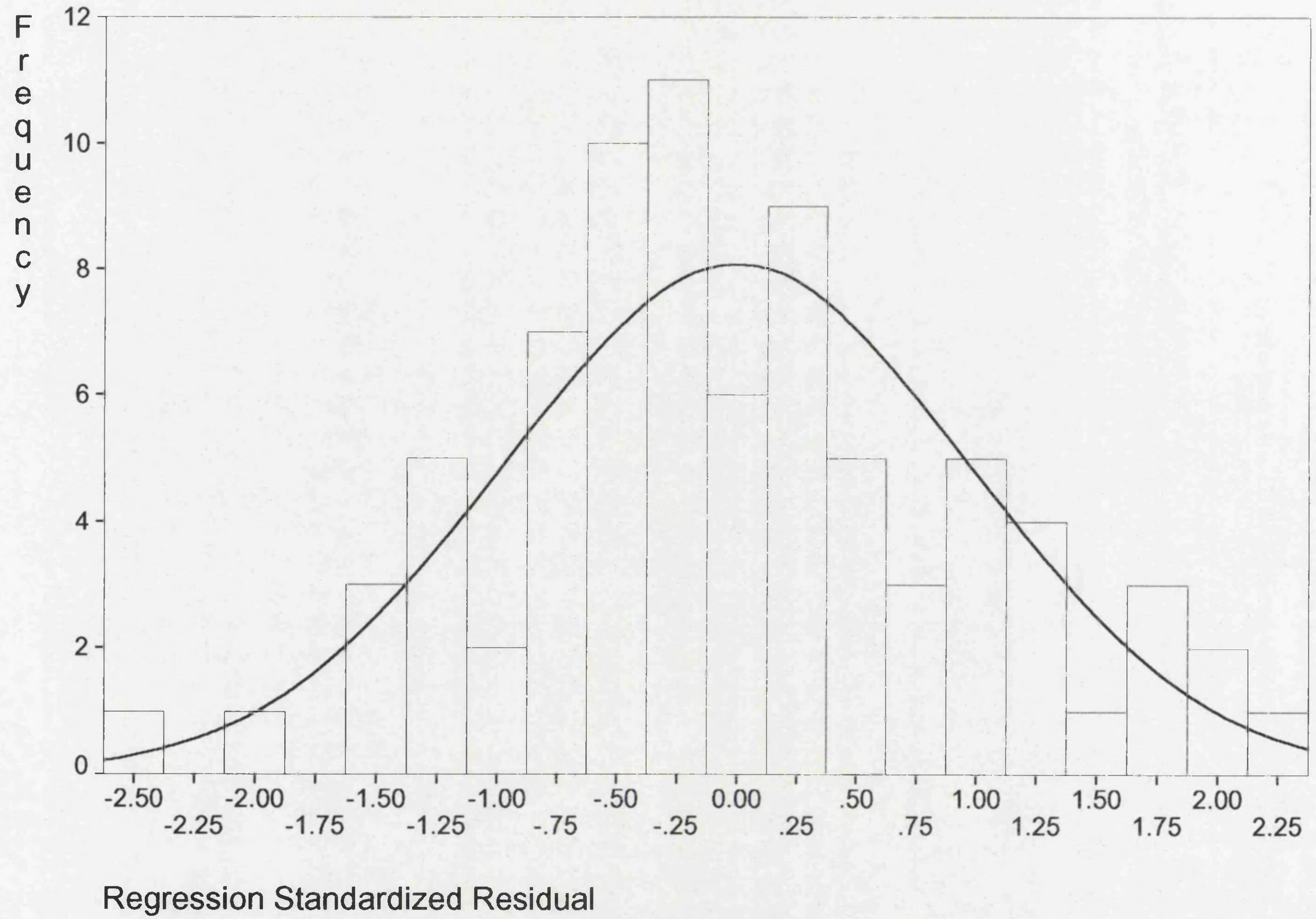


Figure 29

b) Pre-treatment information (Post-treatment PAR/Full fixed)

The variables entered into the equation were:

Pre-treatment Aesthetic Component

Whether or not a first permanent molar had been extracted

Whether or not an anterior crossbite was present

Whether or not oral hygiene was regarded as good

R^2 for this model was low at 0.23 (R_a^2=0.19). The assumptions of normality and equal variance were acceptable (Figures 30-31). The Lilliefors statistic for the residuals was 0.09 for which the associated significance level was 0.19 suggesting a distribution not significantly different from normal.

The constructed regression equation was:

POST-TREATMENT PAR =
6.11 + 0.6(AESTHETIC COMPONENT) +
4.35(FIRST MOLAR XTN) + (-2.64(GOOD ORAL HYGIENE)) +
(-2.62(ANTERIOR XB)) + ε

The beta values were:

Variable	Beta
AESTHETIC COMPONENT	.24
FIRST MOLAR EXTRACTION	.27
GOOD ORAL HYGIENE	-.24
ANTERIOR CROSSBITE	-.22

A subsequent regression carried out excluding Aesthetic Component as a possible predictor variable did not introduce any further different into the equation. It was noted however that of those variables not in the equation the next most influential variable, if it had been entered, would have been the pre-treatment PAR score judging by its "Beta in" value of 0.21.

Normal (P-P) plot of standardised residual

Full fixed appliances: Post-treatment PAR

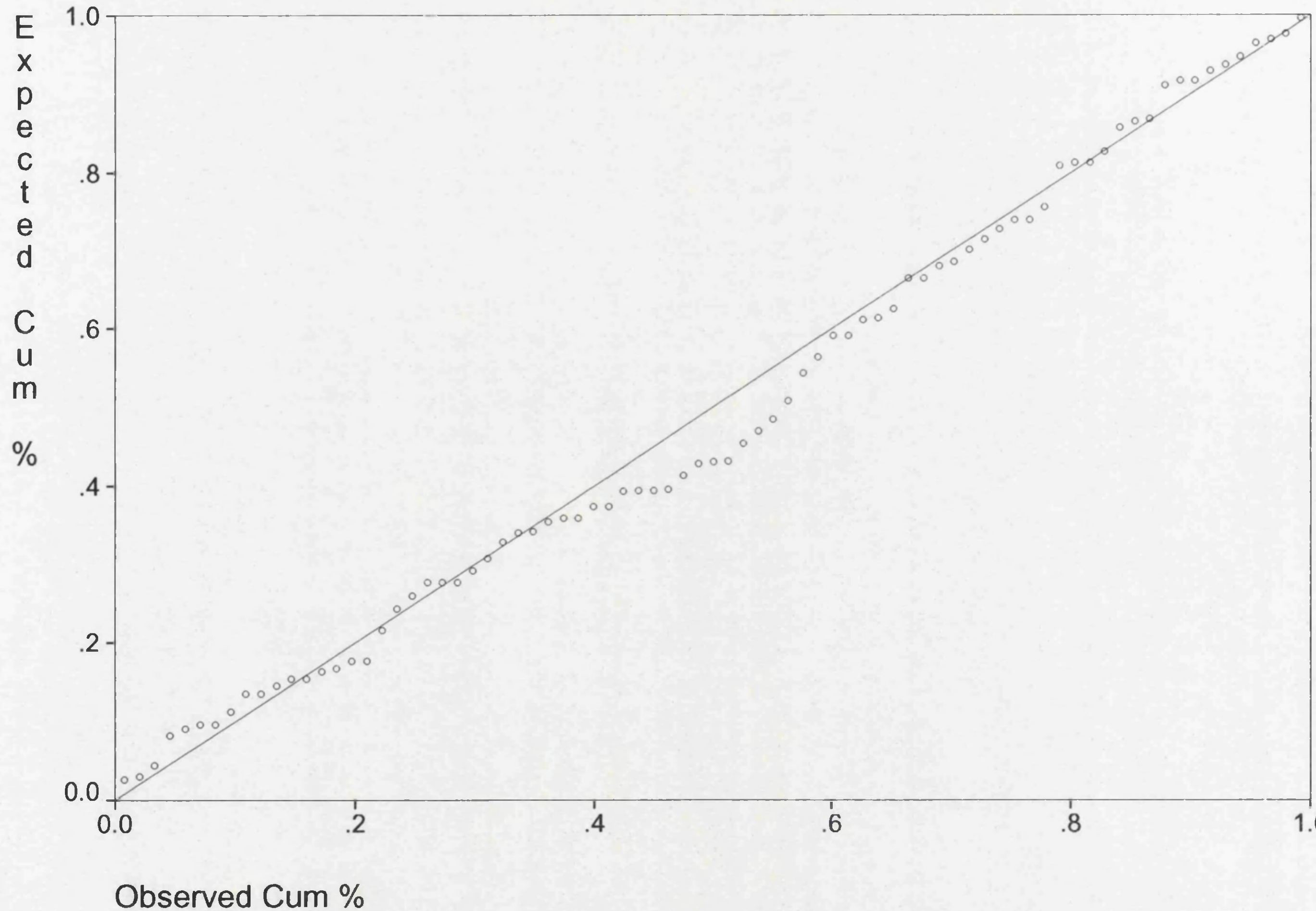


Figure 30

Histogram: Residual distribution

Full fixed appliances: Post-treatment PAR

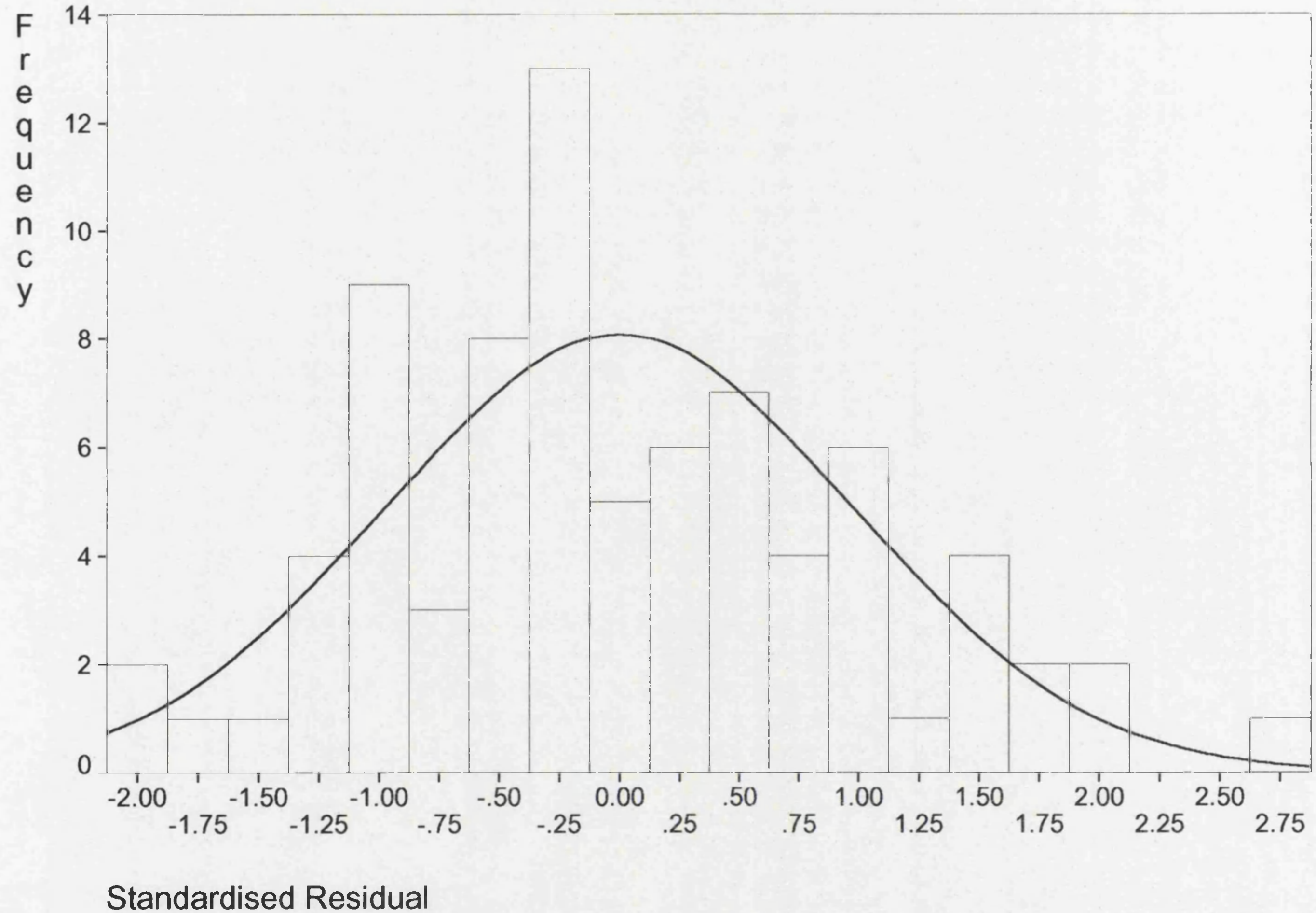


Figure 31

11.1.2 Post-treatment PAR: Other appliances

a) All information (Post-treatment PAR/Other)

Variables entered into the equation were:

Whether or not the operator was an undergraduate

Whether or not the treatment plan had been altered

Whether or not the initial incisor relationship was Class II division 1

Whether or not the incisor relationship was Class II division 2

Whether or not headgear had been worn

Number of appointments during treatment

R^2 for this model was 0.41 (R_a^2=0.35) with acceptable assumptions of regression as illustrated by Figure 32 and by a Lilliefors statistic of 0.69 indicating no difference from a normal distribution (P>0.20).

Normal (P-P) plot of standardised residual

Other appliances: Post-treatment PAR

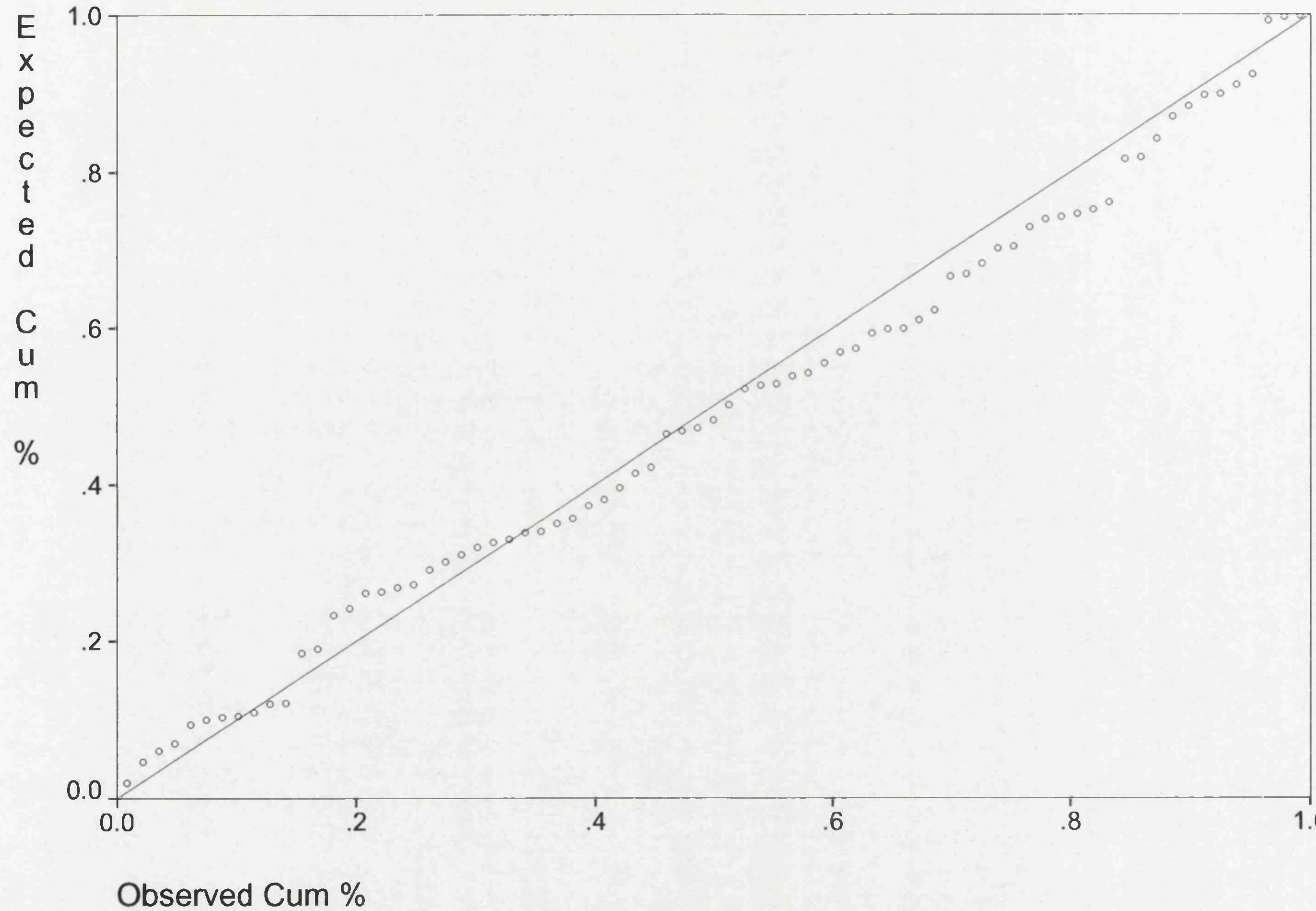


Figure 32

The constructed regression equation was:

POST-TREATMENT PAR =

$$12.29 + 3.89(\text{UNDERGRADUATE}) + 5.31(\text{T/P ALTERED}) + 8.50(\text{CLASS II/2}) + 4.60(\text{CLASS II/1}) + (-3.36(\text{HEADGEAR})) + (-0.24(\text{No. OF APPOINTMENTS})) + \varepsilon$$

with the following beta values:

Variable	Beta
UNDERGRADUATE	0.30
T/P ALTERED	0.35
CLASS II/2	0.42
CLASS II/1	0.36
HEADGEAR	-0.23
No. OF APPOINTMENTS	-0.23

Mean and individual 95% prediction intervals for post-treatment PAR with number of appointments and whether or not the treatment plan was altered are shown in Table 44.

No. of appts.		6	10	15
T/P altered				
	Mean	13.2 - 19.1	12.5 - 17.9	11.2 - 16.8
	Individual	6.2 - 26.1	5.3 - 25.1	4.1 - 23.9
T/P maintained				
	Mean	8.9 - 12.8	8.2 - 11.6	6.8 - 10.6
	Individual	1.2 - 20.6	0.2 - 19.5	0.0 - 18.4

Table 44

b) Pre-treatment information (Post-treatment PAR/Other)

The variables included in the equation were:

Whether or not the operator was an undergraduate

Whether the pre-treatment incisor relationship was class II division 2

That the equation contained only the constant and two binomial dummy variables is borne out by the appearance of the plot of residuals against predicted values (Figure 33) which illustrates the fact that the predicted value of post-treatment PAR could be one of only four possible values.

The resulting model had an R^2 of only 0.17 (R_a^2=0.14), indicating a poor fit of the data to the model and that only 17% of the total variability in the dependent variable had been explained by it.

The assumptions of regression were not acceptable for this model. A Lilliefors' statistic of 0.10 suggested the residual distribution was significantly different from normal (P=0.036).

Residuals/Predicted values

Other appliances: Post-treatment PAR

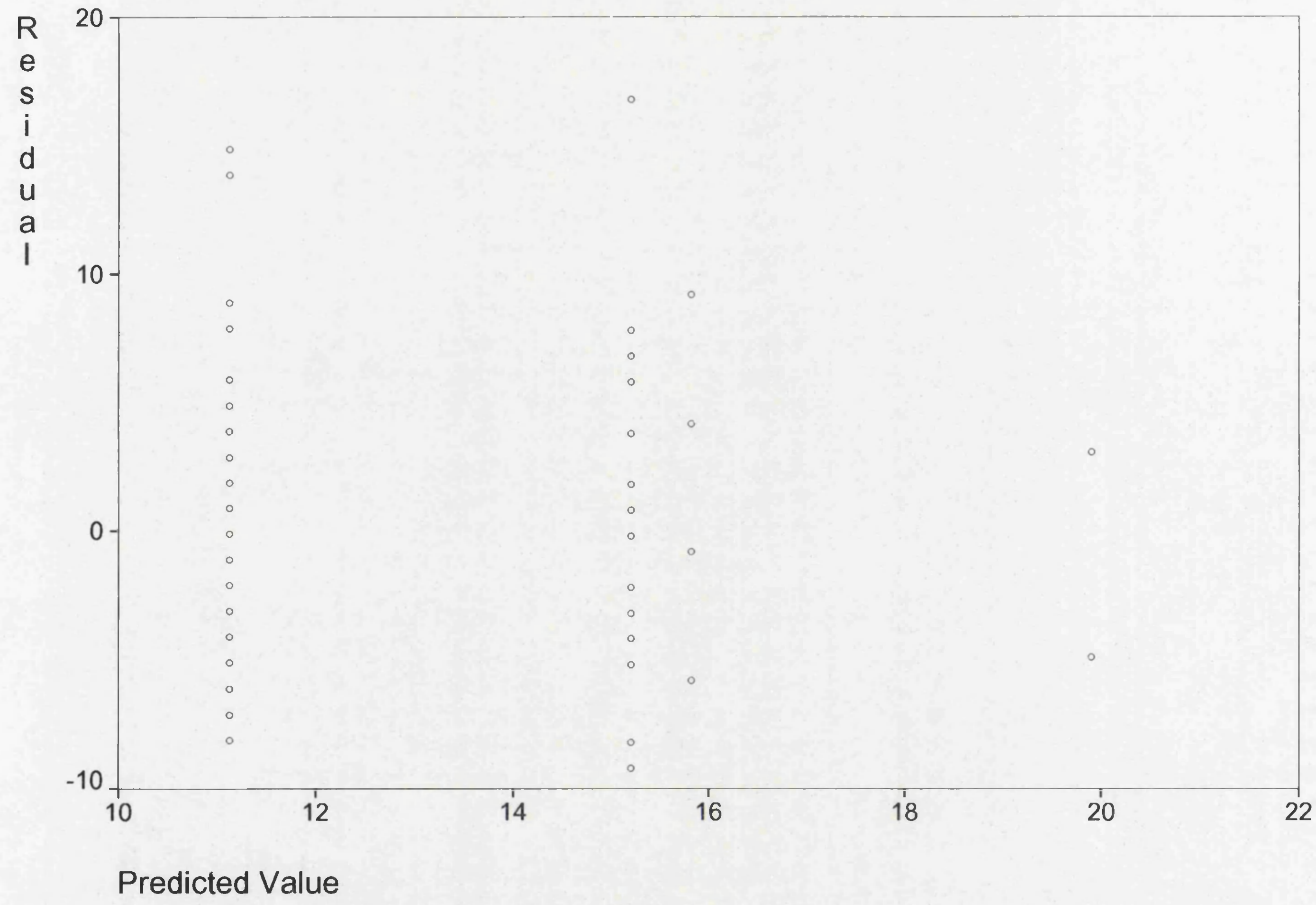


Figure 33

11.2 Change in PAR score

11.2.1 Full fixed appliances

a) All information (Change in PAR/Full fixed)

Regression on the difference between pre-treatment and post-treatment total PAR scores produced a model which included the following predictor variables:

- Pre-treatment PAR score
- Whether the original treatment plan was altered
- Whether an anterior crossbite was present initially
- Whether treatment had involved extraction of a first permanent molar

Much of the variability in change in PAR appeared to be explained by this model (R^2=0.84, R_a^2=0.83).

