CHAPTER 4: RESULTS

4.1 DESCRIPTION OF SAMPLE

Although every attempt was made to contact all 120 patients qualifying for inclusion in this study, only 31 patients were able to present themselves for record taking. This group thus constituted the sample which comprised 14 males (45.2 %) and 17 females (54.8 %) (Table 4.1.1). The mean pretreatment (T1) and post-treatment (T2) ages for the sample was 13.2 years and 15.2 years respectively, while the mean postretention (T3) age for the same sample was 28.1 years. The mean duration of treatment was 24.4 months, with a range of 14 to 25 months, and the mean postretention (T3) period was just under 13 years (154.7 months) with a range of 10 to 13.25 years (159 months).

Twenty-four patients (77.4 %) had had teeth extracted as part of their treatment, while 7 patients (22.6 %) had been treated without extraction. Maxillary Hawley retainers and mandibular canine to canine fixed lingual retainers were used as retention devices in all patients. Edgewise mechanics with full banding was the treatment regimen for the entire sample.

Table 4.1.1 Gender by class (Angle).

GENDER	CLASS			TOTAL
	I	II DIVISION I	II DIVISION II	
FEMALE	3	13	1	17
MALE	2	11	1	14
TOTAL	5	24	2	31

When considering subdivision of the sample by both class and gender, it was noted that most of the subjects were categorised as Angle Class II division I, as seen in Table 4.1.1 . As a result, any correlation coefficients calculated by gender alone would be strongly influenced by Class II division I patients. Thus, when examining the findings for gender alone and the findings for gender and class simultaneously, similar outcomes were to be expected although with larger p-values in the latter case (due to the smaller sample sizes). Due to the small sample sizes in Class I and Class II division II, it was difficult to assess their effect on the findings for gender alone and in particular, it was difficult to detect spurious correlations when looking at gender alone that may have been induced by class differences. From Table 4.1.1, it was also seen that no Class III malocclusions were encountered in the sample group.

4.2 DATA ANALYSIS

When the individual malocclusion scores for each variable for each patient in the sample were added together, a total malocclusion score for each patient was obtained. These total malocclusion scores for each patient were then used to compute a mean score for the sample, for each stage of treatment, viz, pretreatment (T1), post-treatment (T2) and postretention (T3). These mean scores are shown in Table 4.2.1.

Table 4.2.1 Mean malocclusion scores at the three stages of treatment.

	PRETREATMENT (T1)	POST-TREATMENT (T2)	POSTRETENTION (T3)
MEAN	8.8	0.3	1.5
SD	3.6	1.0	1.7
MINIMUM	2.0	0	0
MAXIMUM	16.0	4.0	7.0
MEDIAN	9.0	0	2.0
RANGE	14	4.0	7.0
N	31	31	31

The mean malocclusion score for the sample at T1 was calculated to be 8.8 while the mean malocclusion score at T2 was found to be 0.3 (Table 4.2.1). This dramatic change in the score from T1 to T2 suggested a meaningful improvement in those variables as a result of the treatment. The mean score at T3, however, was found to be 1.5 (Table 4.2.1) which indicated a marginal deterioration in the scores obtained at T2. Using the Student's paired t-test (Goulden, 1956), the changes in the scores between the three stages of treatment were found to be highly significant ($p \leq 0.05$) (Table 4.2.2).

Table 4.2.2 Mean scores paired t-test results.

		N	p-value
All Classes:	T1 to T2	31	<.0001 *
	T2 to T3	31	<.0001 *
	T1 to T3	31	<.0001 *
Class I:	T1 to T2	5	0.06 **
	T2 to T3	5	1.00
	T1 to T3	5	0.06 **
Class II division I :	T1 to T2	24	<.0001 *
	T2 to T3	24	<.0001 *
	T1 to T3	24	<.0001 *

* Significant at the 5% level. ** Significant at the 10% level.

The distribution of the scores of each variable at the three stages of treatment is shown in Table 4.2.3.1 and 4.2.3.2. Compared with the scores at T1, marked improvements were noted in all the dental parameters studied at T2, and T3. Non-ideal relationships, as indicated by scores greater than zero, were recorded for some variables at both the T2 and T3 stages of treatment.

At T2, two patients exhibited Angle Class II molar relationships, while one patient presented with a mild overjet (Table 4.2.3.1). These non-ideal relationships constituted the mean T2 score of 0.3. This indicated that not all of the variables studied were treated to an ideal relationship as defined in this study and in that of Sadowsky and Sakols, (1982).

The malocclusion scores noted at T3 again showed evidence of non-ideal occlusal relationships (Tables 4.2.3.1 and 4.2.3.2). There was an increase in the incidence of occlusal abnormalities when the T2 and the T3 stages of treatment were compared.

Tables 4.2.3.1 and 4.2.3.2 recorded the values for each variable from T2 to T3. It was evident that certain ideal occlusal relationships observed at T2 deteriorated with the passage of time. This was particularly true for features such as overjet and overbite.

