AMERICAN UNIVERSITY OF BEIRUT

VALIDATION AND COMPONENT ASSESSMENT OF ORTHODONTIC TREATMENT NEED, COMPLEXITY AND OUTCOME

by
MOHANNAD NAWAF KHANDAKJI

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science in Orthodontics
to the Department of Otolaryngology - Head & Neck Surgery
of the Faculty of Medicine
at the American University of Beirut

Beirut, Lebanon
August 2017

AMERICAN UNIVERSITY OF BEIRUT

VALIDATION AND COMPONENT ASSESSMENT OF ORTHODONTIC TREATMENT NEED, COMPLEXITY AND OUTCOME INDICES

by
MOHANNAD NAWAF KHANDAKJI

Approved by:

Dr. Joseph Ghafari, Professor and Head — Advisor
Orthodontics and Dentofacial Orthopedics

Dr. Hani Tamim, Associate Professor and Director — Member of Committee
Biostatistics Unit at the Clinical Research Institute

Dr. Elie Akl, Professor and Director — Member of Committee
Clinical Research Institute

Dr. Anthony Macari, Assistant Professor — Member of Committee
Orthodontics and Dentofacial Orthopedics

Date of thesis defense: August 29, 2017

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Khandakji Mohannad Nawaf

Last First Middle

☑ Master's Thesis ◯ Master's Project ◯ Doctoral Dissertation

☐ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☑ I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes

after: **One ---- year from the date of submission of my thesis, dissertation, or project.**
Two ---- years from the date of submission of my thesis, dissertation, or project.
Three ---- years from the date of submission of my thesis, dissertation, or project.

Sep 7 2017

Signature Date

ACKNOWLEDGMENTS

It is a pleasure to express my utmost appreciation to all the individuals who made this thesis possible; therefore, I would like to seize this opportunity to thank everyone who contributed to this work in one way or another.

I am sincerely grateful to my committee members

Dr Jospeh Ghafari, not only for being my primary advisor, but also for pushing me to pursue this second master degree, which provided me the opportunity to further expand my capabilities;

Dr Hani Tamim, for your time and generous advice. I also truly appreciate your encouraging comments and support;

Dr Elie Akl, for your optimisms, kindness, and valuable comments. All of which kept me going;

Dr Anthony Macari, not only for being the teacher and the friend, but also for giving me the opportunity to be your research partner;

Special recognition goes to

Dr Samar Bou Assi, Roy Sabri, Patrick Anhoury, and *Felipe Rizk*, who generously gave their time to participant in this research; I am deeply grateful for all the hours they spent assessing the cases.

Moreover, I would like to extend my gratitude to:

Dr Pamela Genno, Maria Doughan, Manoel Khoury, Nancy Abdo and *Christophe Zoughaib*, not only for being amazing collogues, but also for helping me in recording the panel assessments when I was away, so thank you.

Ms Michella Semaan and Lynn kreik for helping me in finding and adjusting the dental models that are included in this study.

and last but not least, to *my family*, it is only through their love and never-ending support that I have reached this point of my academic endeavors. Thank you for helping me to accomplish all of my goals.

AN ABSTRACT OF THE THESIS OF

Mohannad Nawaf Khandakji for Master of Science
Major: Orthodontics

Title: Validation and component assessment of orthodontic treatment need, complexity and outcome indices

Introduction:
One of the major challenges in orthodontics is the availability of a suitable objective method for recording the severity and complexity of an orthodontic problem. As a result, occlusal indices were realized initially as a tool to classify or rank occlusion. A great number of occlusal indices appeared since the 50's and 60's to minimize the subjectivity related to the diagnosis, severity, complexity and treatment outcome of malocclusions. Despite the great number of occlusal indices, there have been many disagreements among authors, with no universally accepted index for measuring malocclusion.

Objective:
The aims of this study were to -1- evaluate and validate the most commonly used indices: Index of Orthodontic Treatment Need (IOTN) that includes a dental health component (DHC) and an aesthetic component (AC), Peer Assessment Rating Index (PAR), Index of Complexity, Outcome, and Need (ICON), Discrepancy Index (DI) as well as a recently published Facial Aesthetic Index (FAI); -2- introduce modifications encompassing teeth, cephalometrics and facial esthetics; -3- compare the indices with each other. Each index was assessed as a measure of orthodontic treatment need, complexity, and outcome. Thus, not restricted to the respective original objective for which it was developed.