The regression equation was:

CHANGE IN PAR =
-4.84 + 0.87(PRE-TREATMENT PAR) +
(-6.43(T/P ALTERED)) + 3.29(ANTERIOR CROSSBITE) +
(-4.32(FIRST MOLAR XTN)) + ε

The beta values for the variables in the equation were:

Variable	Beta
PRE-TREATMENT PAR	0.83
T/P ALTERED	-0.25
ANTERIOR CROSSBITE	0.14
FIRST MOLAR EXTRACTED	-0.14

A scatterplot of residuals against predicted values showed the plotted points to be confined to a broad horizontal band except there may have been a slight tendency for the spread of the residuals to become greater with increasing magnitude of the predicted value (Figure 34). The distribution of the residuals appeared to be reasonably normal (Figure 35, Lilliefors test $P>0.20$) and the points for detrended normal plot (Figure 36) did not produce any definite pattern. The normal plot of residuals (Figure 37) however did deviate from a straight line to some degree. It was considered that the assumption of equal variance in this case might be at the borderline of acceptability.

In an attempt to stabilise this variance, regression was carried out on the difference between the natural logs of pre- and post-treatment PAR scores, with the natural log of the pre-treatment PAR total included as an independent variable.

It was noted that the same predictor variables, in transformed guise as appropriate, were entered into this equation as in the previous regression.

R^2 for the transformed data was 0.37 (R_a^2=0.33) however the normal plot of the residuals was not markedly improved (Figure 38).

Scatterplot: Residuals/Predicted values

Full fixed appliances: Change in PAR

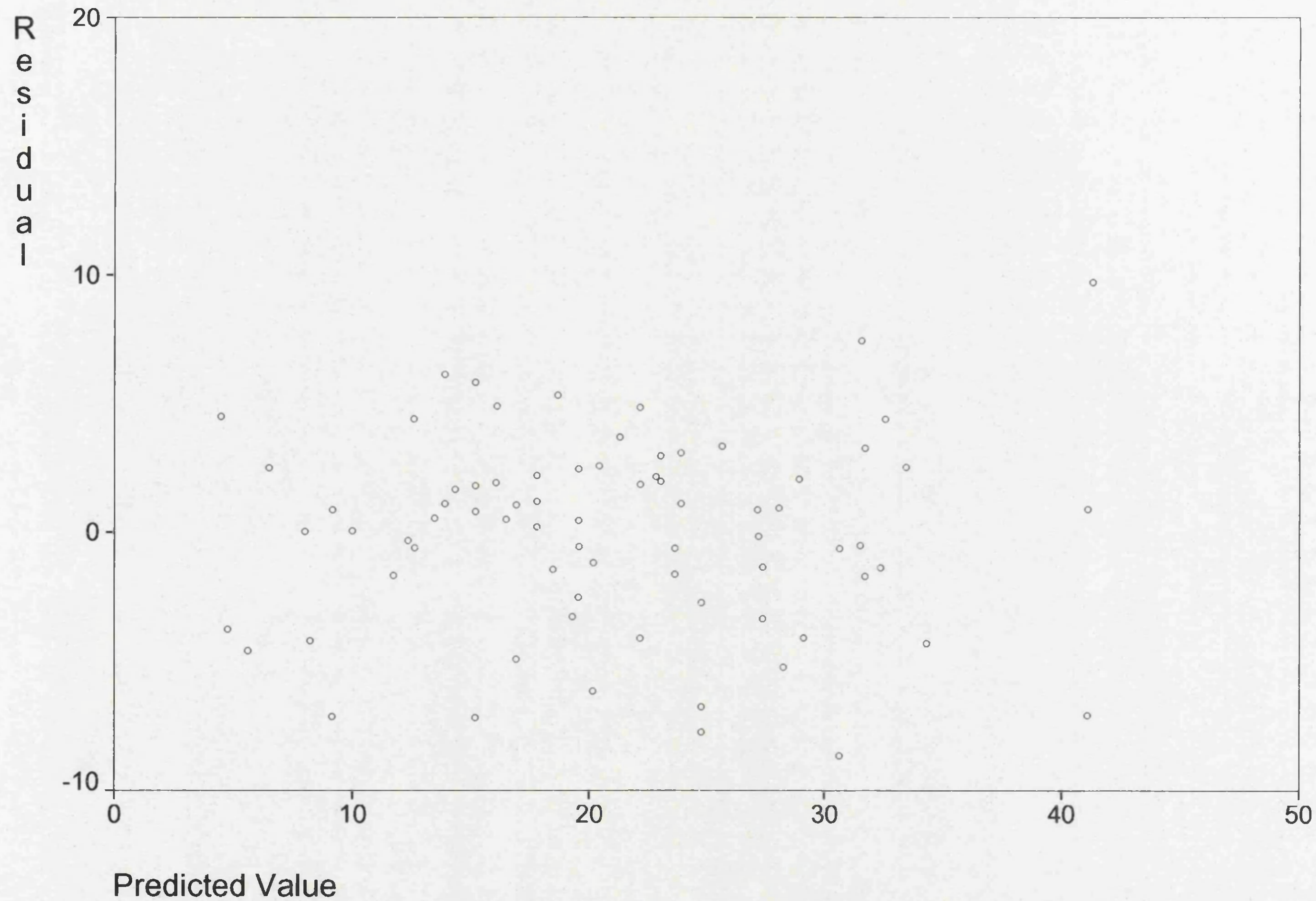


Figure 34

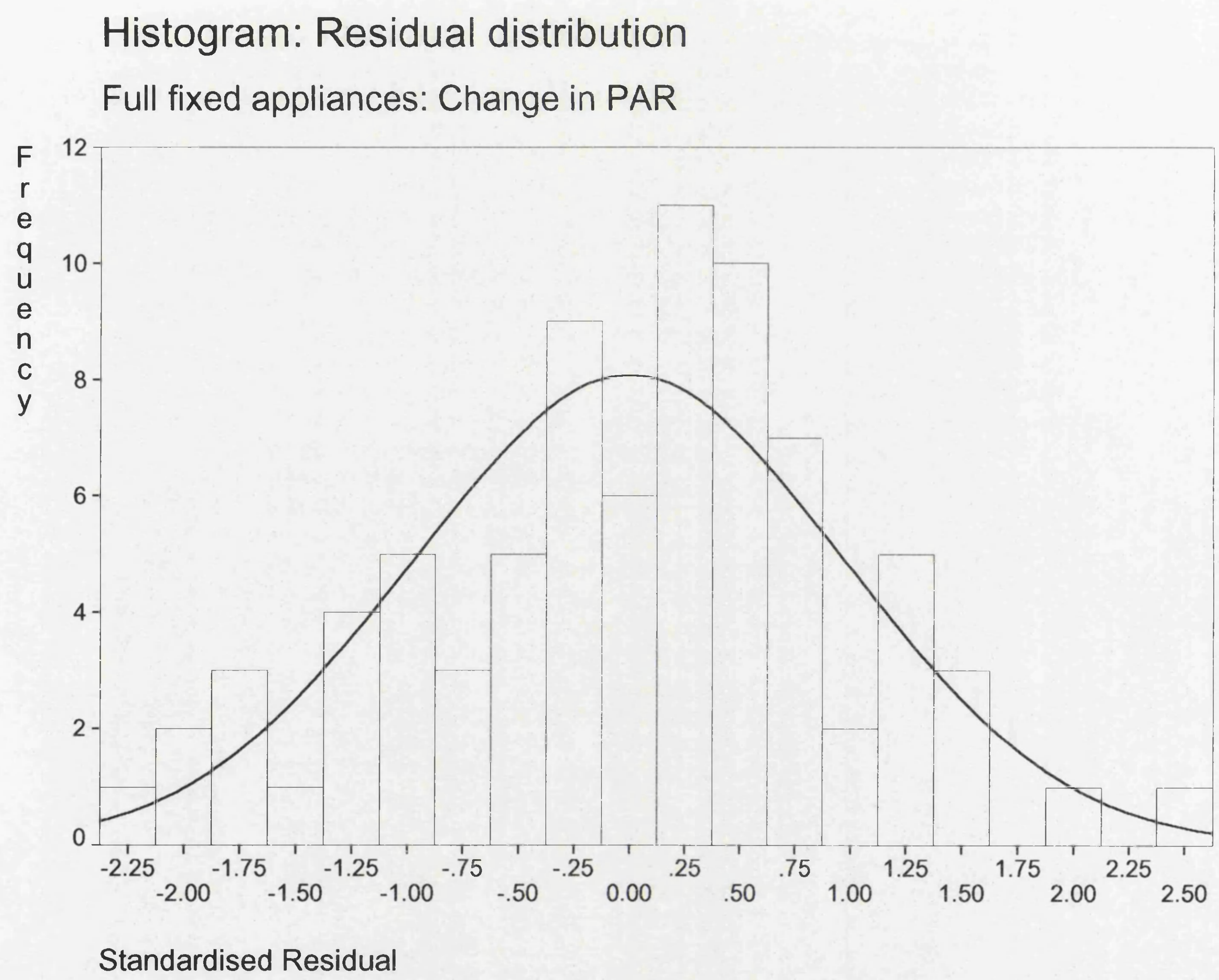

Histogram: Residual distribution
Full fixed appliances: Change in PAR
Frequency
12
10
8
6
4
2
0
-2.25
-2.00
-1.75
-1.50
-1.25
-1.00
-.75
-.50
-.25
0.00
.25
.50
.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Standardised Residual

Figure 35

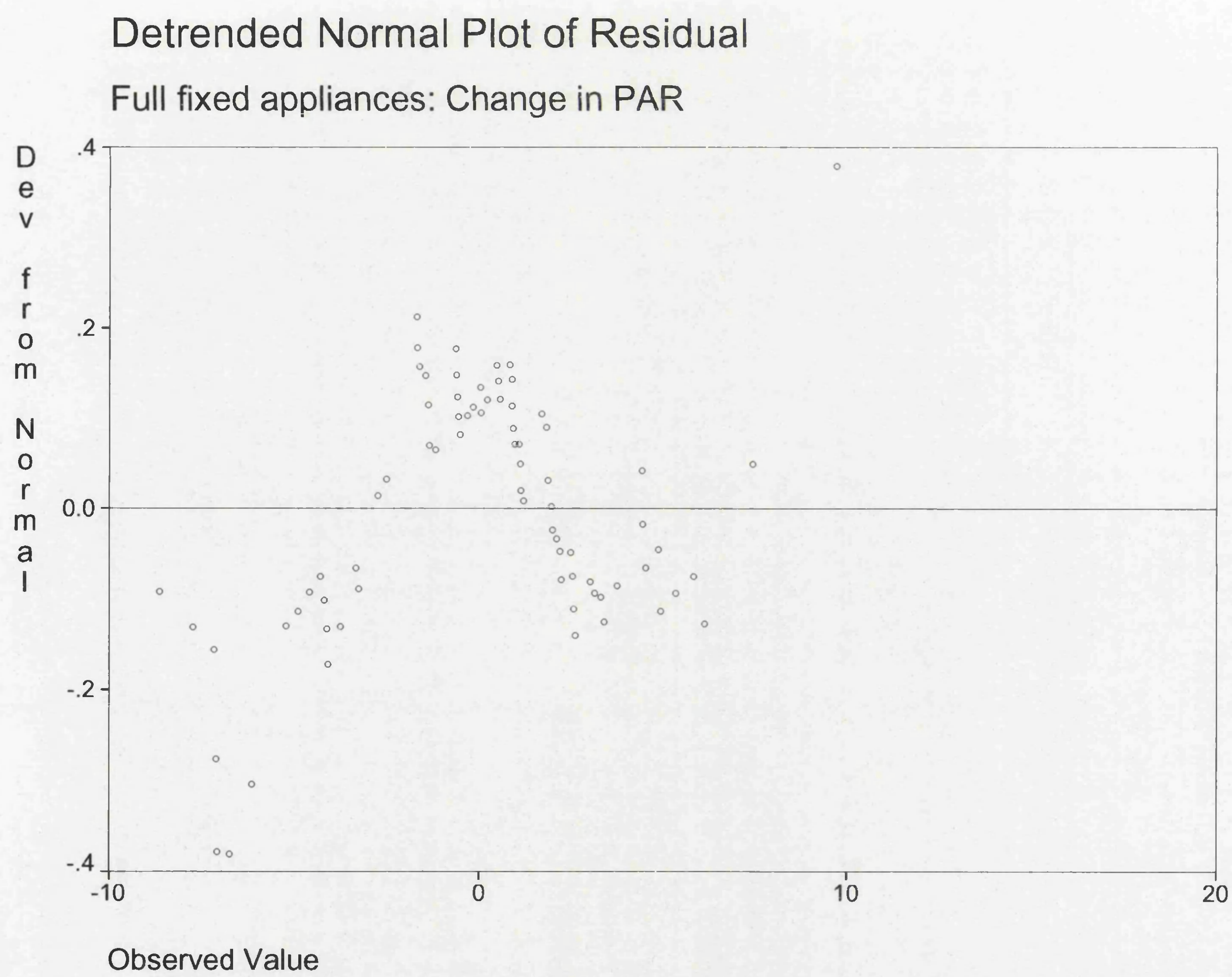

Detrended Normal Plot of Residual
Full fixed appliances: Change in PAR
Dev from Normal
.4
.2
0.0
-.2
-.4
-10
0
10
20
Observed Value

Figure 36

Normal Plot of Residual

Full fixed appliances: Change in PAR

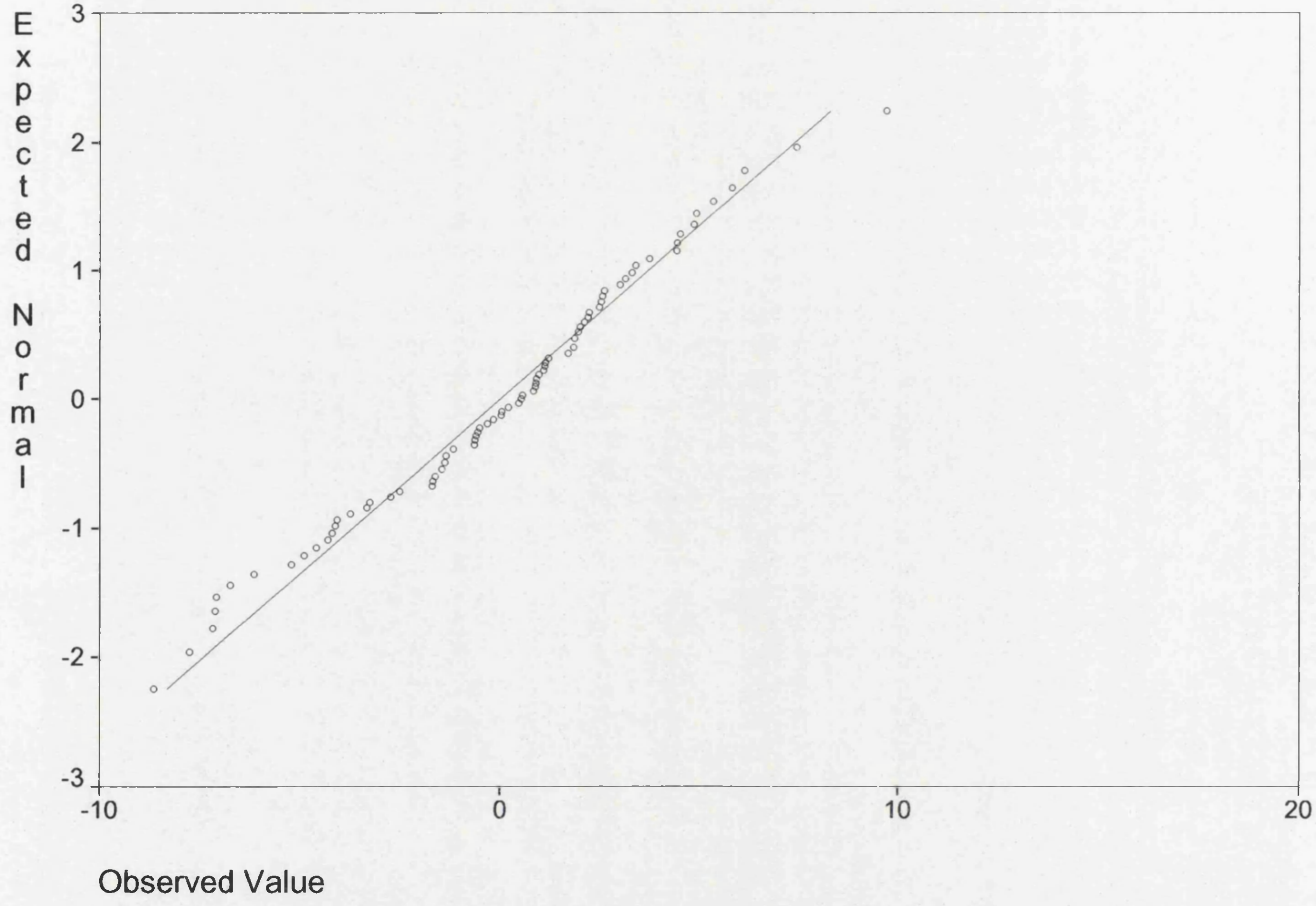


Figure 37

Normal Plot of Residual

Full fixed appliances: Ln(change in PAR)

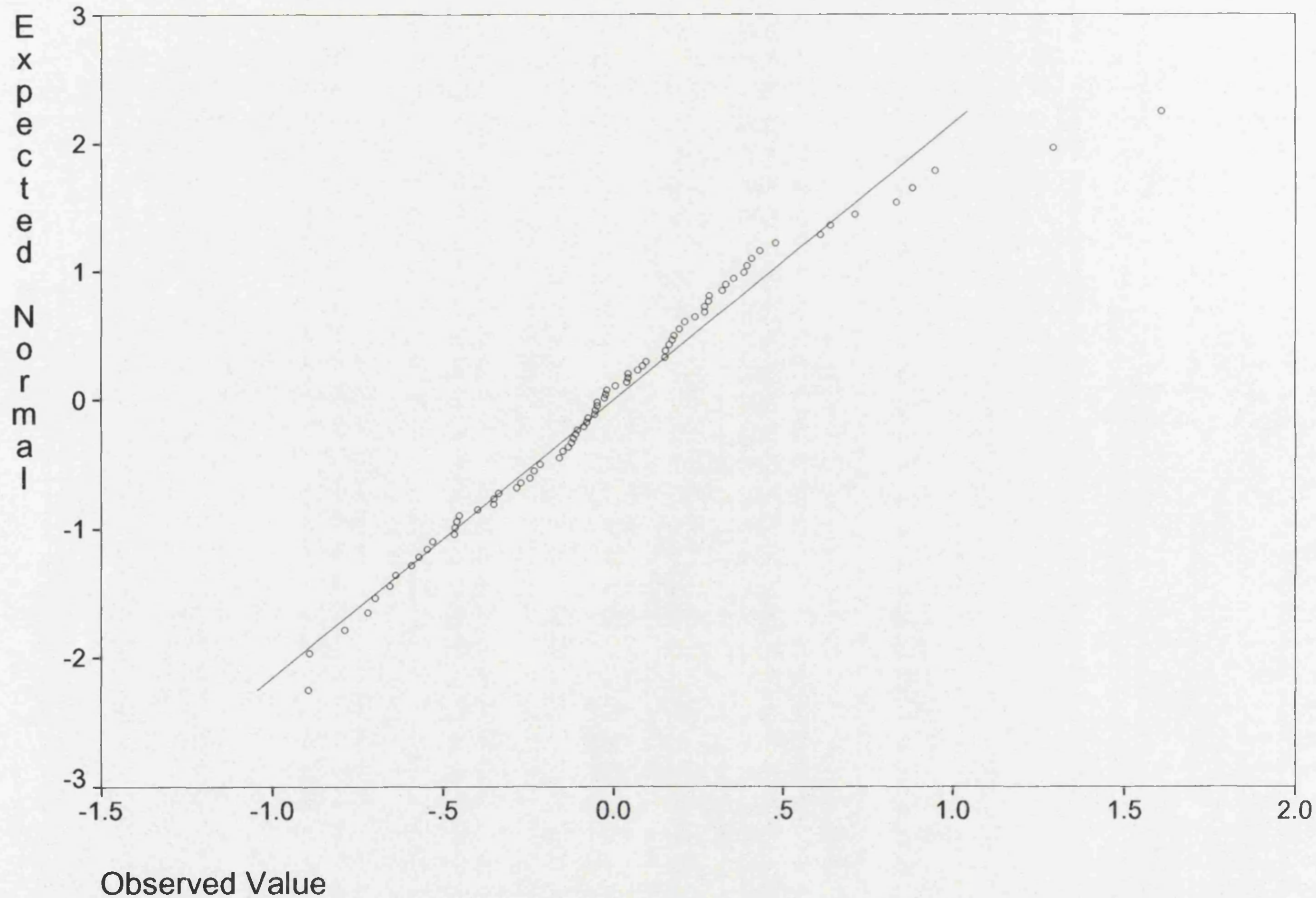


Figure 38

The prediction intervals for change in PAR given various levels of pre-treatment PAR and using the other predictor variables from the regression equation are given below (Table 45).