Tables of p-values for both paired t-tests (Goulden, 1956), and Wilcoxon signed rank tests (Wilcoxon, 1945) are recorded in Appendix A for the total sample i.e. Class I and Class II division I. These tests were performed to determine whether a statistically significant difference in the median total score between different stages of treatment existed. These tests were performed from T1 to both T2 and T3, as well as from T2 to T3. A summary of the p-values determined is recorded in Table III. In the total sample as well as in the Class II division I subgroup, a statistically significant difference existed in the mean total score from zero at the 5 percent significance level ($p \leq 0.05$) for all stages. The difference in the median T1 to T2 and T1 to T3 scores in the Class I sub-group were found to be only marginally significant at the 5 percent level, and statistically significant at the 10 percent level.

Table 4.2.3.1 Scores and distribution of the observations of the variables at the three stages of treatment (T1, T2, T3).

VARIABLE	S	T1		T2		T3	
		No	%	No	%	No	%
<u>MOLAR-LEFT +RIGHT</u>							
IDEAL - CLASS I	0	33	53.2	58	93.5	57	91.9
MODERATE - CLASS II	2	23	37.1	4	6.5	5	8.1
SEVERE - CLASS III	3	6	9.7	0	0	0	0
<u>CANINE- LEFT + RIGHT</u>							
IDEAL - CLASS I	0	37	59.7	62	100	62	100
MODERATE - CLASS II	2	21	33.9	0	0	0	0
SEVERE - CLASS III	3	4	6.5	0	0	0	0
<u>OVERJET</u>							
IDEAL - 0.0 - 3.0	0	7	2.6	30	96.8	24	77.4
MILD - 3.5 - 6.0	2	17	54.8	1	3.2	7	22.6
MODERATE - 6.5 - 9.0	3	6	19.4	0	0	0	0
SEVERE - 9.5 +	4	1	3.2	0	0	0	0
<u>OVERBITE</u>							
IDEAL - 0.0 - 3.0	0	12	38.7	31	100	22	71
MODERATE - 3.5 - 5.0	2	13	41.9	0	0	9	29
SEVERE - 5.5 +	3	6	19.4	0	0	0	0
<u>OPEN BITE</u>							
IDEAL - NONE	0	28	90.4	31	100	30	96.8
MILD WITH OVERBITE	1	1	3.2	0	0	0	0
MODERATE - 0.0 - 3.0	2	1	3.2	0	0	1	3.2
SEVERE - 3.0 +	3	1	3.2	0	0	0	0
<u>ANTERIOR CROSS BITE</u>							
IDEAL - NONE	0	28	90.3	31	100	30	96.8
MODERATE 1-2 TEETH	1	3	9.7	0	0	1	3.2
SEVERE 3 - 4 TEETH	2	0	0	0	0	0	0

(S = Score)

Table 4.2.3.2 Scores and distribution of the observations of the variables at the three stages of treatment (T1, T2, T3).

VARIABLE	S	T1		T2		T3	
		No	%	No	%	No	%
POSTERIOR CROSSBITE							
IDEAL - NONE	0	27	87.1	31	100	30	96.8
MODERATE - UNI-LATERAL	1	4	12.9	0	100	1	3.2
SEVERE - BILATERAL	2	0	0	0	0	0	0
MAX. ANTERIOR CROWDING							
IDEAL - 0.0 - 3.0	0	14	45.2	31	100	31	100
MODERATE - 3.5 - 6.0	1	12	38.7	0	0	0	0
SEVERE - 6.5 +	2	5	16.1	0	0	0	0
MANDIB. ANTERIOR CROWDING							
IDEAL - 0.0 - 3.0	0	17	54.9	31	100	30	96.8
MODERATE - 3.5 - 6.0	1	9	29.0	0	0	1	3 .2
SEVERE - 6.5 +	2	5	16.1	0	0	0	0
LOWER INTERCANINE WIDTH (mm)		28.6		29.8		29.9	
UPPER INTERMOLAR WIDTH (mm)		32.6		32.4		32.0	
TOTAL MALOCCLUSION SCORE		8.8		0.3		1.5	