Methods:
This validation study included the pre- and post-treatment lateral cephalometric and panoramic x-rays, dental casts, and photographs of 101 participants. Data were collected from the files and records of patients treated at the Division of Orthodontics and Dentofacial Orthopedics at the American University of Beirut Medical Center AUBMC. The participants were randomly selected, from a list containing all patients whose orthodontic treatment was completed and appliances removed by the orthodontic residents of the same class (2015). The indices were validated against the independent judgments of a panel of four experienced orthodontists 'gold standard' who met strict criteria for inclusion. Prior to the initiation of the study, the examiner was calibrated against experienced orthodontists to ensure precise assessment of the chosen occlusal indices. The cephalograms used to construct the DI but also for general comparative computations, were digitized and measured using the Dolphin digitizing program. Measures were repeated in 10% of the sample to assess reliability. The scales were compared to each other and to the gold standard before and after treatment. The correlations between the scales were computed using Spearman's correlation and their reliability were determined through intra-class correlation coefficients. Receiver operating curves (ROC) were derived for each index, and the corresponding areas under the curve (AUC) were used to find the best cutoff points for the different indices.

Results:

The total sample consisted of 39 males and 62 females with an average age of 17.2 years. All occlusal indices had excellent levels of intra-examiner reliability. Based on the panel assessments 70 participants had definitive treatment need and 31 had no definitive need. The FAI did not enhance the correlation of indices with need. The summation of the AC and the DHC components into a proposed combined index score (cIOTN) yielded the strongest correlation with need ($r = 0.79$), and had the highest AUC value of 88.5% with sensitivity of 91% and specificity of 71% at the cutpoint ≥6. The DI had the weakest correlation with the need; however, its dental component (dDI) was substantially better ($r = 63$). The weighted and raw scores of both ICON and PAR had almost the same discriminating abilities in identifying definitive treatment need. In assessing complexity, the DI had one of the lowest correlation coefficients ($r=0.43$), and the simple DHC had the highest coefficient ($r = 0.72$). Except for the DHC and the dDI, the correlation between the indices and treatment complexity increased by adding the supplemental FAI. This increase was more significant in some indices than others. On average, the post-treatment index scores had 3 times the weight of the pre-treatment scores in predicting the amount of malocclusion improvement after orthodontic treatment. The ICON, the cIOTN, and the DI had almost equal correlations with malocclusion improvement after treatment (r = 0.65, 0.64, and 0.64, respectively). The ICON, the cIOTN, and the dDI demonstrated good psychometric properties in discriminating unacceptable from acceptable outcomes (AUC = 0.82, 0.85, and 0.86, respectively).

Conclusion:

General conclusion: This study was the first to address occlusal indices with various constructs against one gold standard, and to gauge their multifunctional potential. New findings include: -1- the potential for the proposed combined IOTN (cIOTN) as an excellent multifunctional index; -2- the ability of the dental DI (dDI) to reflect need, complexity and outcome success better than the original DI; and -3- the failure of non-dental measures, specifically lateral cephalometric and facial soft-tissue, to represent an added value to the accuracy of the respective indices in assessing need and outcome.

Detailed conclusions: Except for the DI, all studied indices reliably captured the silent dimension that orthodontists use in determining need for treatment. While the proposed cIOTN yielded better psychometric properties than its components and all other included indices, our findings did not support the PAR and the ICON as measures of treatment complexity. The DI had one of the lowest correlation coefficients with complexity assessment; however, using the dDI alone had a significantly stronger correlation. The simple DHC was the only index that achieved the recommended levels in measuring complexity. The FAI could not be validated as a supplemental measure of treatment need; however, it slightly increased the correlations of some occlusal indices with the treatment complexity. The DI, the cIOTN and the ICON are considered valid measures of malocclusion improvement after treatment. In directly assessing the success of orthodontic treatment, the ICON, the cIOTN and the dDI had a good ability to correctly classify those with and without acceptable outcomes, unlike the DHC, the PAR, and the DI. The cIOTN demonstrated surprising abilities in measuring all different constructs (need, complexity, improvement, and outcome success); yet, in assessing complexity, the DHC scores alone should be used.