Pre-treatment PAR	22	35	44
T/P altered, no crossbite			
Mean	5.5 - 10.4	17.0 - 21.6	24.5 - 29.7
Individual	0.0 - 15.9	11.3 - 27.2	19.1 - 35.1
T/P maintained, no crossbite			
Mean	13.0 - 15.6	24.4 - 26.8	31.7 - 35.2
Individual	6.6 - 22.0	17.9 - 33.3	25.7 - 41.2

Pre-treatment PAR	23	35	39
T/P maintained, no crossbite			
Mean	13.9 - 16.4	24.4 - 26.8	27.7 - 30.5
Individual	7.5 - 22.9	17.9 - 33.3	21.4 - 36.8
T/P maintained, with crossbite			
Mean	16.2 - 20.7	26.9 - 30.9	30.4 - 34.4
Individual	10.6 - 26.4	21.1 - 36.7	24.5 - 40.2

Pre-treatment PAR	21	29	39
T/P maintained, no 1st molar xtn			
Mean	12.1 - 14.8	19.3 - 21.5	22.8 - 25.0
Individual	5.7 - 21.1	12.7 - 28.1	16.2 - 31.5
T/P maintained, + 1st molar xtn			
Mean	6.3 - 11.9	13.4 - 18.8	16.8 - 22.3
Individual	1.0 - 17.2	8.0 - 24.1	11.5 - 27.6

Table 45

b) Pre-treatment information (Change in PAR/Full fixed)

The included variables were:

Pre-treatment PAR score

Whether or not an anterior crossbite was present

Whether or not a first permanent molar was extracted as part of the treatment plan

The constructed regression equation was as follows:

CHANGE IN PAR =

$$-5.41 + 0.86(\text{PRE-TREATMENT PAR}) + 3.06(\text{ANTERIOR XB}) + (-4.17(\text{FIRST MOLAR XTN})) + \varepsilon$$

with beta values,

Variable	Beta
PRE-TREATMENT PAR	0.82
ANTERIOR CROSSBITE	0.13
FIRST PERM.MOLAR EXTRACTED	-0.14

R^2 for this model was 0.78 (R_a^2=0.77). Again the assumptions of regression for this model were at the borderline of acceptability (Figures 39-41).

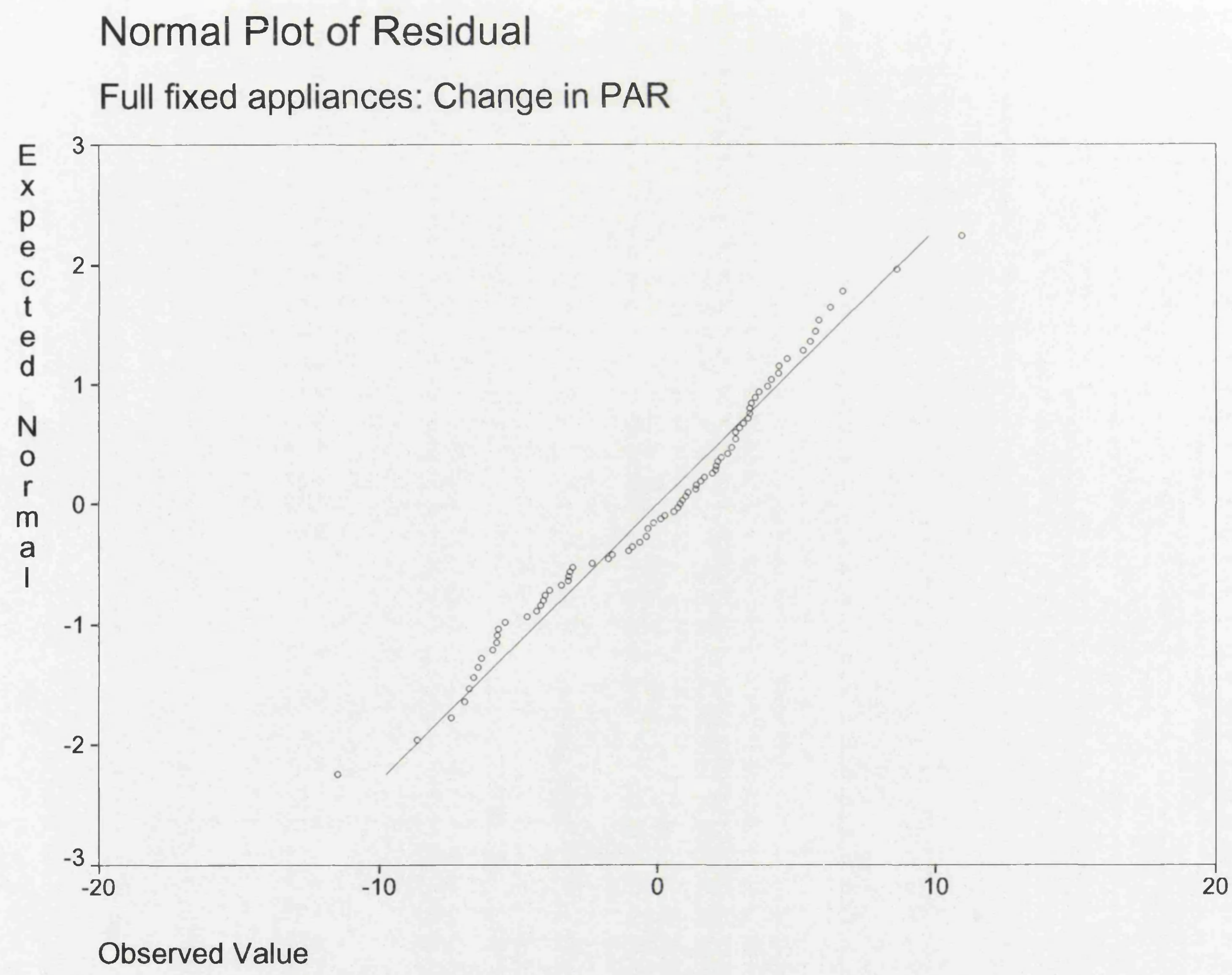

Normal Plot of Residual
Full fixed appliances: Change in PAR
Expected Normal
3
2
1
0
-1
-2
-3
-20
-10
0
10
20
Observed Value

Figure 39

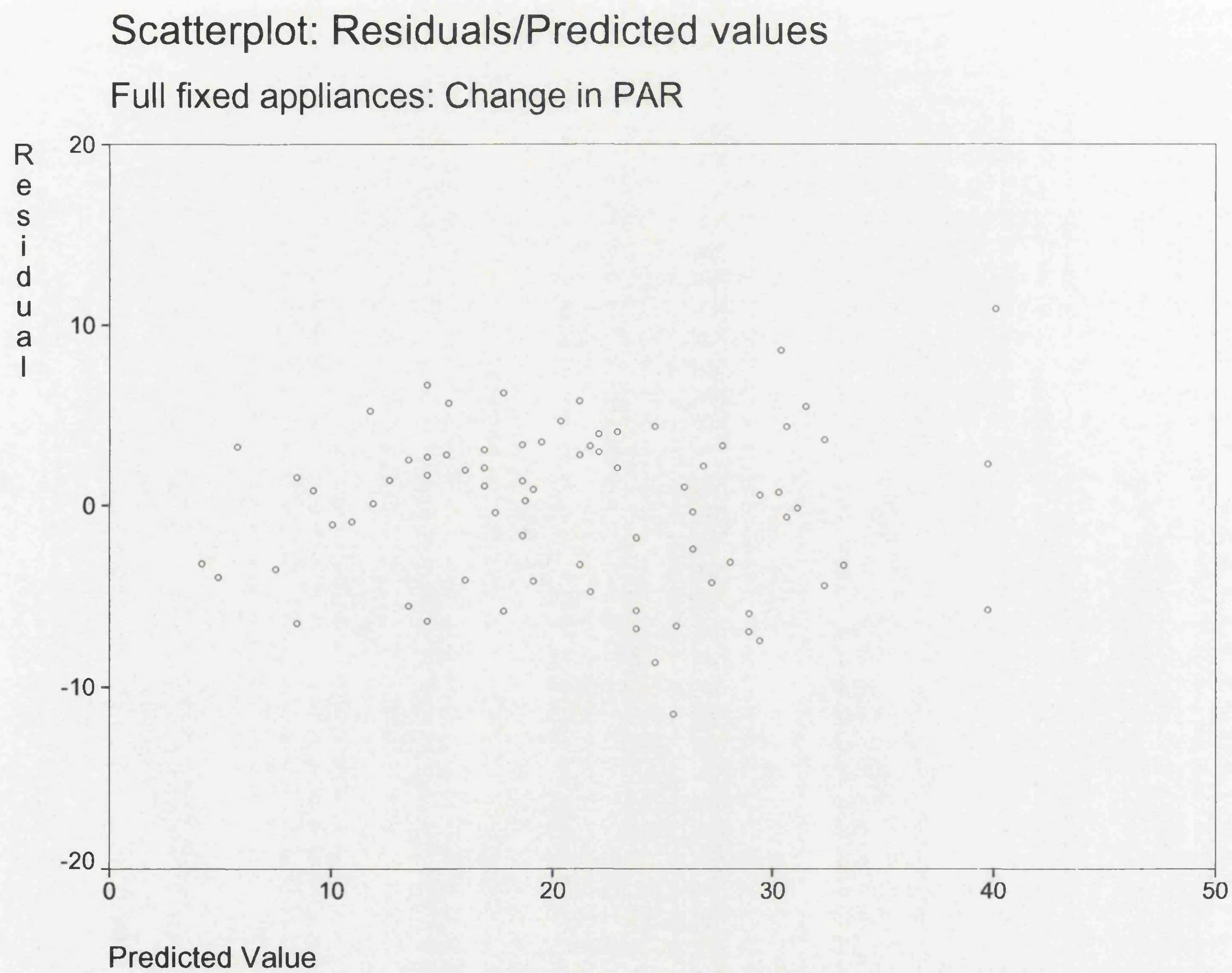
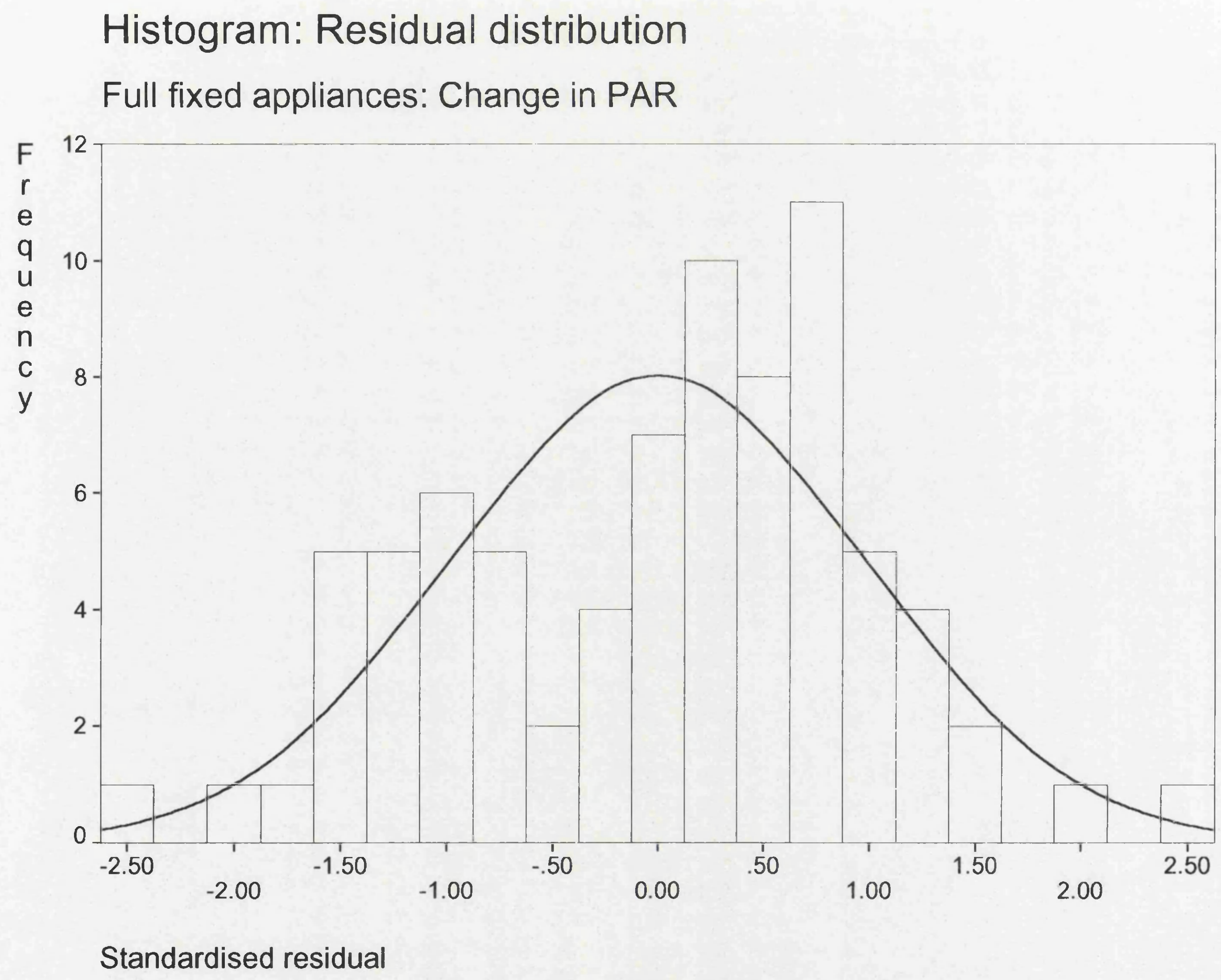

Scatterplot: Residuals/Predicted values
Full fixed appliances: Change in PAR
Residual
20
10
0
-10
-20
0
10
20
30
40
50
Predicted Value

Figure 40



Histogram: Residual distribution
Full fixed appliances: Change in PAR
Frequency
0
2
4
6
8
10
12
-2.50
-2.00
-1.50
-1.00
-.50
0.00
.50
1.00
1.50
2.00
2.50
Standardised residual

Figure 41

95% prediction intervals for change in PAR at various levels of the predictor variables is given in Table 46.

Pre-treatment PAR	21	29	33
No 1st molar xtn			
Mean	11.0 - 14.2	18.3 - 20.7	21.7 - 24.1
Individual	3.6 - 21.6	10.6 - 28.4	14.0 - 31.8
1st molar xtn			
Mean	5.2 - 11.7	12.2 - 18.5	15.5 - 21.9
Individual	0.0 - 17.9	5.9 - 24 7	9.3 - 28.1

Pre-treatment PAR	23	33	39
No crossbite			
Mean	12.9 - 15.7	21.7 - 24.1	26.5 - 29.6
Individual	5.4 - 23.3	14.0 - 31.8	19.1 - 37.0
Crossbite			
Mean	14.8 - 20.0	23.7 - 28.2	28.8 - 33.4
Individual	8.2 - 26.6	16.9 - 35.1	22.0 - 40.3

Table 46

11.2.2 Other appliances

a) All information (Change in PAR/Other)

Regressing on change in PAR produced an equation which included the following variables:

Pre-treatment PAR score

Whether or not the operator had been an undergraduate student

Whether or not the treatment plan had been altered

The number of appointments during active treatment

Whether or not the incisor relationship was Class II division 1

Whether or not the incisor relationship was Class II division 2

R^2 for this model was high at 0.79 (R_a^2=0.77). The regression equation and corresponding beta values in this case were:

CHANGE IN PAR =
-10.33 + 0.91(PRE-TREATMENT PAR) +
(-3.92(UNDERGRAD.)) + (-5.29(T/P ALTERED)) +
(-7.34(CLASS II/2)) + 0.28(No. APPOINTMENTS)
+ (-3.12(CLASS II/1)) + ε

Variable	Beta
PRE-TREATMENT PAR	0.87
UNDERGRADUATE	-0.18
T/P ALTERED	-0.20
CLASS II/2	-0.21
No APPOINTMENTS	0.15
CLASS II/1	-0.14

Plots for examination of the residuals are shown in Figures 42-43. Again the regression assumptions might be thought to be at the boundary of acceptability, however the residual distribution does not differ significantly from normal (Lilliefors P>0.20).

Normal(P-P) plot of standardised residual

Other appliances: Change in PAR

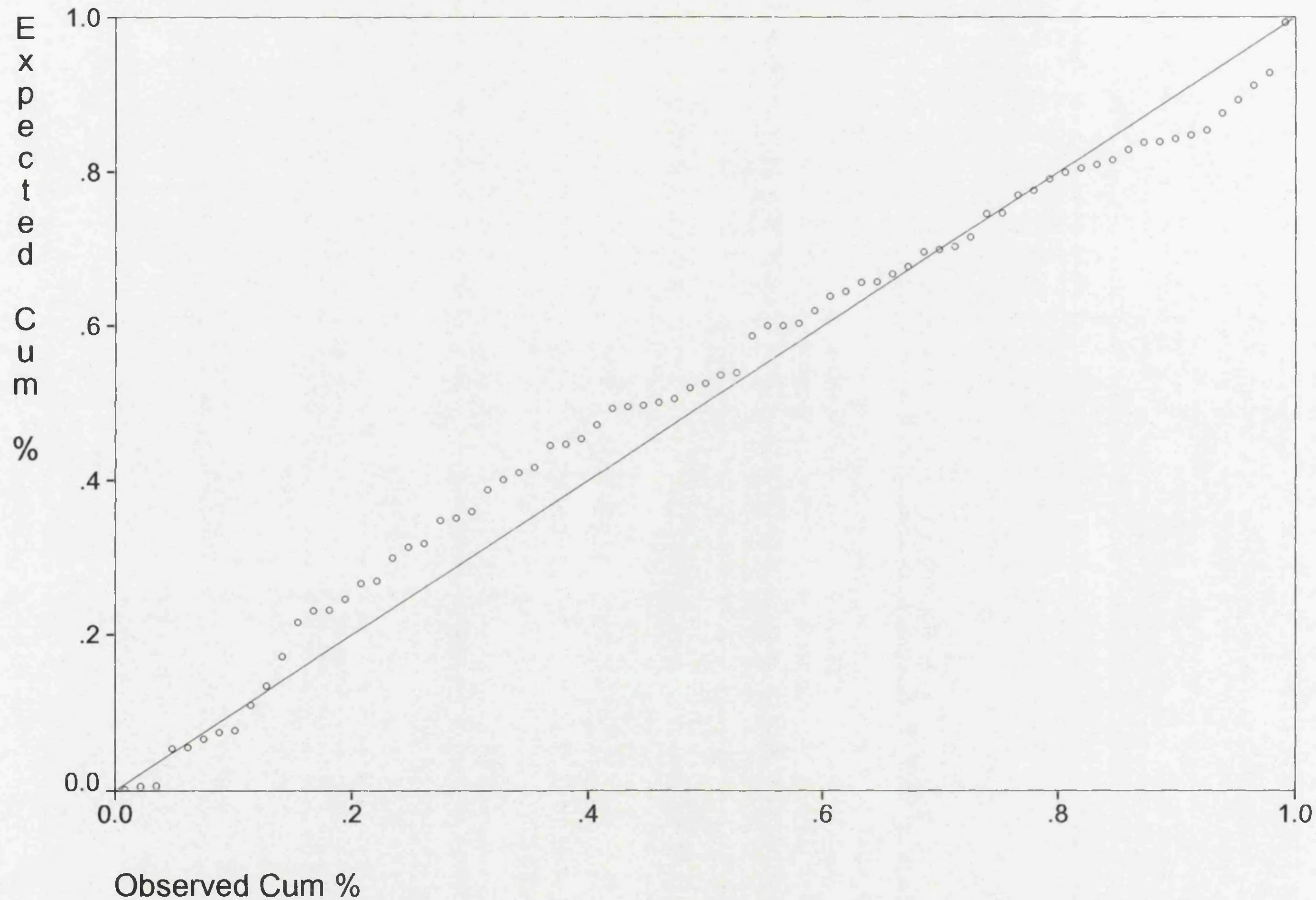


Figure 42

Scatterplot: Residual/Predicted value

Other appliances: Change in PAR

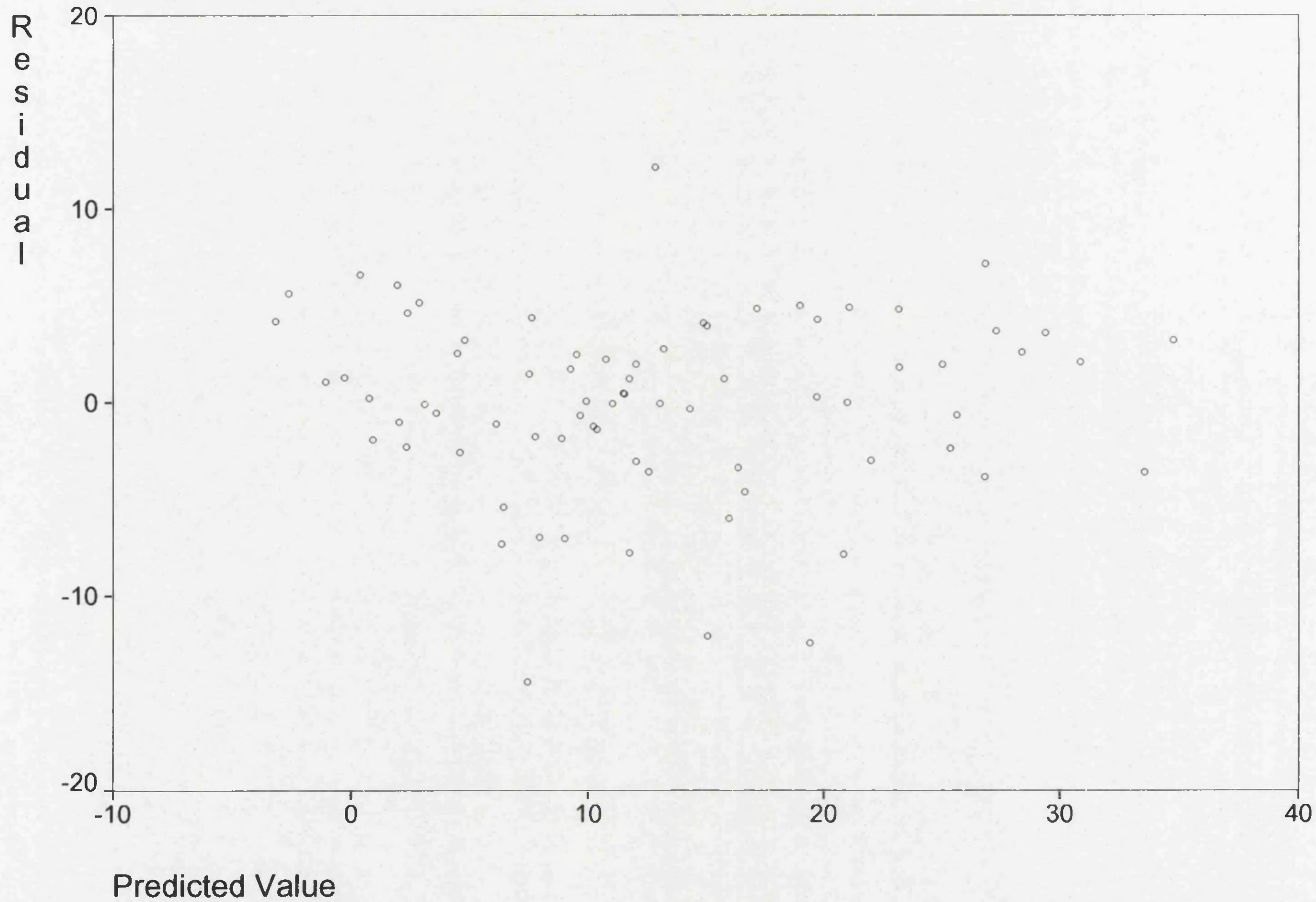


Figure 43

b) Pre-treatment information (Change in PAR/Other)