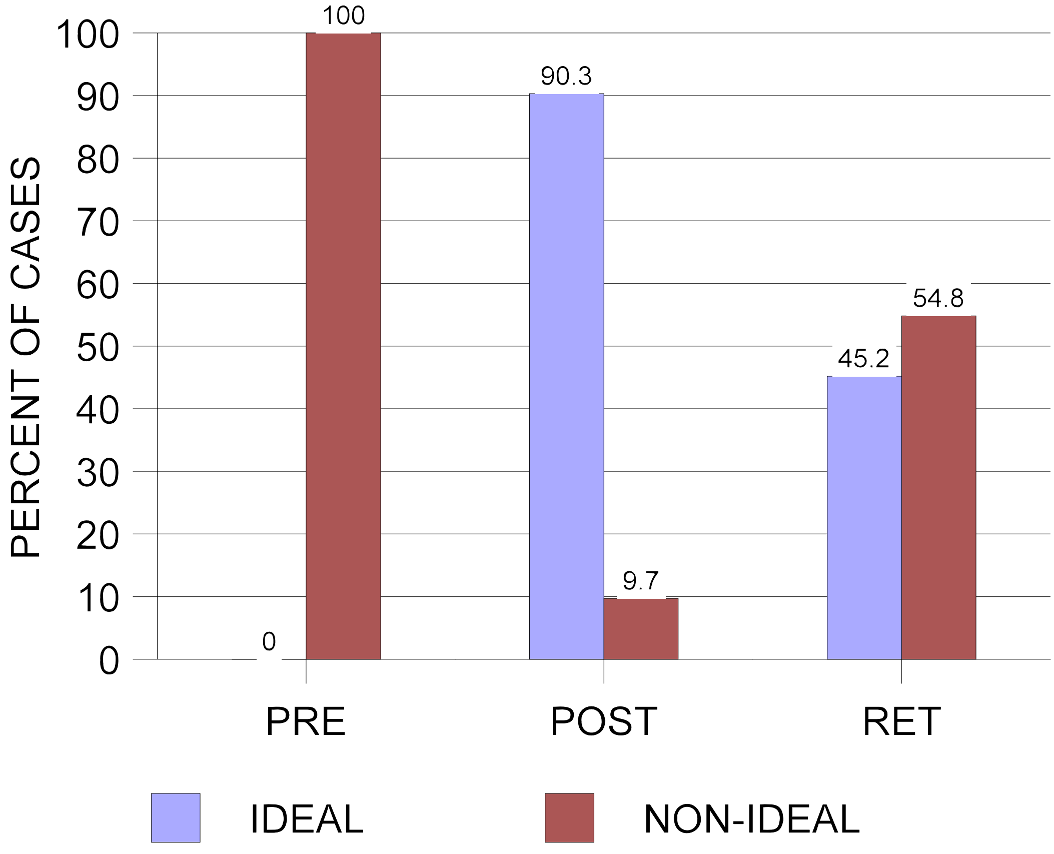
(S = Score)

Relationships seen in Tables 4.2.3.1 and 4.2.3.2 which were assessed as being moderate (according to the parameters set out by Sadowsky and Sakols, 1982) at the commencement of treatment (T1) tended to revert to this status at the end of the retention period (T3). A number of occlusal traits in this category did, however, show some improvement following the retention phase. Noted especially were improvements observed in canine and overjet relationships and maxillary crowding.

Occlusal traits which were initially (T1) classified as being severe were observed in 23 of the sample (Tables 4.2.3.1 and 4.2.3.2). These occlusal traits comprised molar (6 cases) and canine (4 cases) relationships, overbite (6 cases), 5 cases each of maxillary and mandibular crowding, and one case each of open bite and overjet. All of these traits were treated to an ideal relationship at the end of treatment(T2). Of these, 4 of the cases deteriorated in the postretention stage (T3). The severe overjet showed mild relapse at T3, and 3 of the overbite cases showed moderate relapse at this same stage. In total, 13 cases (41.9%) presented with at least one non-ideal occlusal relationship at T3 and four cases presented with more than one non-ideal occlusal relationship at this stage.

Non-ideal occlusal relationships at T3 were noted in 54.8 percent of cases. This was 45.1 percent greater than the incidence recorded at T2 (Figure 4.2.4). As was expected at T1, all of the patients presented with one or more non-ideal dental relationships. It was, however, disappointing to find that 50 percent of the cases which exhibited ideal relationships at T2, deteriorated at T3 (Figure 4.2.4).

Although the number of cases which displayed non-ideal relationships at T3 was high, the analysis of each of the discrete variables studied was more encouraging for all of those variables, with the more ideal relationships predominating over those that were less acceptable (Figures 4.2.5 and 4.2.6). In decreasing order of prevalence of relapse, the following variables accounted for the mean postretention score of 1.5; deep bite (overbite), overjet, molar relationship, anterior and posterior crossbite (XBITE), mandibular crowding, open bite, canine relationship and maxillary crowding.



(PRE = T1), (POST = T2), (RET = T3)

Figure 4.2.4 Ideal vs non-ideal relationships.