CONTENTS

ACKNOWLEDGEMENTS v
3.5. Calibration .... 49 vi
3.5. Calibration .... 49 xii
3.5. Calibration .... 49 xiii
3.5. Calibration .... 49 xvi

Chapter

1. INTRODUCTION 1

2. LITERATURE REVIEW 7

2.1. Dental occlusion 7
2.1.1. Concepts and definitions 7
2.1.2. Assessments and classifications 8
2.2. Occlusal indices 11
2.2.1. Background 11
2.2.2. Requirements for an ideal index 13
2.3. Orthodontic treatment need indices 15
2.3.1. Index of Orthodontic Treatment Need (IOTN) 16
2.3.2. Modified IOTN 18
2.3.3. Index of Complexity, Outcome and Need (ICON) ... 19
2.3.4. Facial Aesthetics Index (FAI) 20
2.4. Orthodontic treatment complexity indices 21
2.4.1. Discrepancy Index (DI) 22
2.4.2. Index of Complexity, Outcome and Need (ICON) 23
2.5. Orthodontic treatment outcome indices 25
2.5.1. Peer Assessment Rating (PAR) 25
2.5.2. Index of Complexity, Outcome and Need (ICON) 27
2.6. Significance 28

2.7. Specific Objectives 30
2.7.1. Objectives of the treatment need study 30
2.7.2. Objectives of the treatment complexity study 30
2.7.3. Objectives of the treatment outcome study 31

3. METHODS 32
3.1. Research Design 32
3.2. Sample Selection 32
3.3. Data collection 33
3.4. Measures 33
3.4.1. Index of Orthodontic Treatment Need 34
3.4.2. Peer Assessment Rating 37
3.4.3. Index of Complexity, Outcome and Need 39
3.4.4. Discrepancy Index 41
3.4.5. Facial Aesthetic Index 45
3.4.6. Cephalometric measures 45
3.4.7. Panel measures 48
3.4.8. Other measures 48
3.5. Calibration 49
3.6. Statistical Analyses 49
3.7. Ethical Considerations 50

4. RESULTS 52
4.1. Descriptive Statistics 53
4.1.1. Participant's characteristics 53
4.1.2. Occlusal indices 53
4.1.3. Cephalometric measurements 54
4.1.4. Panel assessments 54

4.2. Analysis of Reliability ………….…………………….….…………...…….……. 55
4.3. Validation of Occlusal Indices as measures of need …...………...…….….…. 55
4.3.1. Panel assessment of need …..…………………………………………..... 55
4.3.2. Correlation analysis against panel assessment …...…………...….... 56
4.3.3. Receiver operating characteristic (ROC) analysis …..…………….. 57
4.3.4. Multivariable analysis ……....………………………………………..... 58
4.4. Validation of Occlusal Indices as measures of complexity .....………...…..…. 60
4.4.1. Panel assessment of complexity …..…………………………………... 60
4.4.2. Correlation analysis against panel assessment …………………..... 60
4.4.3. Correlation analysis against estimated treatment duration ……... 61
4.5. Validation of Occlusal Indices as measures of improvement .....………....… 62
4.5.1. Panel assessment of improvement …..…………………….....……... 62
4.5.2. Correlation analysis against panel assessment …………………..... 62
4.5.3. Multivariable analysis ……....……………………………………….…... 63
4.5.4. Correlation analysis of the modified indices ……....……………..... 64
4.6. Validation of Occlusal Indices as measures of outcome .....…...………...…. 66
4.6.1. Panel assessment of outcome …..………………………………………... 66
4.6.2. Receiver operating characteristic (ROC) analysis …..…………….. 66
4.2. Cephalometric Analysis ………….…………………….….………….....……..……. 67

5. DISCUSSION ………………………....………………………………………....……….. 69

5.1. Introduction ………………..……………………...……………………………..…. 69
5.2. Reliability of the Occlusal Indices ……………………………………………...…. 70
5.3. Occlusal Indices as measures of treatment need …..….………………......…. 71
5.4. Occlusal Indices as measures of treatment complexity …..….….………….…. 76
5.5. Occlusal Indices as measures of treatment outcome …..……...….……….…. 79
5.5.1. Improvement measure ……..……………………..……………………...…. 79
5.5.2. Outcome success measure ……..…………….…………………….....…. 81
5.6. Limitations and research considerations …...………………………….……. 82