The following variables were included in the equation:

Pre-treatment PAR

Whether or not the operator was a postgraduate student

Whether or not an anterior crossbite was present initially

R^2 for this model was 0.72 (R_a^2=0.71). The regression equation and corresponding beta values in this case were:

CHANGE IN PAR =
-9.68 + 0.82(PRE-TREATMENT PAR) + 3.89(POSTGRAD.) + 3.13(ANTERIOR XB) + ε

Variable	Beta
PRE-TREATMENT PAR	0.79
POSTGRADUATE	0.14
ANTERIOR CROSSBITE	0.13

The assumptions of regression were acceptable (Figures 44-45) and the residual distribution did not differ substantially from normal (P>0.20).

Mean and individual 95% prediction intervals are shown in Table 47.

Pre-treatment PAR		16	25	35
Postgrad.				
	Mean	3.8 - 11.0	11.4 - 18.1	19.4 - 26.6
	Individual	0.0 - 19.0	3.2 - 26.3	11.4 - 34.6
Not postgrad.				
	Mean	1.6 - 5.4	9.4 - 12.4	17.0 - 21.2
	Individual	0.0 - 14.7	0.0 - 22.0	7.9 - 30.4

Table 47

Normal Plot of Residual

Other appliances: Change in PAR

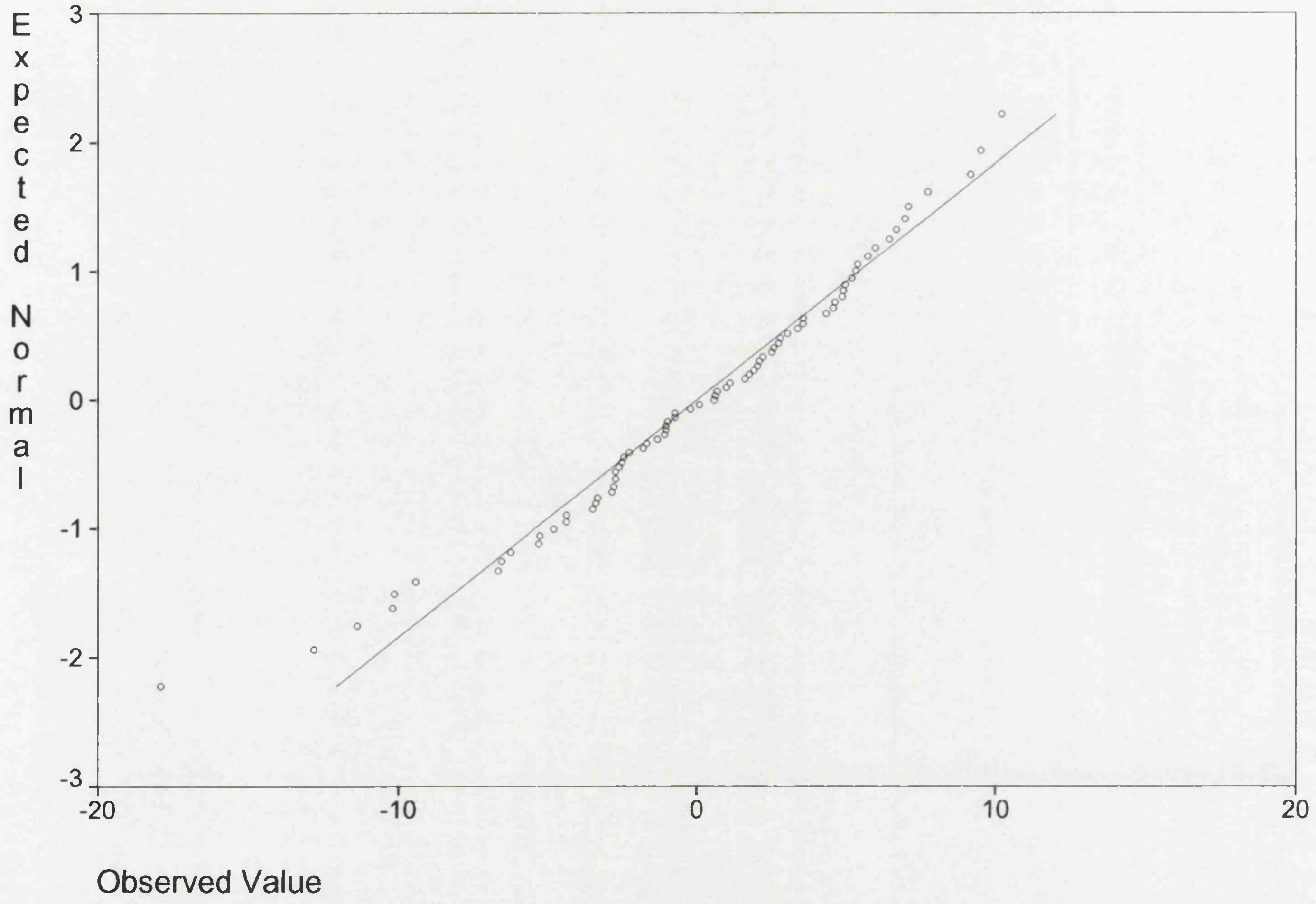


Figure 44

Scatterplot: Residual/Predicted value

Other appliances: Change in PAR

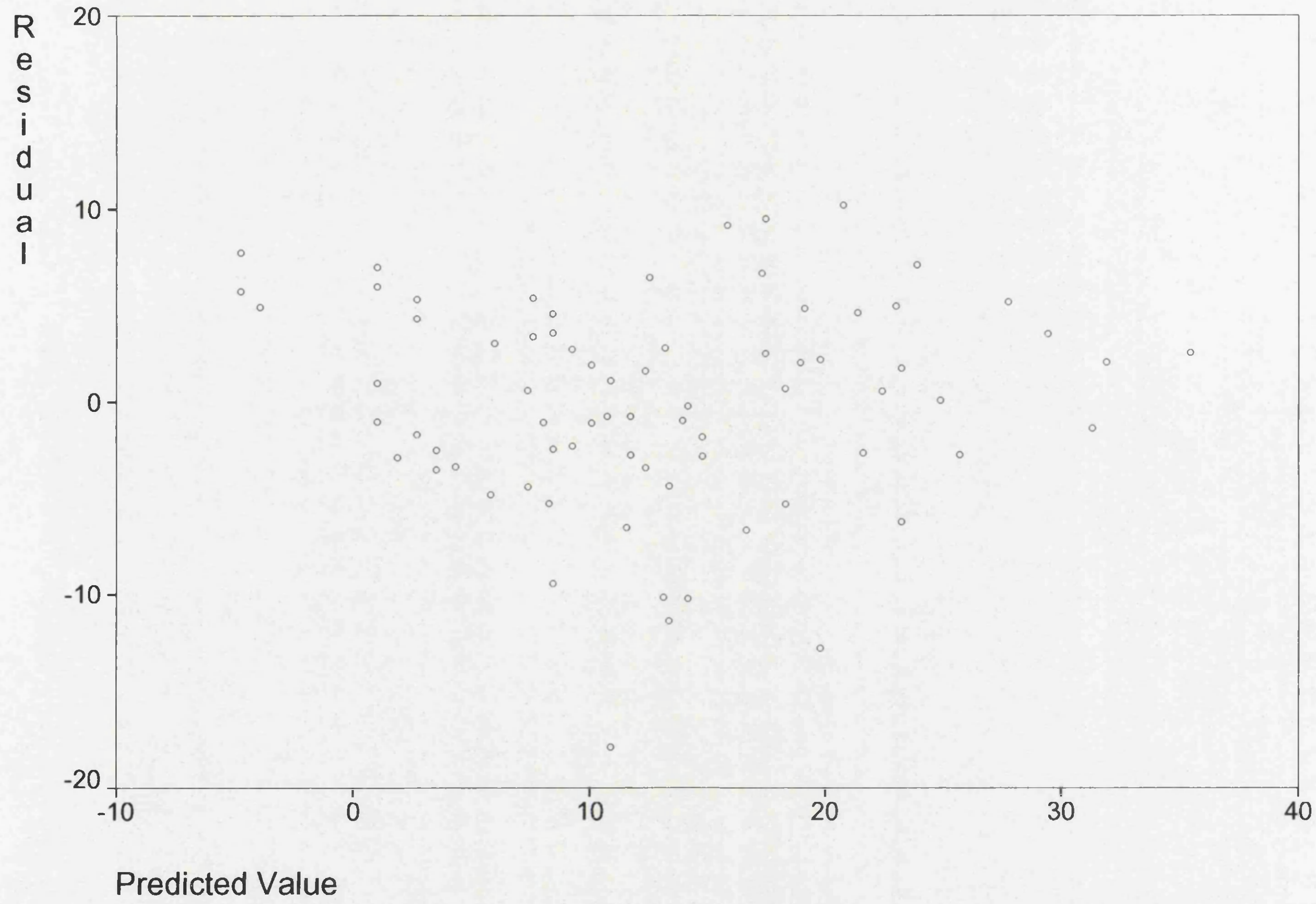


Figure 45

11.3 Duration of treatment

11.3.1 Full fixed appliances

a) All information (Duration/Full fixed)

Regressing on duration of active treatment in months for Full fixed appliances only, produced a model with an R^2 of 0.32 and included pre-treatment PAR, pre-treatment IOTN grade, sex, and incisor classification as predictor variables, however the assumptions of normality and equal variance were not acceptable.

Regressing on the natural log of the duration of active treatment resulted in a model which included the following variables:

Number of appointments during active treatment

Pre-treatment PAR

Presence of an anterior crossbite

R^2 for this model was high at 0.77 (R_a^2=0.76) and the assumptions of normality and equal variance were acceptable (Figure 46-47).

The regression equation and beta values for the predictor variables were:

$$\ln(\text{DURATION}) = 2.04 + 0.04(\text{No. APPTS.}) + 0.01(\text{PRE-TREATMENT PAR}) + (-0.10(\text{ANTERIOR XB})) + \varepsilon$$

Variable	Beta
No OF APPOINTMENTS	0.78
PRE-TREATMENT PAR	0.17
ANTERIOR CROSSBITE	-0.13

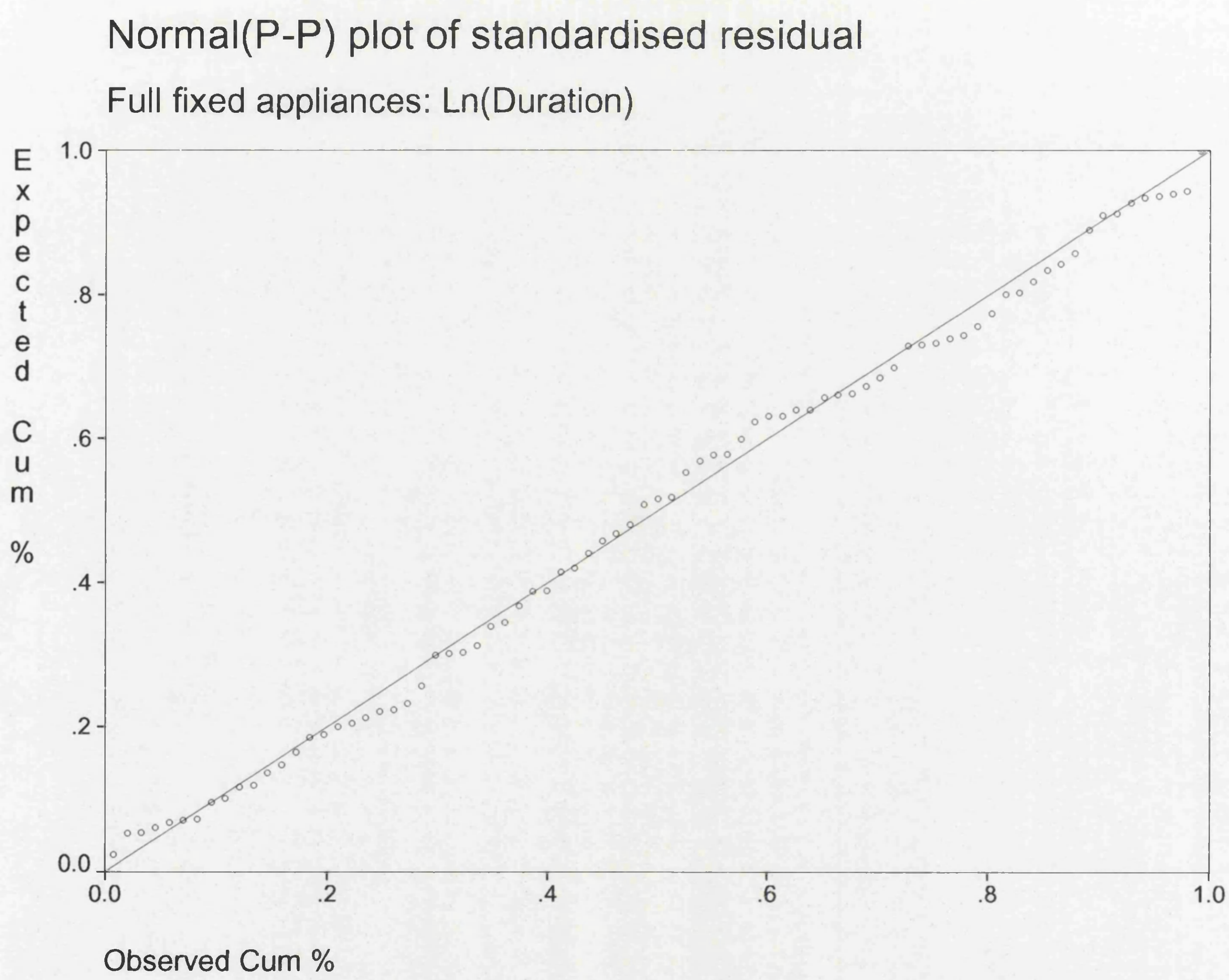

Normal(P-P) plot of standardised residual
Full fixed appliances: Ln(Duration)
Expected Cum %
1.0
.8
.6
.4
.2
0.0
0.0
.2
.4
.6
.8
1.0
Observed Cum %

Figure 46

Histogram: Residual distribution

Full fixed appliances: Ln(Duration)

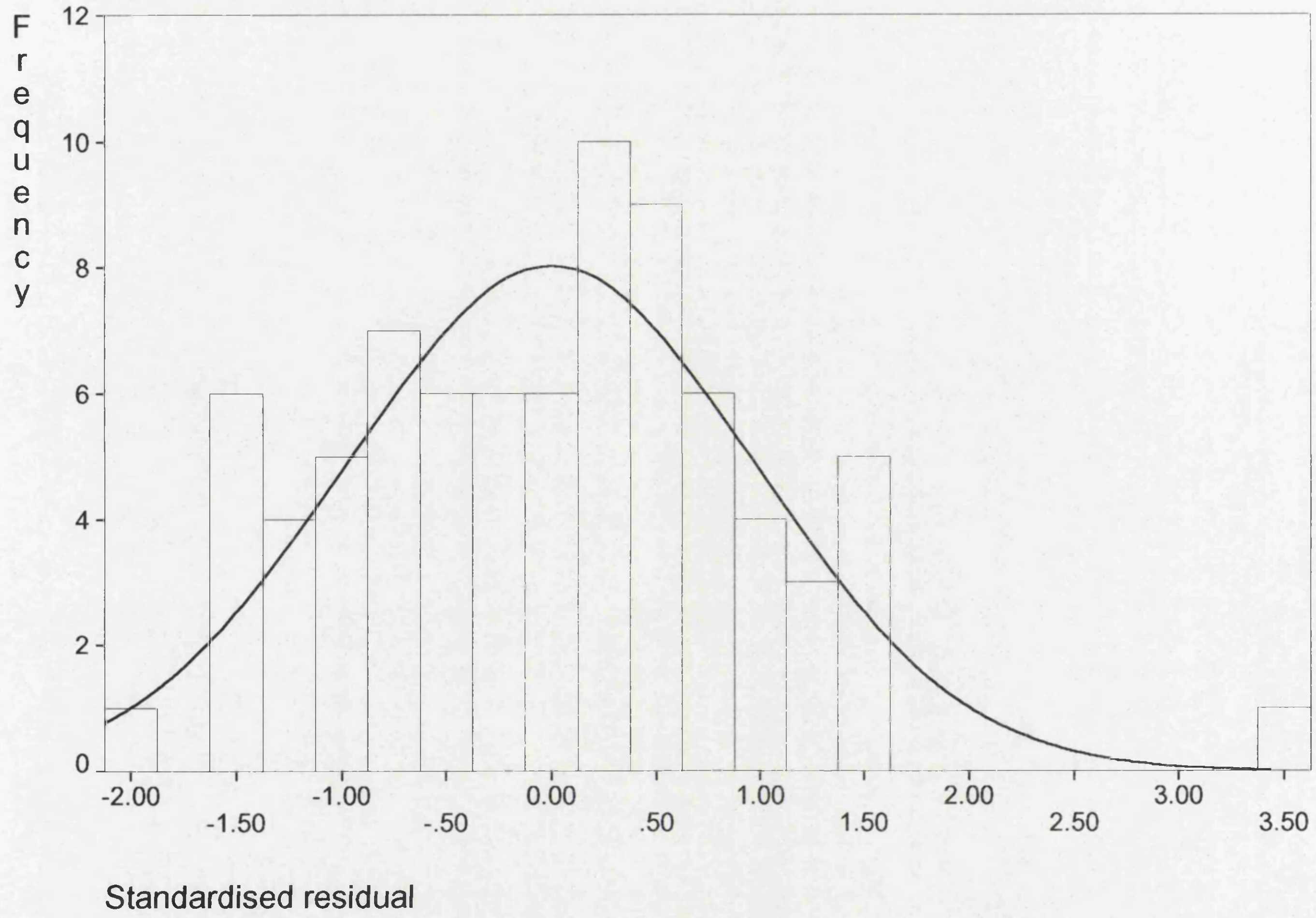


Figure 47

The above regression procedure was repeated but excluding the number of planned appointments as a predictor variable.

In this case the variables entered into the equation were:

Pre-treatment PAR score

Whether or not the incisor classification was Class II division 2

The sex of the patient

Whether or not an anterior crossbite was present initially

R^2 for this model was 0.42 (R_a^2=0.39) and the assumptions required for regression were acceptable (Figure 48, Lilliefors statistic=0.063, P>0.20).