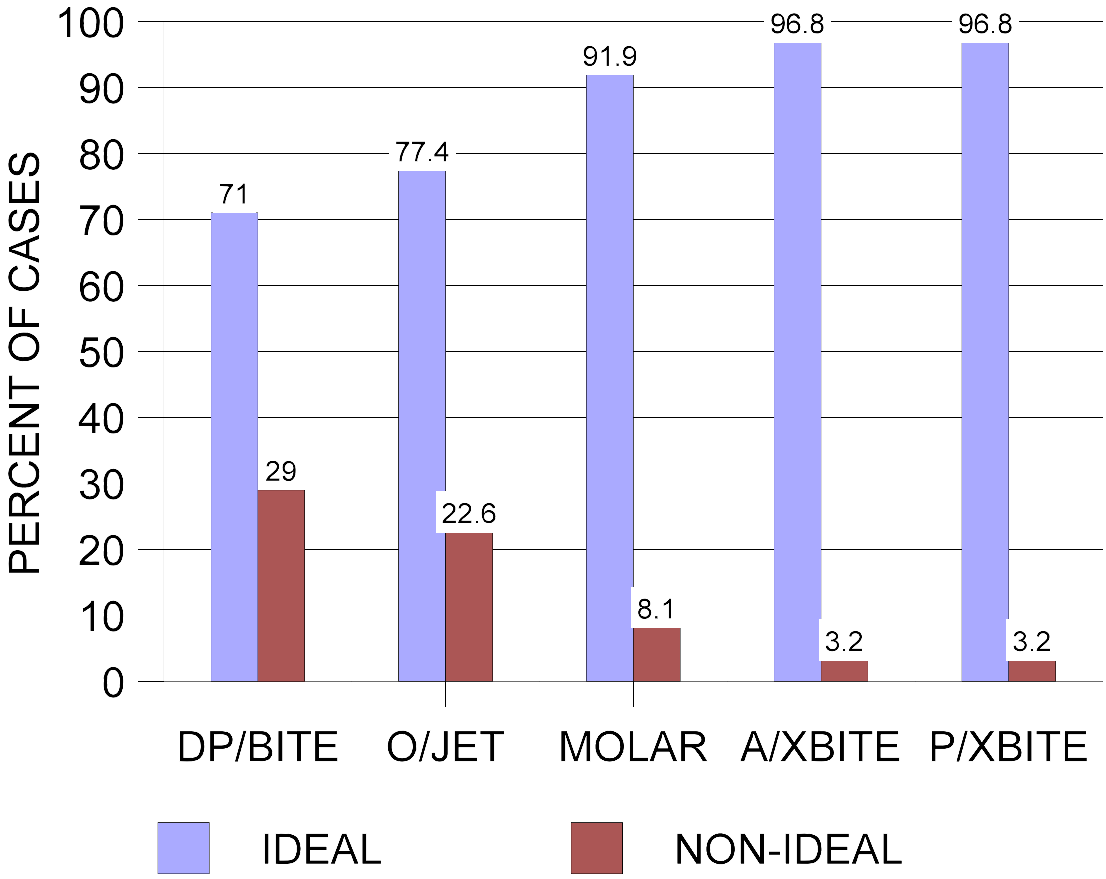


Figure 4.2.5 Ideal vs non-ideal relationships at postretention (T3).