5.7. Strengths 85

6. CONCLUSIONS AND RECOMMENDATIONS 87

6.1. Conclusions 87

6.1.1. Conclusions of the treatment need study 87

6.1.2. Conclusions of the treatment complexity study 88

6.1.3. Conclusions of the improvement study 89

6.1.4. Conclusion of the outcome success study 90

6.1.5. Main contributions of the study 90

6.2. Recommendations for future studies 91

TABLES 93

REFERENCES 121

Appendix

I. DI WORK FORM 127

II. PANEL SCORING SHEET 130

III. SUPPLEMENTARY TABLES 132

ILLUSTRATIONS

Figure		Page
2.1	Ackerman and Proffit classification of malocclusion	10
3.1	The Aesthetic Component (AC) scale of IOTN index	35
3.2	The ABO discrepancy index scoring sheet	44
3.3	Facial Aesthetic Index charts for males and females	45
3.4	Later cephalometric radiograph with location of the cephalometric landmarks and reference lines	47
3.5	Cephalometric radiograph with soft tissue analysis	47
4.1	Pre- and post-treatment scores of the different indices	120

TABLES

Table		Page
2.1	Angle's classification of malocclusion	10
2.2	The different types of occlusal indices	11
2.3	Requirements for an ideal index	13
2.4	Modified IOTN grades	18
2.5	ICON complexity measure grades	24
3.1	Orthodontic treatment need defined by IOTN-DHC	36
3.2	The PAR scoring method and its components	38
3.3	ICON scoring method and its components	40
3.4	Subnasal soft tissue profile measurements of E- and H-lines	47
4.1a	Descriptive statistics of pre- and post-treatment indices' score	93
4.1b	Descriptive statistics of pre- and post-treatment cephalometric measurements	93
4.2	Summary statistics of panel assessment scores	94
4.3	Intra-class correlation coefficients for intra-examiner reliability	94
4.4	Spearman correlation analysis between need for orthodontic treatment and different occlusal indices	95
4.5	The ROC curve analysis of the DHC as measure orthodontic treatment need	95
4.6	The ROC curve analysis of the AC as measure orthodontic treatment need	96
4.7	The ROC curve analysis of the combined IOTN (cIOTN) as measure orthodontic treatment need	96

4.8 The ROC curve analysis of the ICON as a measure of orthodontic treatment need 97

4.9 The ROC curve analysis of the raw ICON scores (rICON) as a measure of orthodontic treatment need 98

4.10 The ROC curve analysis of the PAR as a measure of orthodontic treatment need 100

4.11 The ROC curve analysis of the raw PAR scores (rPAR) as a measure of orthodontic treatment need 101

4.12 The ROC curve analysis of the DI as a measure of orthodontic treatment need 102

4.13 The ROC curve analysis of the DI dental component (dDI) as a measure of orthodontic treatment need 103

4.14 Multiple linear regression analysis showing association between need for orthodontic treatment and cIOTN component 104

4.15 Multiple linear regression analysis showing association between need for orthodontic treatment and rPAR components 104

4.16 Multiple linear regression analysis showing association between need for orthodontic treatment and rICON components 105

4.17 Multiple linear regression analysis showing association between need for orthodontic treatment and rICON aesthetic and dental components 105

4.18 Multiple linear regression analysis showing association between need for orthodontic treatment and dDI components 106

4.19 ROC analysis of nICON as measure orthodontic treatment need 107

4.20 ROC analysis of nPAR as measure orthodontic treatment need 108

4.21 ROC analysis of ndDI as measure orthodontic treatment need 109

4.22 Spearman correlation analysis between the complexity of orthodontic treatment and different occlusal indices 110

4.23 Spearman correlation analysis between the estimated treatment duration and the different occlusal indices 110

4.24 Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the different occlusal indices 111

4.25a Multivariable analysis showing association between the malocclusion improvement by orthodontic treatment and the occlusal indices 112

4.25b Multivariable analysis showing association between the malocclusion improvement by orthodontic treatment and the occlusal indices 113

4.26a	Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the weighted occlusal indices	113
4.26b	Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the weighted occlusal indices	113
4.27	The ROC curve analysis of the AC as a measure of orthodontic treatment outcome	114
4.28	The ROC curve analysis of the DHC as a measure orthodontic treatment outcome	114
4.29	The ROC curve analysis of the cIOTN as a measure orthodontic treatment outcome	114
4.30	The ROC curve analysis of the rPAR as a measure orthodontic treatment outcome	115
4.31	The ROC curve analysis of the PAR as a measure of orthodontic treatment outcome	115
4.32	The ROC curve analysis of the rICON as a measure of orthodontic treatment outcome	116
4.33	The ROC curve analysis of the ICON as a measure of orthodontic treatment outcome	116
4.34	The ROC curve analysis of the DI as a measure of orthodontic treatment outcome	117
4.35	The ROC curve analysis of the dDI as a measure of orthodontic treatment outcome	118
4.36	Spearman correlation analysis of the need and complexity of orthodontic treatment with different cephalometric measurements	119
3.37	The association between the different cephalometric measurements and the outcome of the orthodontic treatment as assessed by the panel	119