The regression equation in this instance was:

ln(DURATION) =
2.39 + 0.02(PRE-TREATMENT PAR)+0.30(CLASS II/2) + 0.15(FEMALE) + (-0.17(ANTERIOR XB)) + ε

Beta values for these variables were:

PRE-TREATMENT PAR	0.52
CLASS II/2	0.32
FEMALE	0.23
ANTERIOR XB.	-0.23

Normal(P-P) plot of standardised residual

Full fixed appliances: Ln(Duration)

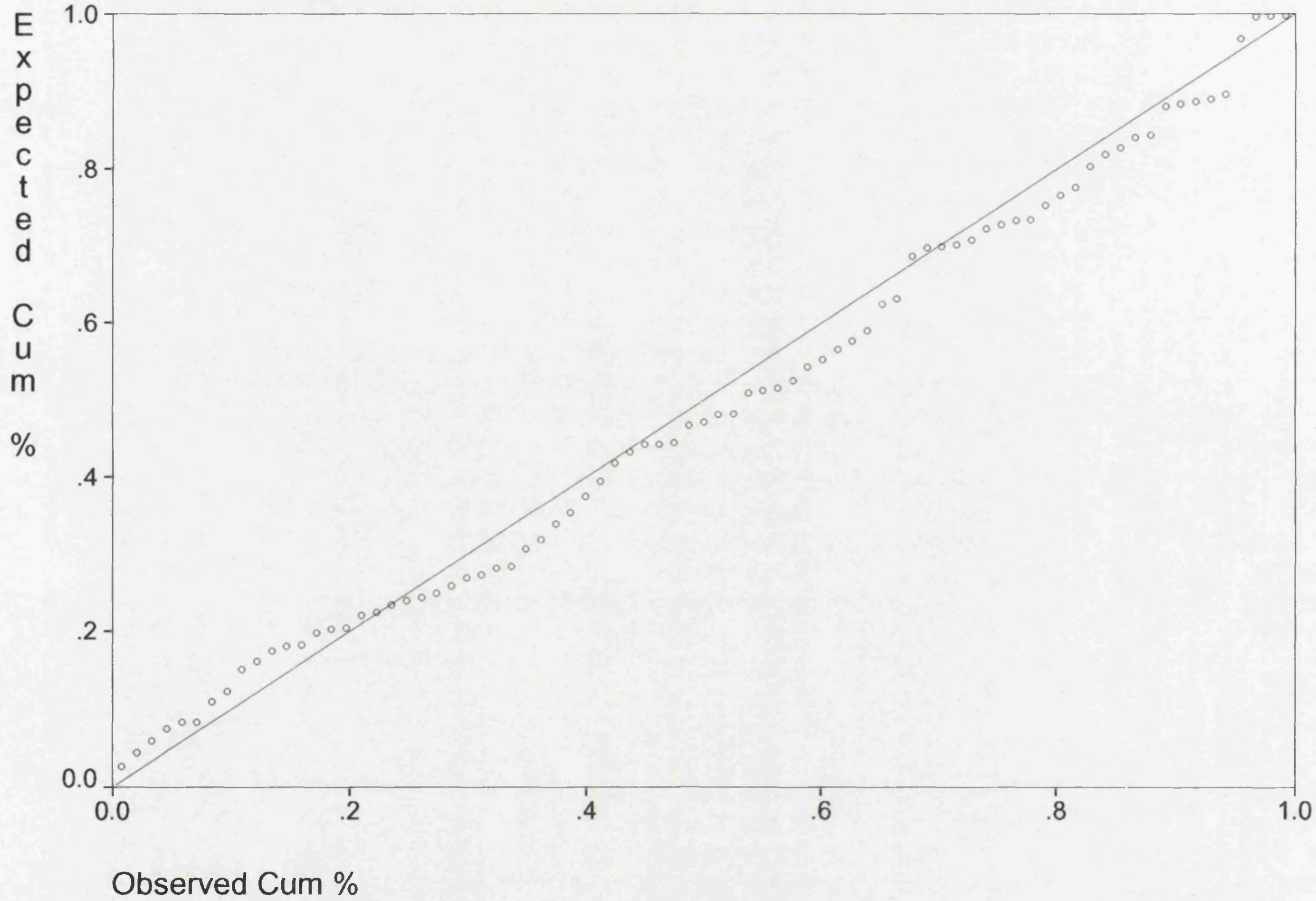


Figure 48

95% prediction levels for duration in the absence of an anterior crossbite and in other than Class II/2 cases were (Table 48):

Pre-treatment PAR	21	32	43
Female			
Mean	16.6 - 20.1	20.6 - 23.9	24.3 - 30.0
Individual	11.4 - 29.3	13.9 - 35.5	16.8 - 43.5
Pre-treatment PAR	20	31	42
Male			
Mean	13.9 - 17.3	17.3 - 20.8	20.2 - 26
Individual	9.6 - 25.0	11.7 - 30.3	14.2 - 37.1

Table 48

b) Pre-treatment information (Duration/Full fixed)

It was observed that the variables appearing in the regression equation for ln(Duration) above all have values which would be known prior to the start of treatment.

11.3.2 Other appliances

a) All information (Duration/Other)

Regressing on duration of treatment with all relevant variables produced a model which included the following as predictor variables:

Number of appointments during active treatment

Number of broken appointments during active treatment

Whether or not incisor relationship was class II division 1

Age at start of treatment

R^2 was high at 0.80 (R_a^2=0.79) however the graphical evidence that the assumptions of normality and of equal variance were acceptable was not convincing and Lilliefors statistic (0.10) indicated that the distribution of the residuals was close to being significantly different from normal (P=0.06).

Transforming the dependent variable to the natural log of duration and repeating the regression yielded an equation which involved only the number of appointments during treatment and the number of broken appointments. R^2 for this model was 0.72, however the detrended normal plot (Figure 49) revealed the presence of an outlier. This was identified as the same case as that which was investigated with regard to the minimum value for duration of treatment. In this instance it was decided to withdraw the case due to its excessive influence on the acceptability of the regression.

A further regression was performed without the outlying case and with ln(duration) as the dependent variable. A model with R^2=0.80 (R_a^2=0.79) was constructed. The variables included in the model were:

Number of appointments

Number of broken appointments

Age at start of treatment

Presence of an anterior crossbite initially

The assumptions of regression were acceptable (Figure 50-51).

Detrended Normal Plot of Residual

Other appliances: Ln(Duration)

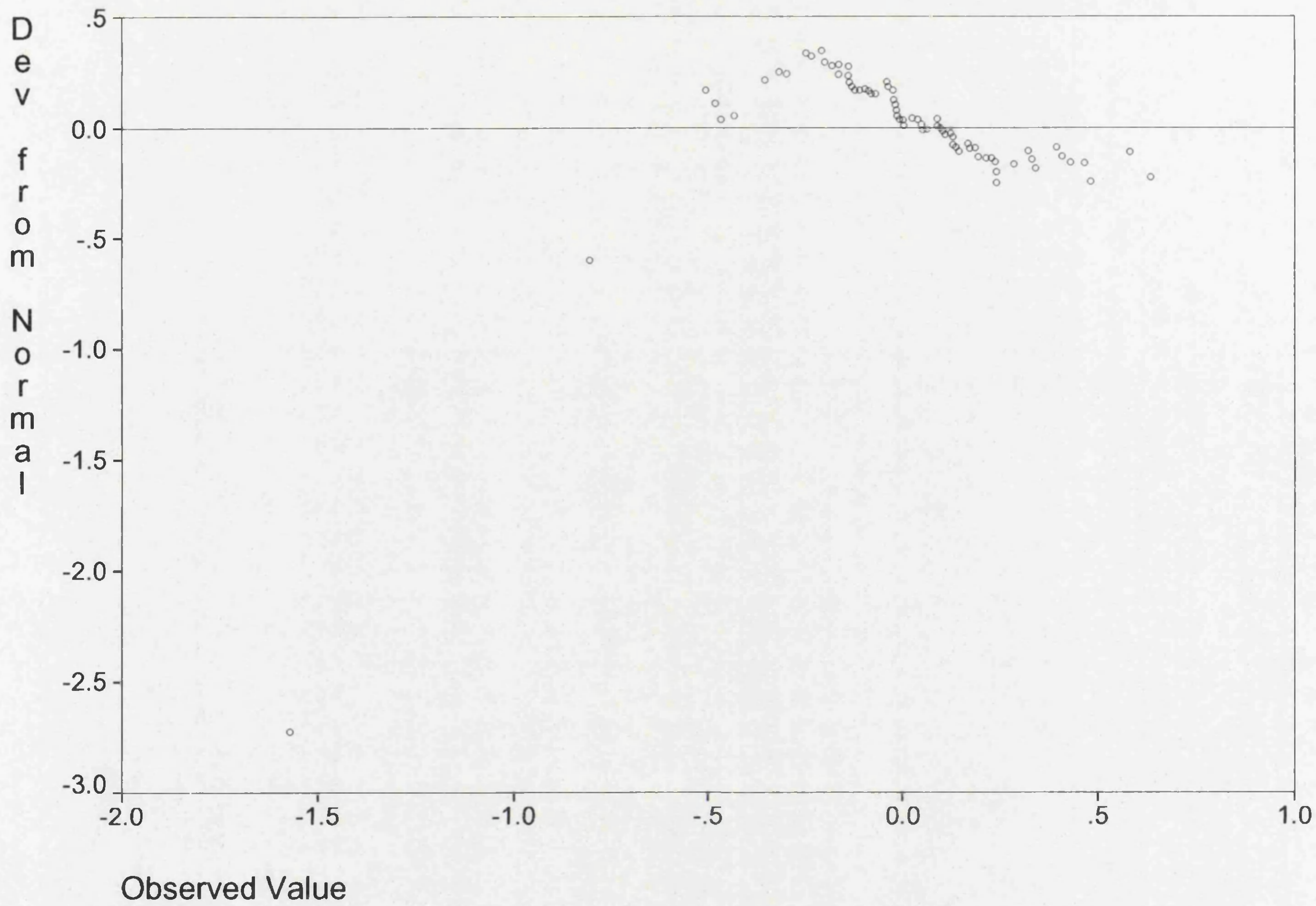


Figure 49

Normal Plot of Residual

Other appliances: Ln(Duration)

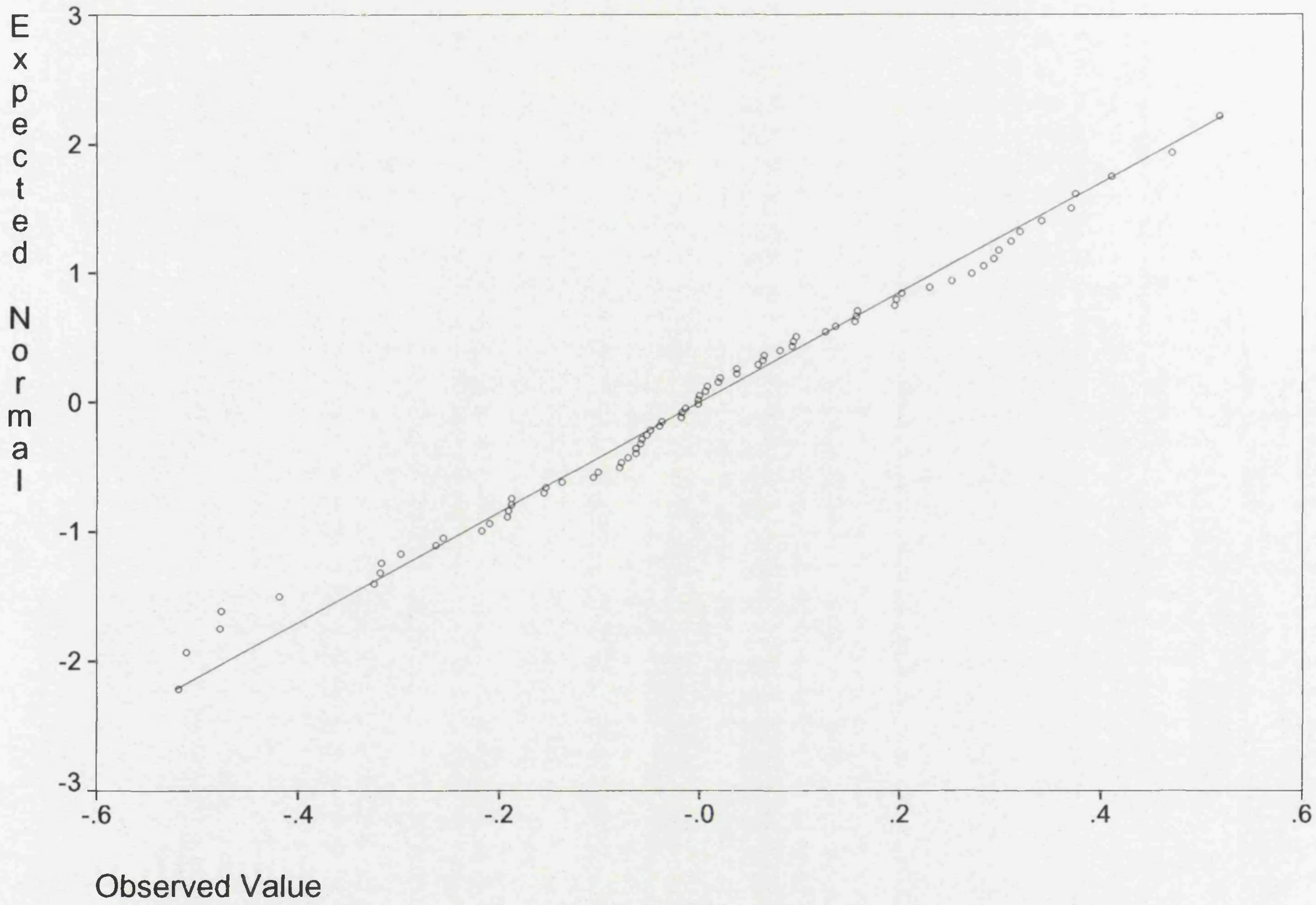


Figure 50

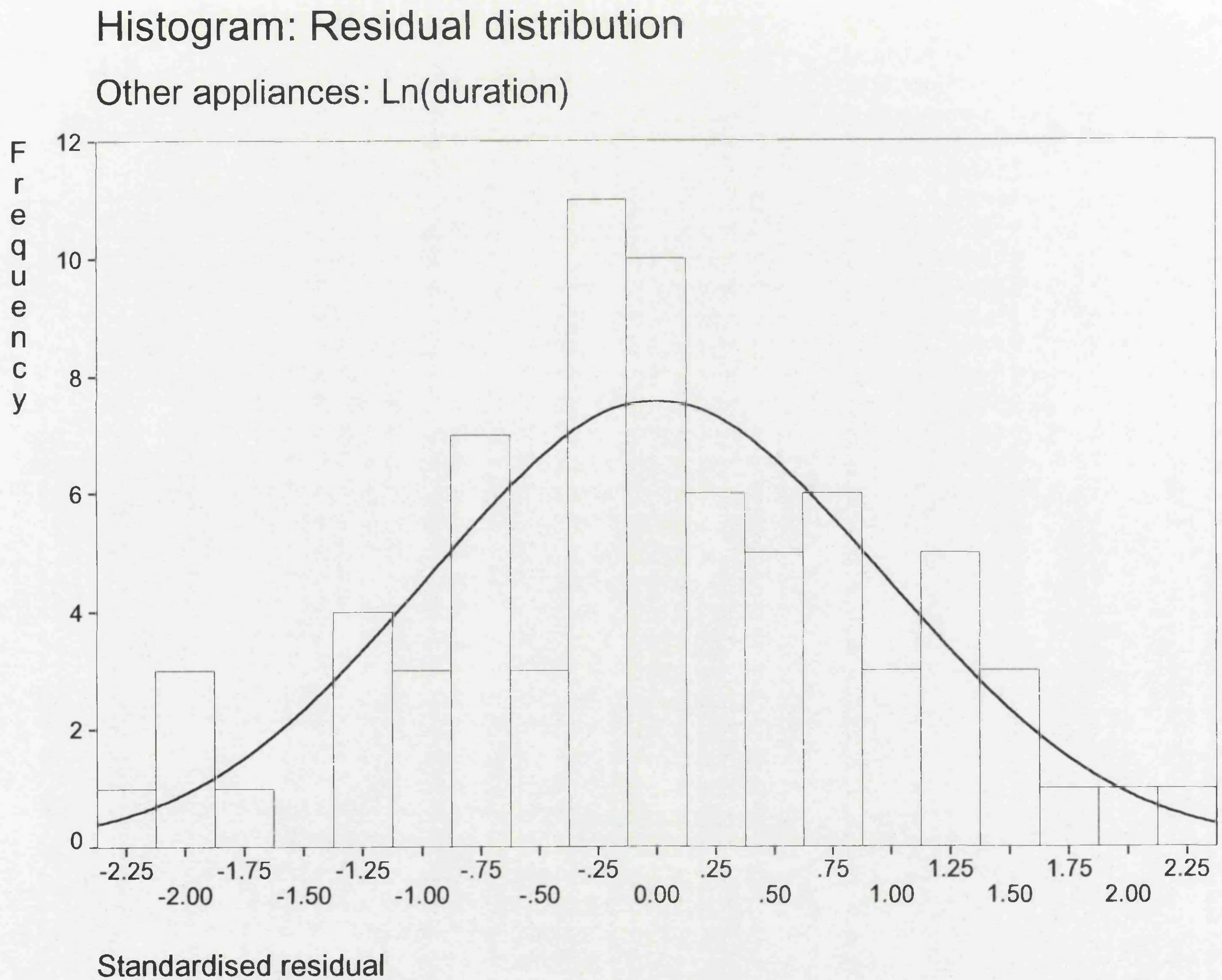

Histogram: Residual distribution
Other appliances: Ln(duration)
Frequency
0
2
4
6
8
10
12
-2.25
-2.00
-1.75
-1.50
-1.25
-1.00
-.75
-.50
-.25
0.00
.25
.50
.75
1.00
1.25
1.50
1.75
2.00
2.25
Standardised residual

Figure 51

The regression equation and beta values for this model are shown below:

ln(DURATION) =
2.00 + 0.07(No APPTS.) + 0.06(No BROKEN APPTS.) + (-0.03(AGE AT START) + (-0.15(ANTERIOR XB)) + ε

Variable	Beta
No OF APPOINTMENTS	0.77
No OF BROKEN APPTS.	0.23
AGE AT START	-0.19
ANTERIOR CROSSBITE	-0.12

A further regression was carried out on ln(duration) with number of appointments withheld from the equation. The following variables were included in the equation:

Number of removable appliances used

Number of fixed components damaged or dislodged

Whether teeth were extracted or not

Number of broken appointments

Whether or not the incisor relationship was Class III

R^2 for this model was only moderate at 0.55 (R_a^2=0.52) and the assumptions of regression in this case were considered to be unacceptable (Figure 52).

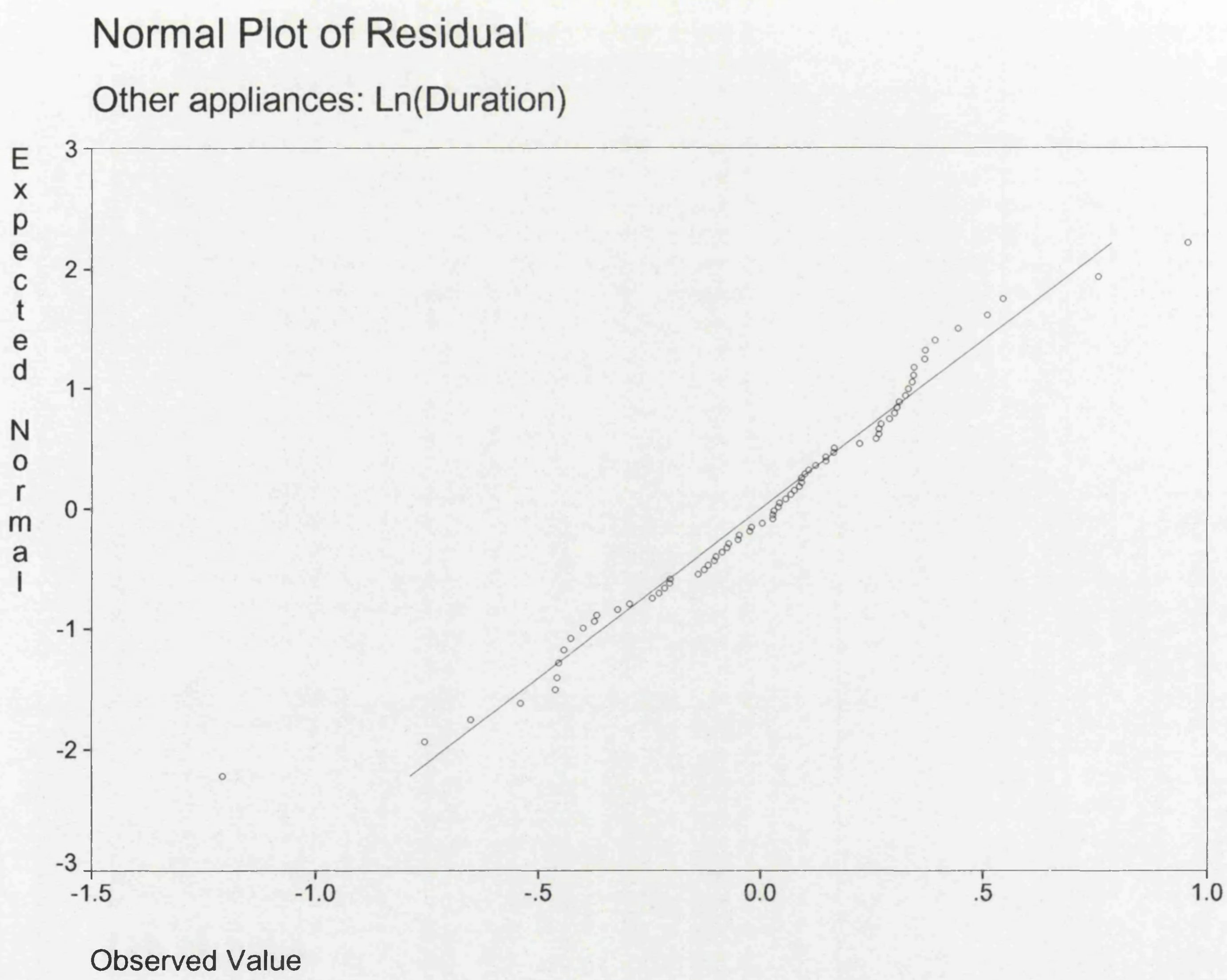

Normal Plot of Residual
Other appliances: Ln(Duration)
Expected Normal
3
2
1
0
-1
-2
-3
-1.5
-1.0
-.5
0.0
.5
1.0
Observed Value

Figure 52

b) Pre-treatment information

In contrast to the number of archwires used in a fixed appliance treatment, where the decision to change an archwire and which wire to place is generally made on a visit to visit basis depending on progress, it may be feasible to plan the number of removable appliances that will be used during a removable appliance case. For this reason the number of removable appliances used was included in this analysis as a predictor variable of known quantity at the start of treatment.