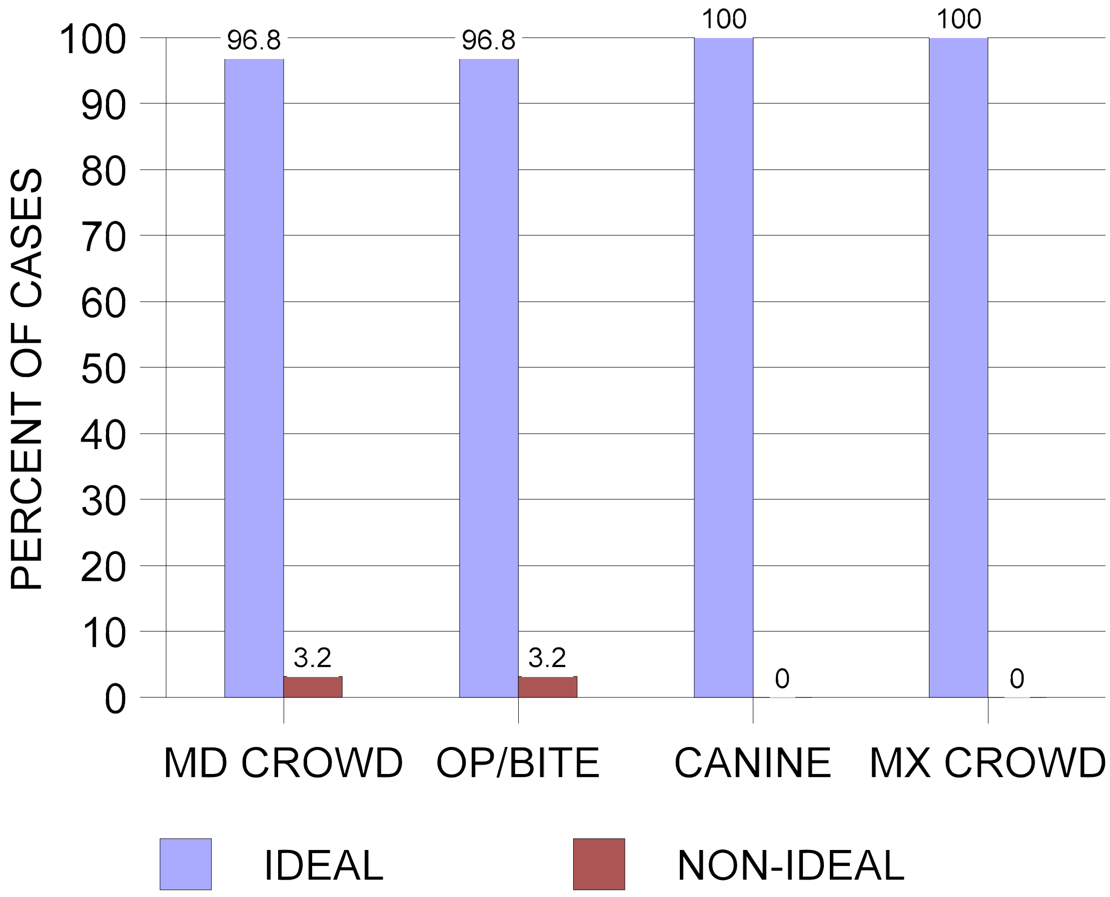


Figure 4.2.6 Ideal vs non-ideal relationships at postretention (T3).

Logistic regressions of each variable (canine relationship, molar relationship, open bite, anterior crossbite, overjet, overbite, posterior crossbite, maxillary and mandibular crowding), dichotomised as 0 (ideal score) and > 0 (non-ideal score) at T3, against 3-3 (T1 and T2) and 6-6 (T1 and T2) as well as changes in these latter variables were performed in order to i) identify statistically significant relationships and ii) determine the point of expansion or contraction where these differences occur (if any). Using backward elimination, no statistically significant predictors (at the 5% significance level) were found for any of the response variables. Consequently, it was not possible to identify any change-points. Furthermore, in order to identify relationships between the following two variables, irregularity index (T3) and inter-incisal angle (T3), as well as changes therein, ordinary (least squares) linear regression was performed on 3-3 (T1 and T2) and 6-6 (T1 and T2) and changes in these variables. No statistically significant predictors (at the 5% significance level) were found for either irregularity index (T3) or inter-incisal angle (T3).

The irregularity index was then dichotomized ($\leq$3 and >3) and logistic regression on 3-3 and 6-6 at T1 and T2 as well as the changes in 3-3 and 6-6 (T1-T2) by gender and class was performed.

Considering the whole sample, including class and gender, the T2 3-3 measurement was found to be a statistically significant predictor (p-value = 0.025) of the probability of the irregularity index at T3 being less than or equal to three (Appendix J). All other predictors were found not to be statistically significant (at the 5% level).

The table referring to the analysis of maximum likelihood estimates indicates a negative association (-0.745) of the post-treatment 3-3 width with the postretention irregularity index(Appendix J). This indicated that as the lower intercanine width was expanded during treatment, the probability of finding a low irregularity index following retention was reduced.

Spearman correlation coefficients were calculated between each of the following variables at T3; canine and molar relationships, overjet and deep bite, mandibular crowding and irregularity index. These correlation coefficients were performed within each class and gender, and by gender and class, against all other T1 and T2 variables (Appendix K).

In all patients with Class I malocclusions, a large open bite at T1, as well as a large change in the open bite relationship between T1-T2 was positively associated with a poor T3 score. Similarly, a poor T2 molar and malocclusion score was associated with a high T3 molar relationship score. When considering the estimates for Class I in females, caution is expressed regarding interpretation of statistically significant results as there were only 3 observations in that group. Also, since there were only 2 observations in males with Class I, no meaningful estimates were produced.

In Class II division I subjects, a large difference between T1-T2 molar scores was associated with good T3 molar stability, while a large change between T1-T2 open bite scores was associated with a small incisor irregularity index score, but with a high T3 molar score. Similarly, an increase in T3 intercanine width was positively correlated with a poor T3 molar relationship as well as mandibular crowding. A large T1 intermolar width was, however, associated with low T3 mandibular crowding. Regarding inter-incisal angle measurement, larger T2 angles were associated with less relapse at T3, and larger T3 angles were significantly associated with a large degree of overjet stability at T3. Large T1 malocclusion scores were associated with reduced T3 molar scores, while large T2 and T3 malocclusion scores were associated with poor T3 molar scores, the high T3 malocclusion score also being associated with poor T3 overbite and overjet (scores).

Furthermore, an assessment was made to determine what role gender played regarding postretention stability or relapse. Females in the total sample with poor T1 canine and open bite relationships, as well as changes in these relationships (T1-T2), exhibited high T3 overjet and irregularity index stability respectively. Within all the females in this sample, an increase in 3-3 width (T1-T2) and an increased T3 6-6 width, was associated with small T3 overjet scores in the former, and small T3 overjet and molar scores in the latter. High T1, T2 and T3 scores were positively associated with low overjet, and high molar and overjet T3 scores, respectively.

Males in the total sample showed a positive association between T2 and T3 molar scores, however, large changes in the T1-T2 molar scores were negatively associated with T3 molar scores. In addition, this group showed a high association between a large T1-T2 6-6 change and deep bite stability at T3. Similarly, a high T2 inter-incisal angle was also positively associated with T3 deep bite stability. However, a large change in the inter-incisal angle T1-T2 resulted in a poor T3 deep bite score. High T2 and T3 malocclusion scores were associated with poor T3 molar, overjet and overbite stability, respectively.

Females in the Class II division I subgroup exhibited a good T3 overjet relationship after a poor T1 canine score. A high T1 molar score was also associated with good T3 overjet and overbite scores (marginally significant, $p \leq 0.1$). A high T2 inter-incisal angle was associated with good T3 deep bite and irregularity index stability. A high T1 malocclusion score also showed a low T3 overjet score.

In the male Class II division I subgroup, an increase in 3-3 width T2-T3 as well as an increase in the 6-6 width T1-T2, was associated with good T3 overjet and overbite stability respectively. A high T2 molar score was associated with continued relapse at T3, while a large change between the molar T1-T2 scores was associated with high postretention stability. A similar positive association between a high T2 inter-incisal angle and T3 deep bite stability was seen in this group as was seen in the males in the total sample as well as the females in the Class II division I subgroup. A large change in the inter-incisal angle T1-T2 showed poor deep bite stability at T3.

The findings of the Wilcoxon rank tests for equal medians of the relevant variable in the presence or absence of wisdom teeth overall and by class at T1, are detailed in Appendix C. Where possible, exact p-values were calculated. The nine variables evaluated were those which made up the malocclusion score together with the inter-incisal angle and the irregularity index.

Within the Class II division I subgroup, a statistically significant difference (p-value = 0.03) was noted in the manner in which the canine relationship changed with treatment between T1 and T2, as well as between T2 and T3. A marginally significant difference (p-value = 0.07) was seen in the change in inter-incisal angle measurement between T1 and T2.

A two-sample t-test was performed to determine whether a statistically significant difference existed in the malocclusion scores in the change from T2-T3, in those who pretreatment presented with or without wisdom teeth at T1. The mean change in the scores were 1.23 in the wisdom teeth group (n = 13) and 1.17 and in the wisdom teeth absent group (n = 18). The p-value for this difference was 0.91 which was not statistically significant (Appendix A). This finding was confirmed by the two-sample Wilcoxon rank-sum (Mann Whitney) non-parametric test, where a p-value of 0.70 was obtained (Appendix A).

The results for Wilcoxon rank tests for equal medians (of the relevant variable) in extraction and non-extraction subgroups (overall and by class) are listed in Appendix D. The variables evaluated were the nine which made up the malocclusion score as well as the inter-incisal angle and irregularity index. Where possible, exact p-values were calculated. There were no patients in the Class I subgroup who underwent extraction, and hence no tests for that class were performed.

In the total sample, a statistically significant mean difference (p-value $\leq$ 0.05) between the extraction and non-extraction groups was noted in the T1 irregularity index, and the change from T1-T2 and T1-T3 in the irregularity index and inter-incisal angle (Appendix D). Furthermore, marginally statistically significant differences (p-value $\leq$ 0.1) were seen in the T1 inter-incisal angle, overbite and mandibular crowding measurements. The treatment response of the latter two variables from T1-T2 were also marginally significant (p-value $\leq$ 0.1).

When evaluating Class II division I malocclusion separately, statistically significant differences were seen in the treatment response from T1-T2 for the following variables; inter-incisal angle,

irregularity index, mandibular crowding, molar relationship and overbite. Statistically significant T1-T3 differences in inter-incisal angle, irregularity index, mandibular crowding and molar relationship were also seen.

A two-sample t test was performed to determine whether the differences in the mean total score from T2-T3 (post-treatment relapse) between the extraction and non-extraction groups were statistically significant. The p-value was determined to be 0.66 (not significant), and this value was confirmed by the non-parametric Wilcoxon rank-sum (Mann-Whitney) test (p-value = 0.51) (Appendix A).

The extraction and non-extraction groups for the total sample were dichotomized into those patients who presented with an ideal overjet and overbite relationship (score =0) at T3, and those who did not (score $\geq$0). The Fisher's exact test showed no statistically significant difference between the extraction and non-extraction groups with respect to overjet (p-value = 1.0) and overbite (p-value = 0.15) scores (Table 4.2.7) (Appendix A).

Table 4.2.7 Evaluation of extraction treatment with overjet (T3) and overbite (T3).

T3	NON-EXTRACTION TREATMENT	EXTRACTION TREATMENT	TOTAL	p VALUE
OVERJET - IDEAL	6	19	25	1.0
NON-IDEAL	1	5	6	
OVERBITE - IDEAL	3	19	22	0.15
NON-IDEAL	4	5	9	

Table 4.2.8 Ideal and non-ideal relationships of the variables in extraction cases observed at the three stages of treatment (N=24)