Regressing on the natural log of treatment duration using independent variables with values known at start of treatment resulted in a regression equation with the following factors included:

Whether or not extraction of teeth was planned as part of treatment

Whether or not removable appliances only were used

The number of removable appliances used during treatment

The regression coefficients and beta values for the included variables are given below:

$$\ln(\text{DURATION}) = 2.58 + (-0.30(\text{NON-XTN})) + (-0.39(\text{REMOVABLES ONLY})) + 0.23(\text{No OF REMOVABLES}) + \varepsilon$$

Variable	Beta
NON-EXTRACTION THERAPY	-0.28
REMOVABLES ONLY	-0.36
No OF REMOVABLES	0.51

R^2 for this model was 0.36 (R_a^2=0.33).

The assumptions of regression for this model were found to be acceptable (Figures 53-54). Lilliefors' test statistic to indicate departure from a normal distribution in this instance was 0.053 with an associated level of significance $P>0.20$.

95% prediction intervals are shown below (Table 49):

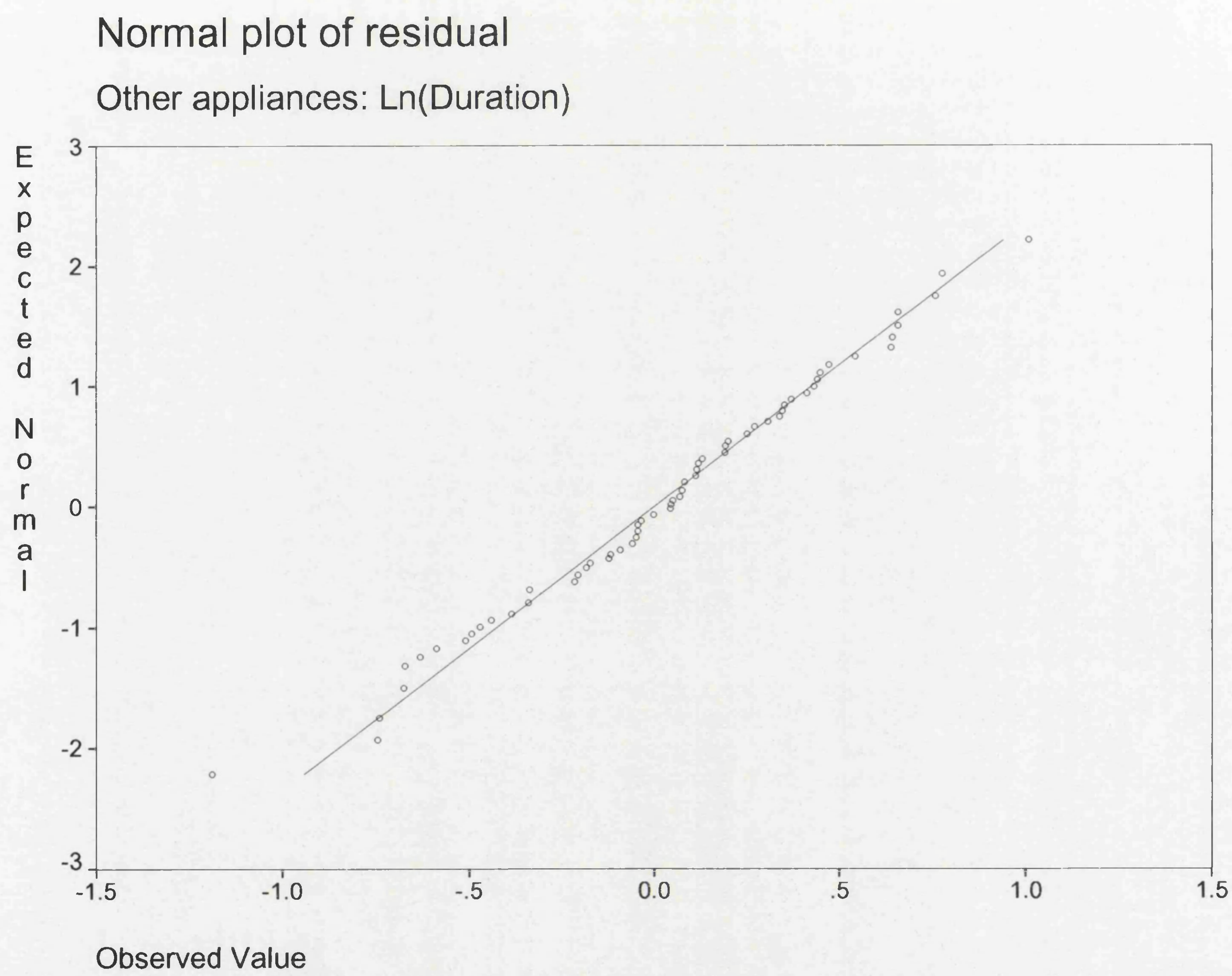
Removable and fixed components

No. of appliances	0 (Mini-fixed)	1	2
Non-extraction			
Mean	8.3 - 11.6	10.6 - 14.5	12.7 - 19.3
Individual	4.1 - 23.8	5.1 - 29.9	6.4 - 38.2
Extraction			
Mean	10.5 - 16.7	13.9 - 20.1	17.3 - 25.7
Individual	5.4 - 32.4	6.9 - 40.5	8.7 - 51.3

Removable only

No. of appliances	1	2	3
Non-extraction			
Mean	6.8 - 10.4	8.6 - 13.0	10.5 - 17.1
Individual	3.4 - 20.5	4.4 - 25.8	5.5 - 32.9
Extraction			
Mean	8.9 - 14.3	11.8 - 17.2	14.8 - 21.8
Individual	4.6 - 27.7	5.9 - 34.6	7.4 - 43.7

Table 49



Normal plot of residual
Other appliances: Ln(Duration)
Expected Normal
3
2
1
0
-1
-2
-3
-1.5
-1.0
-.5
0.0
.5
1.0
1.5
Observed Value

Figure 53

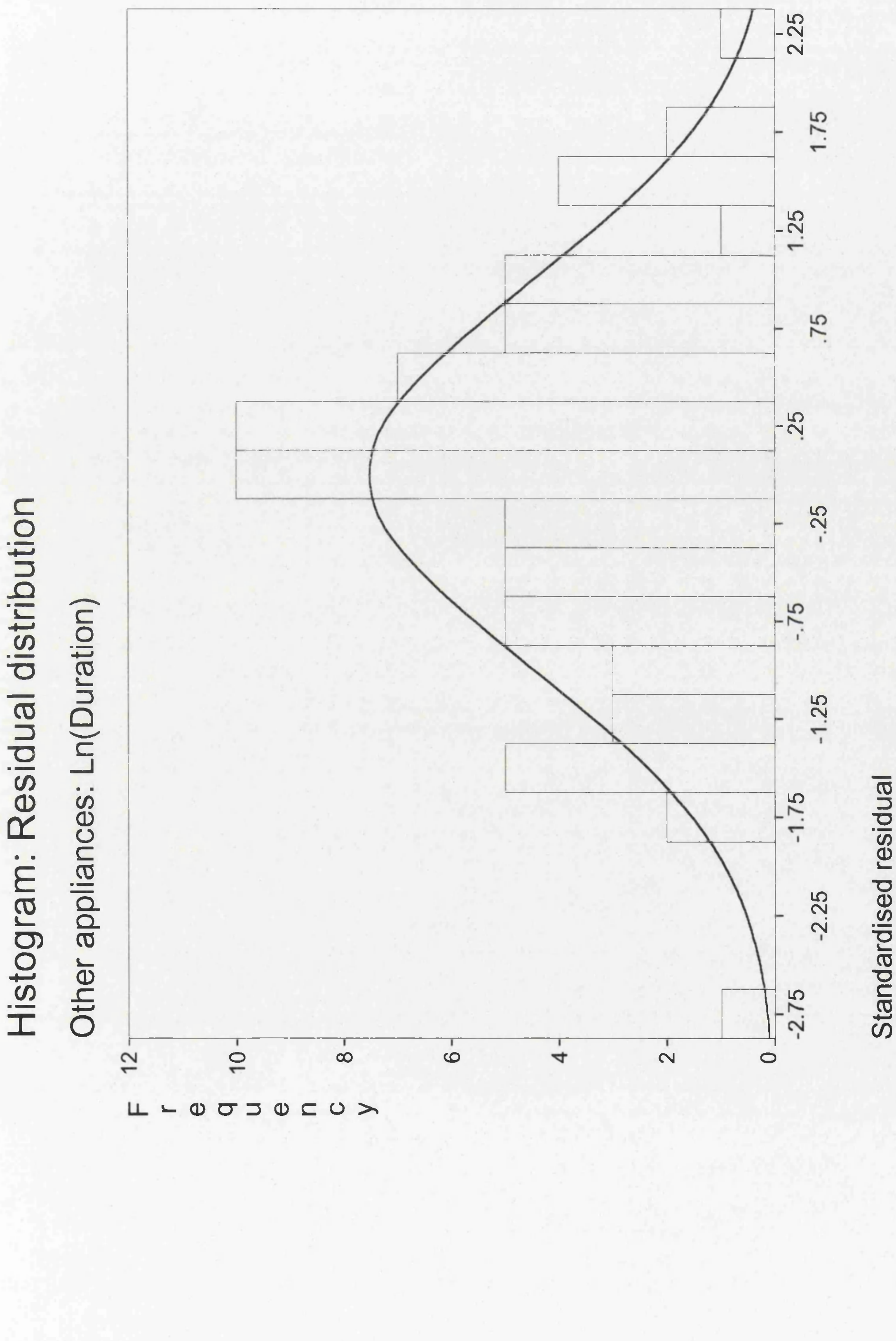

Histogram: Residual distribution
Other appliances: Ln(Duration)
Frequency
0
2
4
6
8
10
12
-2.75
-2.25
-1.75
-1.25
-.75
-.25
.25
.75
1.25
1.75
2.25
Standardised residual

Figure 54

11.4 Regression analysis: Discussion

11.4.1 Post-treatment PAR

In previous attempts at quantitative prediction of treatment outcome, anticipated quality of finished result has not been investigated although prediction of treatment duration (Vig *et al.*, 1990; Fink and Smith, 1992; Kerr *et al.*, 1993) and occlusal change due to treatment (Kerr *et al.*, 1993) have been considered.

The adage coined by Mills (1978) that extraction of permanent first molars "doubles the treatment time and halves the prognosis" has received some support in this study, although not to the point of agreement on degree and certainly not in every case, given the width of the individual 95% prediction intervals of approximately 16 PAR points.

Alteration of the treatment plan had a greater effect on the predicted post-treatment PAR score than any of the other explanatory variables though for full fixed appliances an increase in pre-treatment Aesthetic Component and the extraction of a first permanent molar would also tend to increase the post-treatment score. The presence of an initial anterior crossbite would tend to reduce the post-treatment score.

The Aesthetic Component could be considered a more subjective and less reproducible measurement than PAR as highlighted by the error study. A further regression excluding pre-treatment Aesthetic Component resulted in this variable being substituted by pre-treatment PAR score in the equation. This might be more clinically useful.

Examination of the prediction intervals suggest that, on average, full fixed cases where the treatment plan must be altered will have post-treatment PAR scores approximately 6 points higher than otherwise would be the case. Whether or not a first permanent molar was extracted would make an average difference of approximately 4½ PAR points. The goodness of fit of the model as indicated by the R^2 value of 0.40 was only moderate however and the wide prediction intervals emphasise that the above figures would have only a limited clinical application on an individual basis. For example, a case with pre-treatment PAR of 35, where the treatment plan was adhered to, where treatment did not involve extraction of a first permanent molar and where there was no anterior crossbite would have a 95 percent chance of finishing with a post-treatment PAR score of between 1.7 and 17.1 PAR points.

The inclusion of anterior crossbite in the regression equation may be associated with edge to edge incisor relationships or mild crossbites in crowded Class I cases. Generally speaking correction of these traits is readily achieved. The small changes in incisor position required to correct such relationships result in a considerable reduction in PAR score due to PAR's overjet weighting. Crossbite correction in Class III cases may be more problematical. In this sample there were only a small number of Class III cases.

The mean prediction intervals for cases where a first permanent molar was extracted were larger than those for other cases. This may be a function of the difference in numbers of cases within the two groups.

Subsequent alteration of treatment plan during the course of treatment cannot be known prior to commencement. Exclusion of this factor to determine pre-treatment predictors resulted in a regression with a much lower R^2 of 0.23 confirming the importance of change in treatment plan to the standard of finish. Further support for alteration of the treatment plan being a relatively important factor was a beta value for that variable of 0.48. In this model good oral hygiene was found to reduce the predicted post-treatment PAR score. With so little variability having been explained the clinical usefulness of any prediction must be in doubt, however it is notable that such an important clinical entity as oral hygiene has been highlighted as a predictor variable by the model.

Thus it has not been possible to predict the quality of finished result following fixed appliance therapy with any great degree of accuracy.

Concerning cases which were not treated with full fixed appliances, mean and individual prediction intervals were found to be consistent throughout the range of numbers of active appointments although the intervals themselves were large due to the moderate R^2 of 0.41. Whether or not the original treatment plan had been altered was again found to be an important factor although the presence of a Class II division 2 incisor relationship also contributed substantially to an increased predicted post-treatment PAR score as did the presence of a Class II division 1 relationship and whether or not it had been treatment by an undergraduate student. Predicted post-treatment irregularity reduced with increasing number of appointments and use of headgear.

Where only information known at the start of treatment was included a satisfactory model could not be constructed for post-treatment PAR score with the variables employed in this study.

It is clear that with this sample and the information collected here it is difficult to accurately predict the standard of result as measured by PAR although the factors which have been identified as having a bearing on result quality are not different from those which, from clinical experience, might be expected to be important.

11.4.2 Change in PAR

Kerr *et al.*, (1993) found pre-treatment PAR score to be the most influential predictor of improvement through both removable appliance therapy and functional/headgear therapy. In this study the same variable was found to be by far the most important predictor of change in PAR for both full fixed and non-full fixed treatments.

Change in PAR must depend on pre-treatment PAR to some extent as change in PAR = pre-treatment PAR - post-treatment PAR, thus there is a correlation between initial PAR score and change in PAR which contributes towards the higher values of R^2 for regressions on change in PAR.

It was interesting to note that regarding full fixed appliances the variables which appeared in the regression equation for change in PAR were the same as those for post-treatment PAR. On average the need to deviate from the original treatment plan and the planned loss of a permanent first molar tended to reduce the predicted improvement by close to 6 points and 4 points respectively whilst the presence of an anterior crossbite tended to increase the improvement by 3 PAR points. The 95% prediction intervals for individual cases were consistent at approximately 16 PAR points.

In comparison with the findings reported by Kerr *et al.*, (1993) from a group treated with both removable and fixed components, the models constructed for both Full fixed and Other groups generally showed similar but slightly higher values for R^2. The larger number of explanatory variables used here may explain the difference between the two studies rather than any difference in homogeneity of samples although in this study the Full fixed group formed a more homogenous sample than the Other group.

Details of examination of residuals of regression were not included with the report of Kerr *et al.* so that an indication of the degree of acceptability of the assumptions of regression was not available. In this study the assumptions of regression on change in PAR were acceptable although the normal plots of residuals were not precisely linear.

Restriction of possible predictor variables to only those whose values are known at the beginning of treatment resulted in a regression equation with the same entered variables as that obtained when all factors were included with the exception of alteration of treatment plan. R^2 for this model was still high at 0.78. Consistent with the small reduction in R^2 between the two models, the individual 95% prediction intervals were approximately 18 PAR points wide.

For Full fixed appliance therapy, therefore, it is possible to predict change in PAR through treatment with reasonable accuracy

Concerning cases in the Other appliance group, again the most important determining factor for change in PAR through treatment was the pre-treatment PAR score. The regression performed using all relevant variables produced a model which included five variables along with pre-treatment PAR but beta values for those variables were low. The signs of the correlation coefficients for the entered variables which indicate the direction of effect on the predicted change in PAR were consistent with what might be expected from clinical experience with the exception of incisor classification. Lack of co-operation as indicated by the treatment plan being altered, the operator being an undergraduate, and a Class II incisor relationship all tended to reduce the predicted change in PAR. Increasing the number of appointments during active treatment tended to increase the change in PAR.

It is difficult to explain the effect that the presence of Class II/1 and Class II/2 incisor relationships have on this regression however there may be an association between the higher proportion of teeth in anterior crossbite among the Class I and Class III cases which may have been treated with simple removable appliances.

This possibility is supported by the inclusion of anterior crossbite as an explanatory variable in the model for change in PAR with non-full fixed appliances where only pre-treatment information was considered. Whether or not the operator was a postgraduate student was also included along with pre-treatment PAR. Both the operator being a postgraduate and there being an anterior crossbite present increased the predicted change in PAR. On average having a postgraduate as operator increases the PAR change by approximately 4 points. The individual 95% prediction intervals were wider than those for full fixed appliances at about 23 PAR points which is to be expected given the lower value for R^2 of 0.72.

11.4.3 Duration of treatment

The ability to accurately predict the duration of a particular course of treatment prior to starting might be useful for patient advice and for management of resources. Treatment durations in this sample ranged from one month to four years with a mean of 18.3 months and standard deviation of 7.9 months. Within the Full fixed group a well fitting model (R^2=0.77) for duration was obtained by transforming the dependent variable to its natural logarithm. Once again pre-treatment PAR featured as an predictor variable which tended to increase treatment duration with increasing value whereas the presence of an anterior crossbite tended to reduce the predicted duration. However the most important explanatory variable of this model was number of appointments during treatment.

It is perhaps not surprising that the number of appointments should be such a major factor as appliances require to be maintained and inspected at regular intervals regardless of the rate of progress.

With the exclusion of number of appointments the pre-treatment PAR score appeared to be the most influential factor in determining the overall duration of treatment, with the presence of a Class II/2 incisor relationship and the patient's sex also contributing. On average the predicted treatment duration for females was approximately 3 months longer than that for males. On the other hand, the presence of an anterior crossbite is a factor which tends to a reduction in duration.

It is possible that the appearance of Class II/2 in the equation tending to increase treatment duration is associated with time required for overbite reduction.

Unfortunately only 42% of the variability in treatment duration was explained by this model which compares with a figure of 60% obtained by Kerr *et al.* (1993) from a regression on a sample of removable appliance cases which included pre-treatment PAR as the most important factor. The prediction intervals measured in months were wide and became wider as the value for pre-treatment PAR increased so that prediction of treatment duration became less accurate with increasing pre-treatment PAR score.

It is interesting to observe that female patients within the Full fixed group tended towards longer durations of treatment and, as noted previously, had generally lower PAR scores after treatment than males. This prompts the suggestion that greater time and effort might be spent on achieving higher standards of finish in females.

For appliances other than full fixed appliances, the number of appointments during active treatment was again the most important explanatory variable. As in the study reported by Fink and Smith (1992), number of broken appointments also tended to increase treatment duration. However it may be there is a greater chance of appointments being broken the longer a treatment continues.

In the present study duration tended to be reduced where there was an anterior crossbite initially, whereas the earlier a treatment started the longer the duration. The latter finding is in agreement with that of Vig *et al.*, (1990) who also obtained positive partial correlation coefficients for extraction as opposed to non-extraction and Class II/2 treatments although an R^2 of 0.33 for that model could not be considered high.

A regression performed with number of appointments excluded produced an equation comprising number of broken appointments, whether or not teeth were extracted and number of removable appliances used. These factors were amongst variables highlighted as possible predictors of duration by other studies (Vig *et al.*, 1990; Fink & Smith, 1992; Kerr *et al.*, 1993), though in this study the assumptions for regression were not acceptable. Thus, for duration of treatment with non-full fixed appliances, no suitable predictive model using all variables could be found.

An acceptable model for treatment duration was constructed where only pre-treatment information was considered, although just 36% of the variability in duration was explained. In this model, non-extraction therapies on average took 3.7 months less to complete than extraction treatments where removable appliances only were employed, and 6.5 months less where some fixed components were used. For Mini-fixed cases the mean predicted treatment duration was 10 months in non-extraction treatments and approximately 14 months in extraction cases.

For the removable appliances only group, average predicted duration of treatment was increased by just 2.5 months as a second appliance was used and by another 3.4 months if a third appliance was added.

Despite the assumptions of regression being acceptable for this model, the very wide prediction intervals and the fact that, for Mini-fixed appliances, the only applicable predictor in the equation was whether or not teeth had been extracted, the model must be regarded as having limited usefulness.

A study of factors associated with orthodontic treatment outcome.

Section E

CONCLUSIONS

Chapter 12

Conclusions

1. Significant differences between appliance groups were identified in a number of patient factors, treatment factors and outcome variables:

a) Patient factors: Severity of initial malocclusion as measured by PAR and by size of overjet was generally greater in full fixed appliance cases than non-full fixed cases.

b) Treatment factors: Full fixed appliance therapy involved more cases where four premolars were extracted and fewer non-extraction cases than other types of treatment. Also a greater number of appointments was required to complete full fixed cases. A substantial proportion of non-full fixed cases were carried out by undergraduate students whereas no undergraduate performed full fixed appliance therapy.

c) Outcome variables: Full fixed appliance treatments showed a lower level of residual irregularity and greater improvement through treatment as measured by PAR, but needed longer courses of treatment to achieve this.

2. The differences highlighted suggest that in considering the relative efficacy of different appliance systems, assessments made with PAR should not be directly compared without reference to other group differences, particularly as change in PAR is so strongly correlated with pre-treatment PAR.

3. Although sex distribution was similar in both full fixed and non-full fixed groups, when related to pre-treatment PAR score within the non-full fixed group females had considerably lower pre-treatment PAR scores than males. This suggests that males with severe malocclusions might be given either full fixed appliances or removable appliances whereas females with such malocclusions would be less likely to receive non-full fixed appliances.