NON-IDEAL				IDEAL		
PRE	**POST**	**RET**	**VARIABLE**	**PRE**	**POST**	**RET**
12 (7)(5)	**0**	**0***	**CANINE**	**12** (7)(5)	**24** (14)(12)	**24*** (14)(10)
15 (8)(7)	**2** (1)(1)	**3** (2)(1)	**MOLAR**	**9** (6)(3)	**22** (13)(9)	**21** (12)(9)
3 (3)	**0**	**1** (1)	**OPEN BITE**	**21** (11)(10)	**24** (14)(10)	**23** (13)(10)
3 (2)(1)	**0**	**1** (1)	**ANT X-BITE**	**21** (12)(9)	**24** (14)(10)	**23** (14)(9)
19 (11)(8)	**1** (1)	**6** (3)(3)	**OVERJET**	**5** (3)(2)	**23** (14)(9)	**18** (11)(7)
13 (6)(7)	**0**	**5** (2)(3)	**DEEP BITE**	**11** (8)(3)	**24** (14)(10)	**19** (12)(7)
4 (3)(1)	**0**	**1** (1)	**POST X-BITE**	**20** (11)(9)	**24** (14)(10)	**23** (13)(10)
15 (9)(6)	**0**	**0***	**MX CROWDING**	**9** (5)(4)	**24** (14)(10)	**24*** (14)(10)
13 (7)(6)	**0**	**1** (1)	**MD CROWDING**	**11** (7)(4)	**24** (14)(10)	**23** (10) (13)

(FEMALE) **(MALE)**

PRE = PRETREATMENT (T1), POST = POST-TREATMENT (T2),

RET = POSTRETENTION (T3)

*** RELATIONSHIP REMAINED THE SAME.**

Table 4.2.9 Ideal and non-ideal relationships of the variables in non-extraction cases observed at the three stages of treatment (N=7)

NON-IDEAL				IDEAL		
PRE	**POST**	**RET**	**VARIABLE**	**PRE**	**POST**	**RET**
4 (1) (3)	**0**	**0***	**CANINE**	**3** (2) (1)	**7** (3) (4)	**7*** (3) (4)
6 (3) (3)	**0**	**0***	**MOLAR**	**1** (1)	**7** (3) (4)	**7*** (3) (4)
0	**0**	**0***	**OPEN BITE**	**7** (3) (4)	**7** (3) (4)	**7*** (3) (4)
1 (1)	**0**	**0***	**ANT X-BITE**	**6** (3) (3)	**7** (3) (4)	**7*** (3) (4)
5 (2) (3)	**0**	**1** (1)	**OVERJET**	**2** (1) (1)	**7** (3) (4)	**6** (3) (3)
6 (3) (3)	**0**	**4** (1) (3)	**DEEP BITE**	**1** (1)	**7** (3)(4)	**3** (2) (1)
0	**0**	**0***	**POST X-BITE**	**7** (3) (4)	**7** (3) (4)	**7*** (3) (4)
2 (2)	**0**	**0***	**MX CROWDING**	**5** (3) (2)	**7** (3) (4)	**7*** (3) (4)
1 (1)	**0**	**0***	**MD CROWDING**	**6** (3) (3)	**7** (3) (4)	**7*** (3) (4)

(FEMALE) **(MALE)**

PRE = PRETREATMENT (T1), POST = POST-TREATMENT (T2),

RET = POSTRETENTION (T3)

*** RELATIONSHIP REMAINED THE SAME.**

In extraction cases, some non-ideal relationships deteriorated further following treatment until the end of retention. These were overjet (6 cases) and molar relationship (3 cases) (Table 4.2.8). Similarly, ideal relationships tended to deteriorate during the above mentioned period for all variables except the canine relationship and maxillary crowding which remained the same. This relapse was especially noted with the deep bite (5 cases).

The findings for the non-extraction cases are displayed in Table 4.2.9. All the variables in the seven cases were treated to an ideal relationship at T2. Relapse at T3 was, however, recorded for deep bite in 4 cases and overjet in 1 case.

To consider the possibility of a predictor for relapse at T3, the Wilcoxon rank test was performed to determine whether the treatment time or the amount of time out of retention would have a significant bearing on the T2 and/or T3 variables which constituted the malocclusion score as well as the inter-incisal angle measurement and irregularity index. The duration of treatment was dichotomized into those undergoing treatment for less than 25 months (n = 15), and those whose treatment extended for 25 months or longer (n = 16). Furthermore, the period of time out of retention was also dichotomized into a postretention time of less than 158 months (n = 15), and 158 months or longer (n = 16). The test was performed for the total sample as well as by class. The results obtained are shown in Appendices E and F.

When considering treatment time within the total sample, statistically significant differences ($p \leq 0.05$) were obtained for the two groups in the median T1 inter-incisal angle measurement , as well as for the amount the inter-incisal angle changed from T1-T2. In addition, the response from

T1-T3 in the two groups was found to be significantly different with respect to mandibular crowding (Appendix E).

Marginally significant differences ($p \leq 0.1$) were noted in the two groups following treatment in the overall T2 malocclusion scores as well as the T1 mandibular crowding and irregularity index score and the treatment response from T1-T2 for mandibular crowding and irregularity index (Appendix E).