4. It was not possible to predict post-treatment PAR for any appliance group with any degree of accuracy in individual cases, but on average, alteration of the treatment plan led to a 6 point increase in post-treatment PAR. Models for post-treatment PAR restricted to pre-treatment information explained only a small proportion of the variability and would have limited clinical usefulness.

5. Models were produced for both appliance groups which substantially explained the variability in change in PAR. Reasonably accurate prediction of change in PAR is possible from pre-treatment information.

6. For full fixed appliances, the most important factor affecting all the treatment outcomes examined was the pre-treatment PAR score, with patient co-operation, as indicated by maintenance of the original treatment plan, also of considerable influence.

7. Prediction equations including clinically useful pre-treatment variables were also constructed for duration of active treatment. However accuracy of prediction was low and the model for non-full fixed treatments would not be helpful in cases where only fixed components were used.

8. Better fitting regression models might be obtainable for non-full fixed treatment types if different modes of treatment were considered separately. This would require larger samples of each treatment type than were available in this study.

9. Treatments in this study had been carried out in the orthodontic department of a large dental teaching hospital and so had involved a number of different operators with differing experience, training levels and backgrounds. It may be that more accurately predictive models could be produced by a smaller department or a single practice through reduction in diversity of approaches to treatment, thereby increasing sample homogeneity.

The material for this study originated from cases commenced in 1985. Recent developments such as super-elastic aligning wires and pre-adjusted brackets have become more widely used. Also the attention of the Specialty has become focused lately on treatment standards, efficient use of resources and clinical audit with the increased use of occlusal indices. Alterations in orthodontic practice such as these together with changes in public awareness may lead to changes in profile of the caseload now being undertaken and other factors determining treatment outcome may now be more influential.

Although clinical practice continues to evolve and changes in materials, techniques and attitudes will have taken place, the findings of this study might be regarded as baseline figures against which future information may be compared and trends determined.

BIBLIOGRAPHY

Alexander, A. G. & Tipnis, A. K., (1970); The effect of irregularity of teeth and the degree of overbite and overjet on the gingival health.
British Dental Journal,**128**,539-544

Andrews, L.F., (1972); Six keys to normal occlusion.
American Journal of Orthodontics,**62**,296-309

Angle, E.H. (1899); Classification of malocclusion.
Dental Cosmos,**41**,248-264

Asher-McDade, C., Roberts, C., Shaw, W.C., and Gallager, C., (1992); Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate.
Cleft Palate and Craniofacial Journal,**28**,385-390

Ashley-Miller, M., (1977); Dental Research.
Health Bulletin,**35**,46-47

Bass, N.M., (1988); British Orthodontic Certification Board.
British Journal of Orthodontics,**15**,135-136

Berg, R., (1991); Evaluation of orthodontic results - a discussion of some methodological aspects.
The Angle Orthodontist,**61**,261-265

British Medical Association, (1989); Special report of the Council of the BMA on the Government's White Paper "Working for Patients",BMA SRM2

British Medical Association, (1989); Special report of the Council of the BMA on the Government's White Paper "Working for Patients"; 1989 SRM 2

British Orthodontic Standards Working Party, (1986); 2nd Report.
British Journal of Orthodontics,**12**,165-174

British Orthodontic Standards Working Party (1980); Working Party on British Orthodontic Standards.
British Journal of Orthodontics,**7**,103-107

British Standards Incisor Classification, (1983); 4492, British Standards Instituite, Glossary of Dental Terms.

Brook, P.H., & Shaw, W.C., (1989); The development of an index of orthodontic treatment priority.
European Journal of Orthodontics,**11**,309-320

Brown, W.A.B., Harkness, E.M., Cousins, A.J.P. & Isotupa, K., (1977); Treatment planning from study models: An examiner variability study.
The Angle Orthodontist,**47**,118-122

Carlos, J.P., (1970); Evaluation of indices of malocclusion.
International Dental Journal,**20**,606-617

Ciccheti, D.V., (1976); Assessing inter-rated reliability for rating scales; resolving some basic issues.
British Journal of Psychiatry,**129**,452-456

Cohen, J., (1960); A coefficient of agreement for nominal scales.
Educational and Psychological Measurement,**20**,37-46

Cons, N.C., (1986); DAI: The dental aesthetic index. Monograph. Health Quest. 1517 Derwen Dr. Iowa City

Court, S.D.M., (1976); Fit for the future-Report of the committee on Child Dental Health Services. HMSO

Devenish, E., (1978); Orthodontic treatment difficulty. MSc. thesis, University of Manchester

Donabedian, A., (1966); Evaluating the quality of medical care.
Millbank Memorial Fund Quarterly,**44**,166-204

Draker, H.L., (1960); Handicapping labio-lingual deviations: A proposed index for public health purposes.
American Journal of Orthodontics,**46**,295-305

Draper, N.R. & Smith, H., (1981); Applied Regression Analysis. 2nd Edition. Wiley. New York
ISBN 0-471-02995-5

Eismann, D., (1974); A method of evaluating the efficiency of orthodontic treatment.
Transactions of the European Orthodontic Society,223-232

Elderton, R.J. & Clark, J.D., (1984); Orthodontic treatment in the General Dental Service assessed by the Occlusal Index.
British Journal of Orthodontics,**11**,187-185

Evans, R. & Shaw, W.C., (1987); Preliminary evaluation of an illustrated scale for dental attractiveness.
European Journal of Orthodontics,**9**,314-318

Everitt, B.S., (1977); The analysis of contigency tables.
London:Chapman and Hall

Fink, D.F. & Smith, R.J., (1992); The duration of orthodontic treatment.
American Journal of Orthodontics and Dentofacial Orthopaedics,**102**,45-51

Fleiss, J.L., (1986); Design and analysis of clinical experiments.
Wiley, New York

Fletcher, G.G.H., (1958); The age factor in orthodontics. Transactions of the British Society for the Study of Orthodontics,31-38

Fox, N.A., (1993); The first 100 cases: A personal audit of orthodontic treatment assessed by the PAR(peer assessment rating) index.
British Dental Journal,**174**,290-297

Grainger, R.M., (1967); Orthodontic Treatment Priority Index.
PHS Publication No. 1000, Series 2 No. 25, Washington.
U.S. Government Printing Office.

Gravely, J.F. & Johnson, D.B., (1974); Angle's classification of malocclusion: An assessment of reliability.
British Journal of Orthodontics,**3**,79-86

Haynes, S., (1982); Discontinuation of Orthodontic Treatment in the General Dental Services in England and Wales 1972 to 1979.
British Dental Journal,**152**,127-129

Helm, S. & Petersen, P.E., (1989); Causal relation between malocclusion and caries.
Acta Odontologica Scandinavica,**47**,217-221

Helm, S., (1977); Intra-examiner reliability of epidemiological registrations of malocclusion.
Acta Odontologica Scandinavica,**53**,161-165

Hermanson, P.C. & Grewe, J.M., (1970); Examiner variability of several malocclusion indices.
The Angle Orthodontist,**40**,219-235

Hill, P.A., (1992); The prevalence and severity of malocclusion and the need for orthodontic treatment in 9- 12- and 15-year-old Glasgow Schoolchildren.
British Journal of Orthodontics,**19**,87-96

Hixon, E.H., Maschka, P.J. & Fleming, P.T., (1962); Occlusal status, caries and mastication.
Journal of Dental Research **41**,514-524

Holmes, A., (1992); The prevalence of treatment need.
British Journal of Orthodontics,**19**,177-182

Hooper, J.D., (1966); Orthodontics as a public service: The Wessex survey.
Transaction of the British Society for the Study of Orthodontics,1-10

Hughes, J. & Humphrey, C., (1990); Medical audit in general practice: a practical guide to the literature.
King Edward's Hospital Fund for London, London

Irvine, D. & Irvine, S., (1991); Making Sense of Audit.
Radcliffe Medical Press Ltd., Oxford
ISBN 1 870905 12 1

Isaacson, R.J., Christianson, R.L., Evance, C.A. & Reidel, R.A., (1975); Research on variation in dental occlusion: A state of the art workshop conducted by the Craniofacial Anomalies Program of the National Institute of Dental Research.
American Journal of Orthodontics,**68**,241-255

Isaacson, R.J., (1985); In response -- Oral health in the United States: prevalence of malocclusion.
Journal of Dental Education,**49**,397

Jago, J.D., (1974); The epidemiology of dental occlusion, a critical appraisal.
Journal of Public Dental Health,**34**,80-93

Jarvinen, S., (1979); Traumatic injuries to upper permanent incisors related to age and upper incisor overjet.
Acta Odontologica Scandinavica,**37**,335-338

Joint Working Party on Specialist Registration in Orthodontics, (1990); Joint report on specialist registration in orthodontics.
British Journal of Orthodontics,**17**,256-262

Katz, R. V., (1978); An epidemiologic study of the relationship between various states of the occlusion and the pathologic conditions of dental caries and periodontal disease.
Journal of Dental Research,**57**,433-439

Kenealy, P., Frude, N. & Shaw, W.C., (1989); An evaluation of the psychological and social effects of malocclusion: Some implications for dental policy making.
Social Science and Medicine,**6**,583-591

Kerr, W.J.S., (1984); Appliance breakages.
British Journal of Orthodontics,**11**,137-142

Kerr, W.J.S., Buchanan, I.B. & McColl, J.H., (1993); Use of the PAR index in assessing the effectiveness of removable orthodontic appliances.
British Journal of Orthodontics,**20**,351-357

Kerr, W.J.S., Buchanan, I.B., McNair, F. & McColl, J.H., (1993); Factors influencing the outcome and duration of removable appliance treatment.
European Journal of Orthodontics, (In press)

Landis, J.R. & Koch, G.G., (1977); The measurement of observer agreement for categorical data.
Biometrics,**33**,159-174

Lau, D., Griffiths, G. & Shaw, W.C., (1984); Reproducibility of an index for recording the alignment of individual teeth.
British Journal of Orthodontics,**11**,80-84

Linder-Aronson, S., (1974); Orthodontics in the Swedish Public Dental Health System.
Transactions of the European Orthodontic Society, 1974,233-240

Little, R.M., (1975); The irregularity index: A quantative score of lower anterior alignment.
American Journal of Orthodontics,**68**,554-563

Luffingham, J.K. & Campbell, H.M., (1974); The need for orthodontic treatment. A pilot study of 14 year old school children in Paisley, Scotland.
Transactions of the European Orthodontic Society,259-267

McGuinness, N.J. & Stephens, C.D., (1993); Holograms and study models assessed by the PAR (Peer Assessment Rating) index of malocclusion--A pilot study.
British Journal of Orthodontics,**20**,123-129

McLain, J.B. & Proffit, W.R., (1985); Oral health status in the United States: Prevalence of malocclusion.
Journal of Dental Education,**49**,386-396

Miller, J. & Hobson, P., (1961); The relationship between malocclusion, oral cleanliness, gingival conditions and dental caries in schoolchildren.
British Dental Journal,**111**,43-52

Mills, J.R.E., (1978); The effect of orthodontic treatment on the skeletal pattern.
British Journal of Orthodontics,**5**,133-143

Mohlin, B. & Thilander, B., (1984); The importance of the relationship between malocclusion and mandibular dysfunction and some clinical applications in adults.
European Journal of Orthodontics,**6**,192-204

Mouatt, B. & Heap, D., (1991); DHSS GDS communication 646m/13-8/VG/1

Murray, J.J., (1969); Prevalence of malocclusion in fifteen year old children from fluoride and non-fluoride communities.
British Dental Journal,**127**,128-131

Myrberg, N. & Thilander. B., (1973); An evaluation of the duration and the results of orthodontic treatment.
Scandinavian Journal of Dental Research,**81**,85-91

Norman, R.G. & Streiner, D.L., (1986); PDQ Statistics
B.C. Decker, Toronto

Norusis, M.J., (1993); SPSS for Windows Base System User's Guide.
SPSS Inc.,Chicago,ISBN 0-13-178856-6

O'Brien, K.D. & Corkhill, C., (1990); Regional variation in the provision of orthodontic treatment by the hospital services in England and Wales.
British Journal Of Orthodontics,**17**,187-195

O'Brien, K.D., Shaw, W.C. & Roberts, C.T., (1993); The use of Occlusal indices in assessing the provision of orthodontic treatment by the hospital orthodontic service of England and Wales.
British Journal of Orthodontics,**20**,25-35

Ovretveit, J., (1990); What is quality in health services?
Health Services Management,June 1990,132-133

Parkhouse, R.C., (1987); Orthodontics in the wake of Schanschieff.
British Dental Journal,**162**,95-96

Pender, N., (1982); Recent graduate opinion on orthodontic training.
British Journal of Orthodontics,**9**,73-76

Pickering, E.A. & Vig, P., (1976); The Occlusal Index to assess orthodontic treatment.
Briish Journal of Orthodontics,**2**,47-51

Prahl-Andersen, B., (1978); The need for orthodontic treatment.
The Angle Orthodontist,**48**,1-9

Rendell, J.K., Norton, L.A. & Gay, T., (1992); Orthodontic treatment and temporomandibular joint disorders.
American Journal of Orthodontics and Dentofacial Orthopaedics,**101**,84-87

Report of the Committee into Unecessary Dental Treatment.
Department of Health and Social Security
HMSO. ISBN 0 11 321051 5

Report of the Occlusal Index Committee, (1987).
Department of Health and Social Security
HMSO. 0289d/13/73-26

Richmond, S., (1990); A critical evaluation of orthodontic treatment standards in the General Dental Service in England and Wales.
PhD thesis, University of Manchester

Richmond, S., Shaw, W.C., O'Brien, K.D., Buchanan, I.B., Jones, R., Stephens, C.D., Roberts, C.T. & Andrews, M., (1992a); The development of the PAR index (Peer Assessment Rating): Reliability and validity.
European Journal of Orthodontics,**14**,125-139

Richmond, S., Shaw, W.C., Roberts, C.T. & Andrews, M., (1992b); The PAR index (PAR): Methods to determine outcome of orthodontic treatment in terms of improvement and standards.
European Journal of Orthodontics,**14**,180-187

Richmond, S., (1993); Personal audit in orthodontics.
British Journal of Orthodontics,**20**,135-145

Richmond, S., Shaw, W.C., Stephens, C.D, Webb, W.G., Roberts, C.T. & Andrews, M., (1993); Orthodontics in the General Dental Service of England and Wales: A critical assessment of standards.
British Dental Journal,**174**,315-329

Richmond, S., O'Brien, K.D., Roberts, C.T. & Andrews, M., (1994); Dentists' variation in the determination of orthodontic treatment need.
British Journal of Orthodontics,**21**,65-68

Ringenberg, Q.M., (1967); Influence of serial extraction on growth and development of the maxilla and mandible.
American Journal of Orthodontics,**53**,19-26

Robertson, N.R.E. & Hoyle, B.A., (1983); Orthodontic treatment: Time for a change?
British Journal of Orthodontics,**10**,154-156

Rose, J.S., (1974); A thousand consecutive treated orthodontic cases - a survey.
British Journal of Orthodontics,**1**,45-54

Rosenthal, R. & Jacobsen, C.F., (1968); Pygmalion in the classroom; Teachers' expectations and pupils' intellectual development.
Holt, Rinehart & Winston. New York. 3-31

Salvia, J., Algozine, R. & Sheare, J.B., (1977); Attractiveness and school achievement.
Journal of School Psychology,**15**,60-67

Salzmann, J.A., (1967); Malocclusion severity estimate.
American Journal of Orthodontics,**53**,109-119

Salzmann, J.A., (1968); Handicapping malocclusion assessment to establish treatment priority.
American Journal of Orthodontics,**54**,749-765

Sandli, T., (1973); Irregularities of the teeth and their relation to the periodontal condition with particular reference to the lower labial segment.
Transactions of the European Orthodontic Society, 1973,319-333

Sergl, H.G., (1977); Periodontal prophylaxis assessing need for orthodontic treatment.
Transactions of the European Orthodontic Society,131-139

Sergl, H.G. & Stodt, W., (1970); Experimental investigation of the aesthetic effect of various tooth positions after loss of an incisor tooth.
Transactions of the European Orthodontic Society,497-507

Shaw, C., & Costain, D.W., (1989); Guidelines for medical audit: seven principles.
British Medical Journal,**299**,498-499

Shaw, W.C., Addy, M. & Ray, C., (1980); Dental and social effects of malocclusion and effectiveness of orthodontic treatment: a review.
Community Dentistry and Oral Epidemiology,**8**,36-45

Shaw, W.C., Rees, G., Dawe, M. & Charles, C.R., (1985); The influence of dentofacial appearance on the social attractiveness of young adults.
American Journal of Orthodontics,**85**,21-26

Shaw, W.C., O'Brien, K.D., Richmond, S. & Brook, P., (1991); Quality control in orthodontics: Risk/Benefit considerations.
British Dental Journal,**170**,33-37

Sheiham, A., Hobdell, M.H., Vig, P. & Griffiths, P.J., (1971); Orthodontic treatment in the General Dental Services in England and Wales: a study.
British Dental Journal,**131**,535-538

So, L.L.Y. & Tang, E.L.K., (1993); A comparative study using the Occlusal Index and the Index of Orthodontic Treatment Need.
The Angle Orthodontist,**63**,57-64

Stephens, C.D. & Harradine, N.W., (1988); Changes in the complexity of orthodontic treatment for patients referred to a teaching hospital.
British Journal of Orthodontics,**15**,27-32

Stephens, C.D., Orton, H.S. & Usiskin, L.A., (1985); Future manpower requirements for orthodontics undertaken in the General Dental Services.
British Journal of Orthodontics,**12**,168-175

Stratford, N.M., (1991); Medical Audit - In Relation to Orthodontics.
British Journal of Orthodontics,**18,**143-145

Summers, C.J., (1966); The occlusal index: A system for identifying and scoring occlusal disorders.
American Journal of Orthodontics,**59**,552-567

Tang, E.L.K. & Wei, S.H.Y., (1990); Assessing treatment effectiveness of removable and fixed orthodontic appliances with the Occlusal Index.
American Journal of Orthodontics and Dentofacial Orthopaedics,**99**,550-556

Thilander, B., (1984); A textbook of clinical periodontology. Copenhagen; Munskaard, pp480-500

Todd, J.E. & Dodd, T., (1985);
Childrens' Dental Health in the United Kingdom.
London, HMSO

Todd, J. E. & Dodd, T., (1985); Children's dental health in the United Kingdom 1983.
HSMO ISBN 0 11 691136 0

Turner, S.A.M., (1990); Occlusal indices revisited.
British Journal of Orthodontics,**17**,197-203

Turner, S.A.M., (1983); The feasibility and validity of orthodontic screening of children in their tenth year.
British Journal of Orthodontics,**10**,142-146

Vig, P.S., Weintraub, J.A., Brown, C. & Kowalski, C.J., (1990); The duration of orthodontic treatment with and without extractions: A pilot study of five selected practices.
American Journal of Orthodontics and Dentofacial Orthopaedics,**97**,45-51

Williams, A.C. & Stephens, C.D., (1992); A modification to the incisor classification of malocclusion.
British Journal of Orthodontics,**19**,127-130

Woolass, K.F. & Shaw, W.C., (1987); Validity and reproducibility of rating dental attractiveness from study casts.
British Journal of Orthodontics,**14**,187-190

Working for Patients, (1989). HMSO
ISBN 0 11 493476 2

World Health Organisation, (1987); Oral health surveys: Basic methods. 3rd Edition. Geneva: World Health Organisation

Appendix I

Index of Orthodontic Treatment Need

Criteria for the Dental Health Component

Grade 1 (None)

1 Extremely minor malocclusions including displacements less than 1mm.

Grade 2 (Little)

2.a Increased overjet greater than 3.5mm but less than or equal to 6mm with competent lips.
2.b Reverse overjet greater than 0mm but less than or equal to 1mm.
2.c Anterior or posterior crossbite with less than or equal to 1mm discrepancy between retruded contact position and intercuspal position.
2.d Displacement of teeth greater that 1mm but less than or equal to 2mm.
2.e Anterior or posterior openbite greater than 1mm but less than or equal to 2mm.
2.f Increased overbite greater than or equal to 3.5mm without gingival contact.
2.g Prenormal or postnormal occlusions with no other anomalies. Includes up to half a unit discrepancy.

Grade 3 (Moderate)

3.a Increased overjet greater than 3.5mm but less than or equal to 6mm with incompetent lips.
3.b Reverse overjet greater than 1mm but less than or equal to 3.5mm.
3.c Anterior or posterior crossbites with greater than 1mm bur less than or equal to 2mm discrepancy between retruded contact position and intercuspal position.
3.d Displacement of teeth greater than 2mm but less than or equal to 4mm.
3.e Lateral or anterior open bite greater than 2mm but less than or equal to 4mm.
3.f Increased and complete overbite without gingival or palatal trauma.

Grade 4 (Great)

4.a Increased overjet greater than 6mm but less than or equal to 9mm.
4.b Reverse overjet greater than 3.5mm with no masticatory or speech difficulties.
4.c Anterior or posterior crossbites with greater than 2mm discrepancy between retruded contact position and intercuspal position.
4.d Severe displacements of teeth greater than 4mm.
4.e Extreme lateral or anterior open bites greater than 4mm.
4.f Increased and complete overbite with palatal or gingival trauma.
4.h Less extensive hypodontia requiring pre-restorative orthodontics or orthodontic space closure to obviate the need for a prosthesis.
4.1 Posterior lingual crossbite with no functional occlusal contact in one or both buccal segments.
4.m Reverse overjet greater than 1mm but less than 3.5mm with recorded masticatory and speech difficulties.
4.t Partially erupted teeth, tipped and impacted against adjacent teeth.
4.x Presence of supernumerary teeth.

Grade 5 -- Very great

5.a Increased overjet greater than 9mm.

5.h Extensive hypodontia with restorative implications (more than 1 tooth missing in any quadrant) requiring pre-restorative orthodontics.

5.i Impeded eruption of teeth (except third molars) due to crowding, displacement, the presence of supernumerary teeth, retained deciduous teeth and any pathological cause.