When considering the total malocclusion score, the two sample Wilcoxon rank-sum test showed no statistically significant difference ($p = 0.13$) for the two groups. This was confirmed by the two-sample t-test (p value = 0.12). Furthermore, when considering class differences, Class II division I malocclusion showed statistically significant differences ($p = 0.05$) in its T1 inter-incisal angle measurement and marginally statistically significant differences in the T1-T2 inter-incisal angle measurement ($p = 0.07$) and in the overall (T1-T3) mandibular crowding score (Appendix E).

A Fisher's exact test was performed to determine whether the two categories of treatment time would play a statistically significant role in whether the overjet and overbite score would be ideal (score = 0) or non-ideal (score > 0) at T3. This test produced no statistical significance for either overjet ($p = 0.17$), or overbite ($p = 1.0$) (Appendix E).

The Wilcoxon rank test for postretention time showed a marginally statistically significant difference between the two groups in Class I and Class II division I subgroups (Appendix F). In patients with Class I malocclusions, this difference was expressed in the T1inter-incisal angle ($p = 0.1$) and mandibular crowding score ($p = 0.1$), and the manner in which mandibular crowding changed from T1-T2 ($p = 0.1$) and overall from T1-T3 ($p = 0.1$). In the Class II division I group, a marginally statistically significant difference was recorded in the manner in which the deep bite variable changed from T1-T3.

In addition, a Fisher's exact test was performed to determine whether the two categories of postretention time would play a statistically significant role in determining whether the overjet and/or overbite T3 score would be ideal (score = 0) or non-ideal (score > 0). This test revealed no statistical significance for either the overjet ($p = 0.17$) or overbite ($p = 0.25$) (Appendix E).

In a further search for a predictor for T3 relapse, the Wilcoxon rank test was performed to determine whether the T1 inter-incisal angle could be a statistically significant predictor of T2 and/or T3 variables which constituted the malocclusion score as well as the inter-incisal angle measurement and irregularity index. The T1 inter-incisal angle was further dichotomized into <135 degrees ($n = 19$) and ≥ 135 degrees ($n = 12$), from which the Wilcoxon rank tests for equal medians was performed for the whole sample as well as by class . The findings of this test are listed in Appendix G. Furthermore, the T1 inter-incisal angle was trichotomized into ≤ 127 degrees ($n = 8$), 127-135 degrees ($n = 11$), and ≥ 135 degrees ($n = 12$), and the Kruskal-Wallis test for equal medians of the same variables was performed. The findings of this test are listed in Appendix H.

When analysing the Wilcoxon rank test for the whole sample, statistically significant differences ($p \leq 0.05$) were seen between the two groups in the deep bite T1 median score as well as the score changes from T1-T2 (Appendix G).In addition, statistically significant differences were also seen in the manner in which the inter-incisal angle responded to treatment from T1-T2 and T1-T3 respectively. A marginally statistically significant difference was observed between the T2 and T3 irregularity indices.

The two-sample t test with equal variance displayed a marginally statistically significant difference in the two groups ($p = 0.06$) between the change in median scores from T2-T3. A Fisher's exact test was performed to determine whether the two categories of T1 inter-incisal angle would play a statistically significant role in determining whether the overjet and overbite score would be ideal (score = 0) or non-ideal (score > 0) at T3. This test showed a statistically significant difference ($p = 0.003$) between the two groups for determining overjet, and marginally significant differences ($p = 0.06$) for overbite. When the T1 inter-incisal angle was trichotomized, the Fisher's exact test showed no statistically significant difference for these same two variables.

In patients with Class II division I malocclusions, a statistically significant difference was established ($p \leq 0.05$) between the two dichotomized groups for changes from T1-T2 for inter-incisal angle and deep bite (Appendix G). Similar significance was noted for the T1-T3 inter-incisal angle and mandibular crowding change. In the same group, marginal significance was indicated for the differences between the T1 mandibular crowding and irregularity index, the T1-T2 mandibular crowding, and the T1-T3 changes in irregularity index and overbite variables.

On assessing the total sample when trichotomizing the T1 inter-incisal angle, a statistically significant difference was determined ($p \leq 0.05$) for (i) median T1 and T1-T2 overjet score, (ii) median T3 molar relationship and (iii) median T1-T2 and median T1-T3 inter-incisal angle (Appendix H). A marginally significant difference was seen in the median change in mandibular crowding in all three phases of treatment, as well as the median change in overjet from T2-T3 across the different T1 inter-incisal angle categories defined.

In the Class II division I group, statistically significant differences ($p \leq 0.05$) were found in the three groups regarding the median mandibular crowding T1 score, median change from T1-T2 and median overall change from T1-T3. Furthermore, significant differences were seen in the T1 score as well as changes in deep bite and overjet as a result of treatment (T1-T2). Although the three different T1 scores as well as the difference in the median scores between T1 and T2 were only marginally statistically significant, statistical significance was observed ($p = 0.05$) in the overall difference between the three from T1-T3.

4.3 Intra-Examiner Repeatability

For the determination of intra-examiner variability, the Spearman rank correlation coefficient test revealed a high degree of association and showed no significant differences ($p \leq 0.05$).