5.m Reverse overjet greater than 3.5mm with reported masticatory and speech difficulties.

5.p Defects of cleft lip and palate and other cranio-facial anomalies.

5.s Submerged deciduous teeth.

Certain alterations to the DHC clinical grading are made when DHC is applied to study casts as clinical gradings are dependent on soft tissue and functional information which is not available from study casts.

Clinical grade	Study casts grade
2.a Overjets 3.5mm - 6mm (competent lips)	3.a Overjets 3.5mm - 6mm (incompetent lips)
2.c and 3.c Crossbites (displacement =<2mm)	4.c Crossbites (displacement =>2mm)
4.b Reverse overjets >3.5mm (no masticatory or speech problems)	5.m Reverse overjets >3.5mm (masticatory or speech problems)

A transparent acetate ruler has been designed to facilitate measuring allowing viewing of contact points through the ruler. The ruler also provides summary reference information concerning the index. Below is an example of an IOTN ruler.

Aesthetic Component

The set of ten monochrome photographs used for assessing study casts and representing a ten-point scale is illustrated in Figure 55. For study casts the use of monochrome photographs removes the potentially confusing effect of gingival colour and staining of restorations.

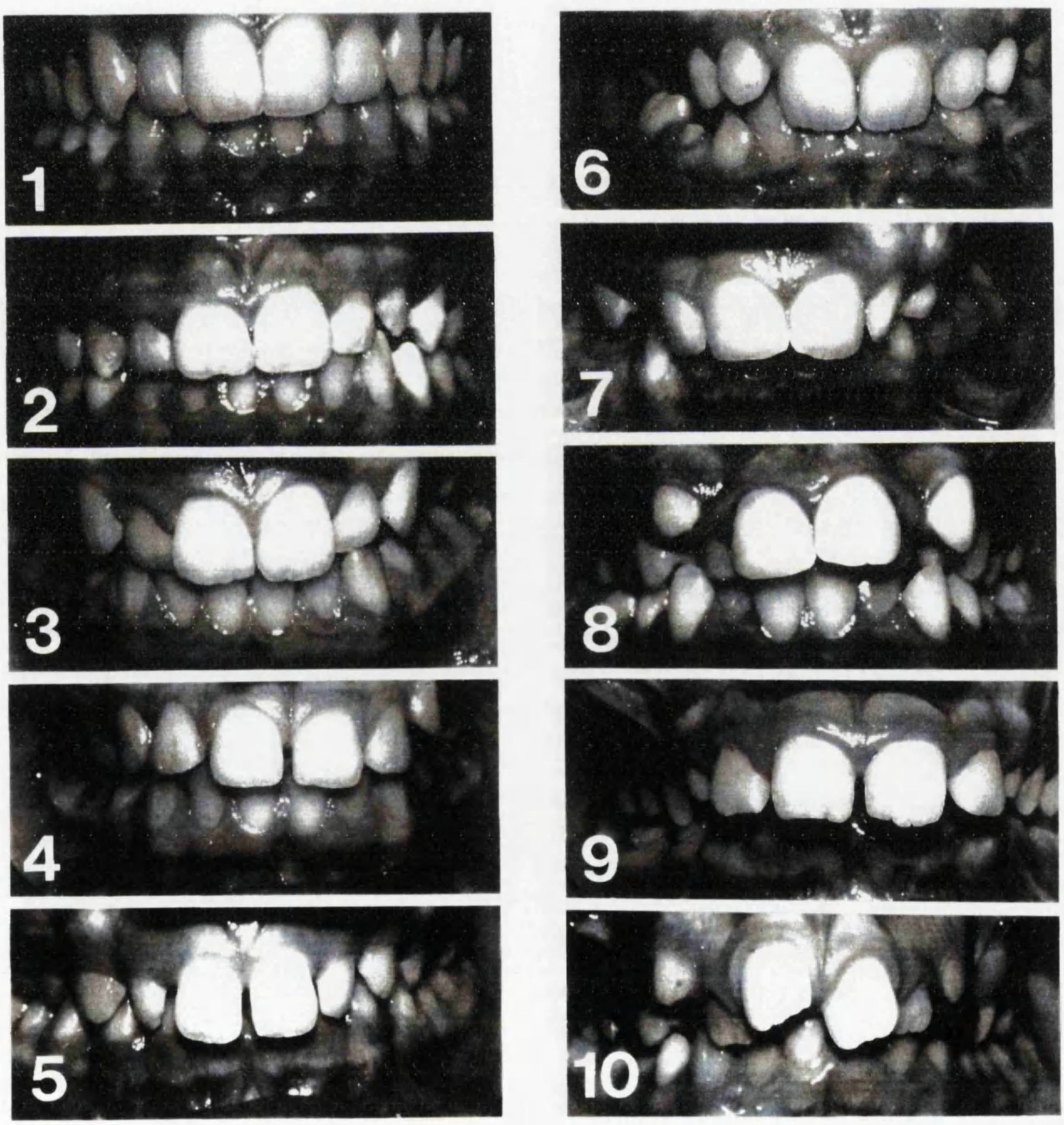


Figure 55

Appendix II

Peer Assessment Rating

Scoring scheme within components

1 Upper and lower anterior segments

Recorded between mesial of the canine on one side to the mesial of the canine on the other. The shortest distances parallel to the occlusal plane between the contact points of the six anterior teeth are measured.

Discrepancy	Score
0 mm - 1 mm	0
1.1 mm - 2.0 mm	1
2.1 mm - 4.0 mm	2
4.1 mm - 8.0 mm	3
> 8.0 mm	4
impacted teeth	5

2 Left and right buccal segments

Recorded from the canine to the last molar. Antero-posterior, transverse and vertical discrepancy scores are summed for each buccal segment.

	Discrepancy	Score
a) Antero-posterior		
	Good intercuspation (Cl 1, II and III)	0
	Less than ½ unit	1
	Cusp to Cusp	2

b) Transverse

No discrepancy	0
Crossbite tendency	1
Single tooth in crossbite	2
> 1 tooth in crossbite	3
> 1 tooth in scissors bite	4

c) Vertical

No discrepancy	0
Lateral open bite>2mm (at least two teeth)	1

3 Overjet

a) Overjet

0 - 3 mm	0
3.1 mm - 5.0 mm	1
5.1 mm - 7.0 mm	2
7.1 mm - 9.0 mm	3
> 9.0 mm	4

b) Anterior crossbite

No discrepancy	0
Edge to edge	1
Single tooth in crossbite	2
Two teeth in crossbite	3
> two teeth in crossbite	4

4 Overbite

a) Overbite

<= 1/3 coverage of lower incisor	0
> 1/3 coverage	1
> 1/3 coverage	2
=> full coverage	3

b) Open bite

No open bite	0
=< 1 mm	1
1.1 mm - 2.0 mm	2
2.1 mm - 3 mm	3
> 4 mm	4

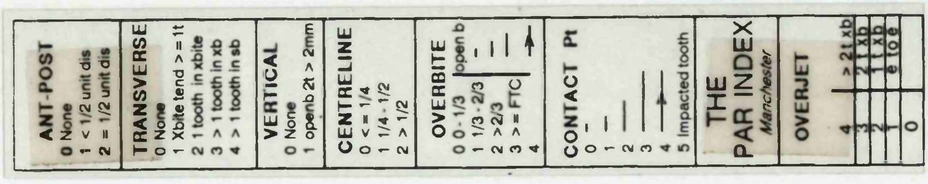
5 Centrelines

Coincident, up to ¼ width of lower incisor	0
¼ - ½ width lower incisor	1
> ½ width of lower incisor	2

The scores within each component are summed, the appropriate weighting for each component applied, and the weighted component scores summed to a total PAR score. Weightings for each component are:

Overjet	6
Centrelines	4
Overbite	2
Right buccal occlusion	1
Left buccal occlusion	1
Upper anterior segment	1
Lower anterior segment	1

A transparent acetate ruler has been designed to facilitate measuring allowing viewing of contact points through the ruler. Below is an example of a PAR ruler.



Appendix III

Calibration exercise

Examiner calibration

Calibration was necessary to ensure that the criteria and definitions of an index were being applied correctly in accordance with the formulation of the indices. Calibration also minimises variation between examiners.

Calibration of the examiner in this study took place in December 1991 at the Occlusal Index Course organised by the University Dental Hospital of Manchester and under the direction of Dr. S. Richmond.

The calibration exercises for PAR and IOTN required the examination of 30 cases and the comparison of index ratings for these cases against the "Gold Standard" scores and rankings agreed by a panel of 74 experienced clinicians during validation of the indices.

Calibration for IOTN

The differences in scores between this examiner and the "Gold Standard" ratings of 30 cases for the Dental Health Component of IOTN are shown in the crosstable.

Crosstable

		Standard				
	IOTN grade	1	2	3	4	5
	1	0	0	0	0	0
	2	0	9	3	0	0
Examiner	3	0	0	1	0	0
	4	0	0	0	13	0
	5	0	0	0	1	3

Values are given for unweighted Kappa *i.e.* all disagreements are equally serious, and weighted Kappa *i.e.* when disagreements occur, the closer the grade given by the examiner is to the standard grade the less serious the discrepancy is.

	Weighted	Unweighted
Kappa	0.89	0.80
Lower 95% confidence limit	0.78	0.65

Both unweighted and weighted kappa were significantly better than chance (P=<0.05) and not significantly different from a value of 0.8 (P>0.05).

The equivalent crosstable for the Aesthetic component of IOTN was:

					Standard					
AC	1	2	3	4	5	6	7	8	9	10
Examiner										
1	1	1	0	0	0	0	0	0	0	0
2	1	3	3	0	0	0	0	0	0	0
3	0	0	1	0	0	1	0	0	0	0
4	0	0	2	1	1	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	2	1	0	0
7	0	0	0	0	1	0	1	0	0	1
8	0	0	0	0	0	0	0	0	2	0
9	0	0	0	0	0	0	0	3	2	0
10	0	0	0	0	0	0	0	0	0	1

	Weighted	Unweighted
Kappa	0.75	0.34
Lower 95% confidence limit	0.67	0.18

For the Aesthetic Component, unweighted Kappa was significantly better than chance (P=<0.05), but was significantly different from a value of 0.8 (P=<0.05).

Weighted Kappa however was significantly better than chance (P=<0.05) and not significantly different from 0.8 (P>0.05). The difference in significance between weighted and unweighted Kappa is not surprising in view of the ordinal scale and progressive nature of the subjectively scored Aesthetic Component.

The ten Aesthetic Component rating groups were condensed into three groups indicating no aesthetic need for treatment (groups 1-4), borderline requirement for treatment (groups 5-7) and definite need for treatment (groups 8-10).

The difference between the ratings of this examiner and the "Gold Standard" when these three groupings are employed are indicated in the crosstable.

Aesthetic Component-3 groups		Standard		
		1	2	3
Examiner	1	13	2	0
	2	0	5	2
	3	0	0	8

	Weighted	Unweighted
Kappa	0.75	0.34
Lower 95% confidence limit	0.67	0.18

Both unweighted and weighted kappa were significantly better than chance (P=<0.05) and not significantly different from a value of 0.8 (P>0.05).

The presence of bias was investigated by the Wilcoxon signed-ranks test. No bias was demonstrated for Dental

Health Component (P=0.3613), Aesthetic Component-10 groups (P=0.4460) or Aesthetic Component-3 groups (P=0.0679) although there was a tendency to underscore for Aesthetic Component.

The examiner in this study was therefore considered calibrated for IOTN.

Calibration for PAR

The actual differences between the "Gold Standard" total PAR scores and those of the examiner in this study and the frequency of their occurrence are shown in Figure 56

A plot of the difference between "Gold Standard" PAR score and this examiners PAR scores against "Gold Standard" scores is shown in Figure 57. Visual examination does not indicate any clear association between error and increasing PAR score.

A simple measure of error in PAR scoring is given by the Root Mean Squared error (RMS). The smaller the value of RMS the more reliable the scoring. In the absence of any significant bias, an examiner with a RMS error of 5 or fewer PAR points was considered to be calibrated for PAR.

RMS error for the examiner in this study = 3.2 PAR points.

Total PAR scores were considered to have a normal distribution allowing the presence of systematic bias to be tested by a paired-t test and the agreement between "Gold Standard" PAR scores and examiner PAR scores to be assessed with the Intraclass correlation coefficient of reliability, (Fleiss, 1986).

CALIBRATION EXERCISE: FREQUENCY OF ERRORS IN PAR SCORING

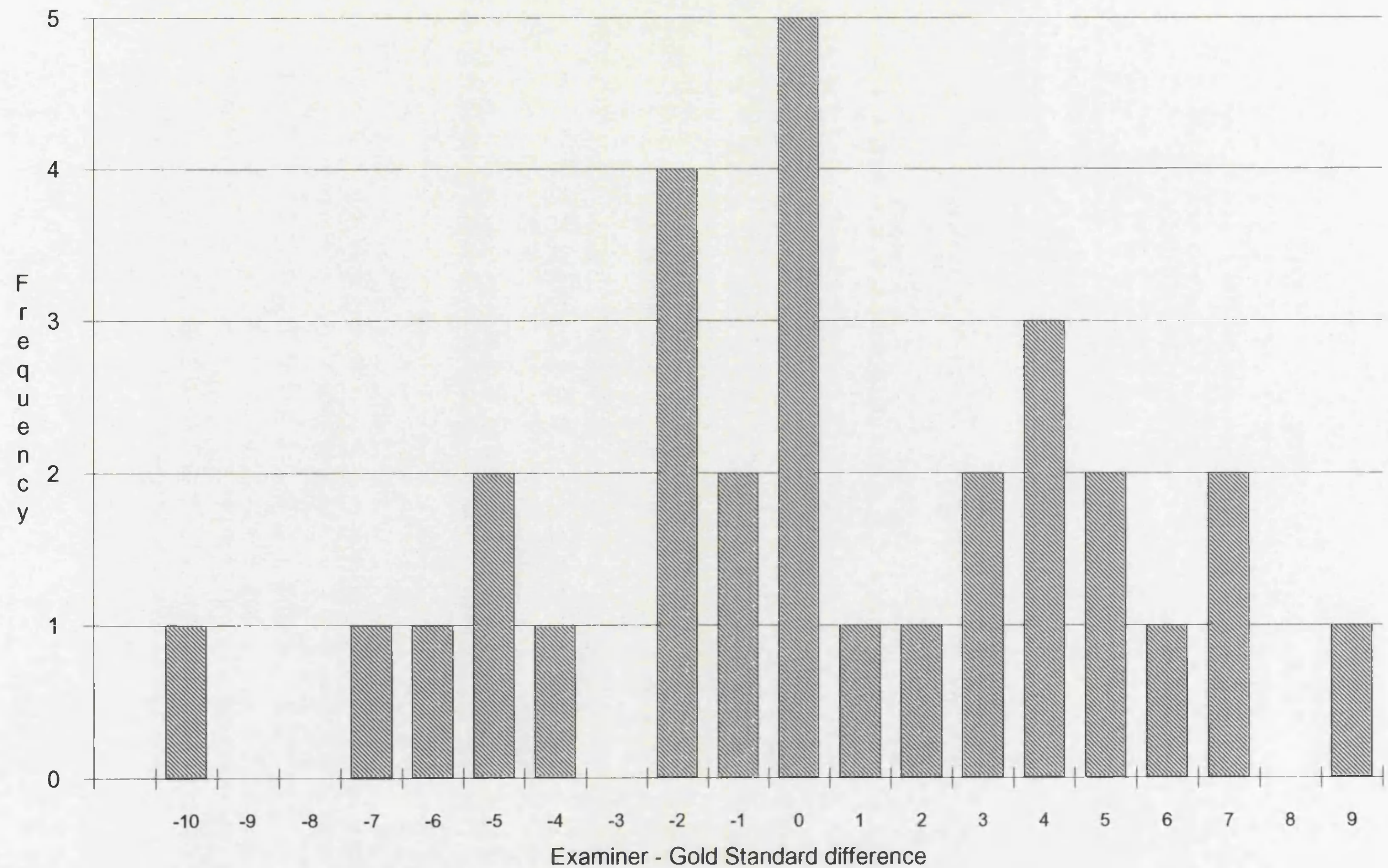


Figure 56

CALIBRATION EXERCISE: ERRORS IN PAR SCORE AGAINST GOLD STANDARD

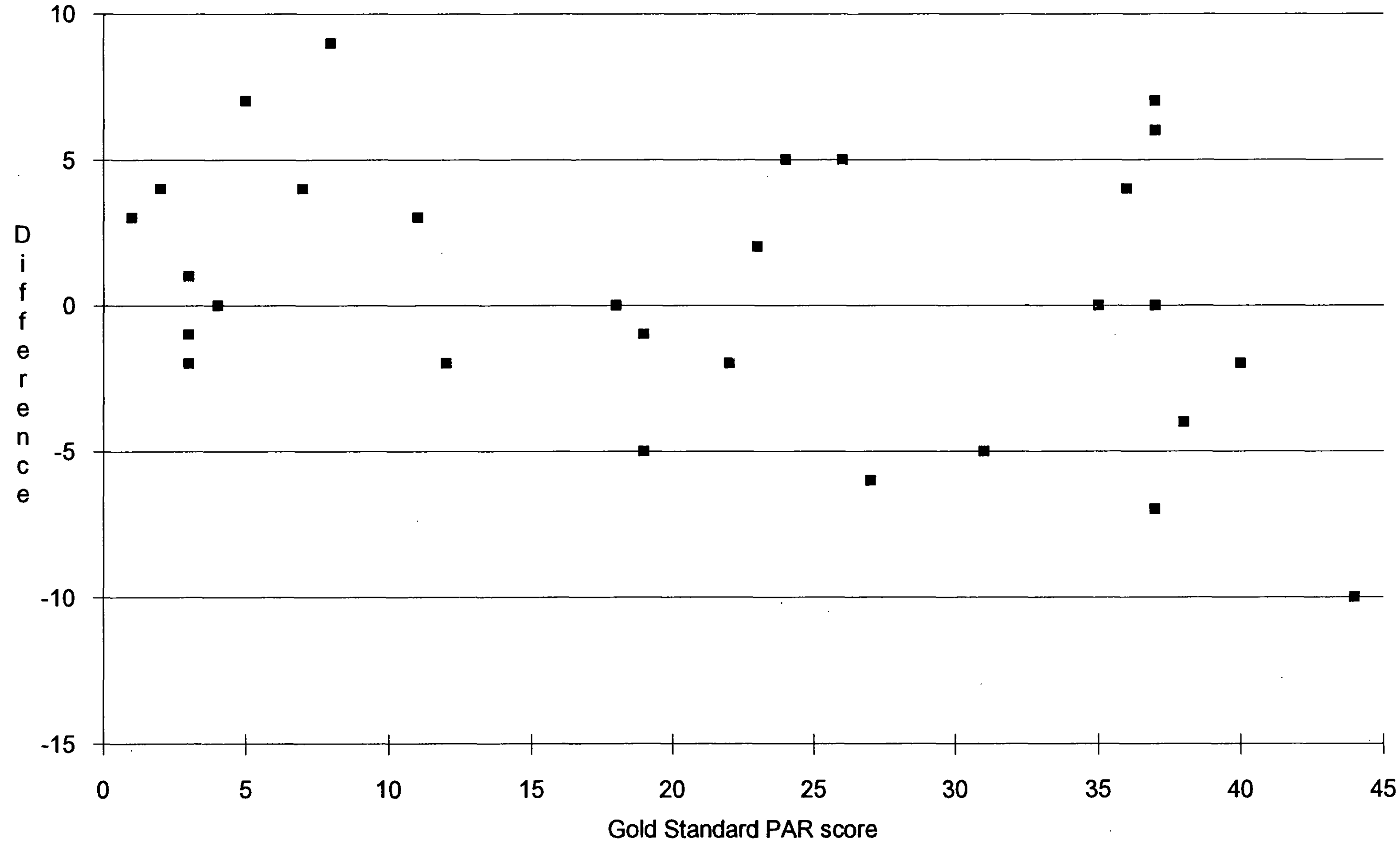


Figure 57

A t-test probability of less than P=0.05 would have indicated that differences were not due to chance and that bias was present. For this examiner bias was not demonstrated (P=0.6).

The Intraclass Correlation Coefficient of Reliability (R) may be directly interpreted as the proportion of the variance of an observation which is due to subject variability. The other contributor to variance is variability amongst raters which includes variability due to random errors. The less the variance due to inter or intra examiner error the closer R will be to 1.

The lower 95% confidence limit for R indicates the limits of uncertainty concerning the degree of reliability present in an examiner's data. If the lower 95% confidence limit is greater than or equal to 0.75, single scores can be relied upon with confidence.

For this examiner R=0.95 whilst the lower 95% confidence limit=0.92.

The examiner in this study was therefore considered calibrated for PAR.

ERRATA

The following works are referred to in the text but do not appear in the bibliography:

Backlund, E., (1963); Tooth form and overbite.
Transactions of the European Orthodontic Society,**36**,97-103

Ballard, C.F. & Wayman, J.B., (1964); Report on a survey of the orthodontic requirements of 310 Army apprentices.
Dental Practitioner,**15**,221-226

Berg, R., (1979); Post-retention analysis of treatment problems and failures in 264 consecutively treated cases.
European Journal of Orthodontics,**1**,55-68

Bjork, A., Krebs, A. & Solow, B., (1964); A method for epidemiological registration of malocclusion.
Acta Odontologica Scandinavica,**20**,27-40

Grewe, J.M. & Hagan, D.V., (1972); Malocclusion indices: A comparative evaluation.
American Journal of Orthodontics,**61**,286-294

O'Brien, K.D.; Presentation given to the British Orthodontic Conference, Manchester, (1992)

In addition:

Page 21 line 23 should read "ensuring the maintenance of the standards taught during training."

The reference to Vig *et al.* on page 26 should be "Vig, Weintraub, Brown and Kowalski (1990)".

The reference to Isaacson (1975) on page 34 should be "Isaacson *et al.*, (1975)".