ABBREVIATIONS

AAO	American Association of Orthodontists
ABO	American Board of Orthodontics
ADA	American Dental Association
AC	Aesthetic Component of IOTN
AUBMC	American University of Beirut Medical Center
cDI	Cephalometric component of DI
cIOTN	combined score of IOTN-AC and -DHC
DI	Discrepancy Index
dDI	dental component of Discrepancy Index
cDI	cephalometric component of Discrepancy Index
DHC	Dental Health Component of IOTN
FAI	Facial Aesthetic Index
ICC	Intra-class correlation coefficient
ICON	Index of Complexity, Outcome and Need
IOTN	Index of Orthodontic Treatment Need
MNOC	Medically Necessary Orthodontic Care
OBS	Objective Grading System
PAR	Peer Assessment Rating
rICON	row scores of Index of Complexity, Outcome and Need
rPAR	row scores of Peer Assessment Rating
UK	United Kingdom
US	United States
WHO	World Health Organization

for

MY LATE FATHER

DR. NAWAF WASFI KHANDAKJI

Who always wanted me
to be an orthodontist

CHAPTER 1

INTRODUCTION

Occlusion in dentistry refers to the relationship between the maxillary and mandibular teeth at both rest and during function. Any deviation from the normal relationship of the maxillary and mandibular teeth is called malocclusion. While any malocclusion can benefit from treatment, not every malocclusion needs treatment. Malocclusions are often imprecisely interpreted as an invariable state that is either present or not, when in fact it is a continuum of many variables that encompasses all variations (Summers, 1971). For this reason, the accurate evaluation of malocclusion is crucial in establishing the diagnosis and treatment need of the orthodontic patient. However, one of the major challenges in orthodontics is the availability of a suitable objective method for recording the severity and complexity of an orthodontic problem.

Occlusal indices were realized initially as an objective tool to classify or rank occlusion, and by far the most commonly used of these indices is the Angle classification of malocclusion (Angle, 1899). A large number of occlusal indices started to appear in the 50's and 60's to assist epidemiological studies (Draker, 1958; Grainger, 1967; Salzmann, 1968; Summers, 1966). Whereas numerous occlusal indices have been proposed over the years, it is important to distinguish between those that classify malocclusion into types and those that record prevalence of malocclusion from indices that attempt to record treatment need. In fact, Shaw et al. (1995) divided occlusal indices into 5 different categories; diagnostic, epidemiological, treatment need, treatment outcome, and treatment complexity indices. Their use minimizes the subjectivity related to the diagnosis, severity, complexity and treatment outcome of malocclusions (Richmond & Daniels, 1998a). Occlusal indices are able to benefit the orthodontic profession in many ways, they can help in; resources allocation and manpower planning, assessing the relationship between malocclusions and other medical or dental conditions, assessing the orthodontic treatment outcome and clinical performance, evaluating the malocclusion treatment complexity and its effect on treatment time, and in assessing the cost effectiveness of orthodontic treatment.

Angle’s classification remains the most widely used system for clinical and epidemiological purposes, a practice largely attributed to the simplicity of the method. Nevertheless, it is a basic qualitative index that does not provide any information on the severity nor the complexity of the malocclusion, therefore, it has no usefulness in recoding treatment priority and has been shown to have poor reproducibility (Gravely & Johnson, 1974). Treatment need indices are a category of occlusal indices used to prioritize the malocclusions need for treatment, and have been used in planning the provision of orthodontic treatment as well as in epidemiological studies. The Index of Orthodontic Treatment Need (IOTN) is perhaps the most common and most accepted index for assessing the orthodontic treatment need (Borzabadi-Farahani, 2011). The perception of treatment need can be different between different countries, thus affecting the cutoff point at which the treatment is indicated (Richmond & Daniels, 1998a). Accordingly, the IOTN with its two components (dental health and aesthetic) has been developed and validated against a panel of orthodontists in Great Britain as well as in different geographical locations (Beglin et al., 2001; Shaw, Richmond, & O'Brien, 1995; Younis, Vig, Rinchuse, & Weyant, 1997), and was assessed several times against different indices (Fox, Daniels, & Gilgrass, 2002; Johnson, Harkness, Crowther, & Herbison, 2000).

In addressing the need for an objective method of measuring the amount of malocclusion improvement and treatment efficacy, the peer assessment rating (PAR) index was developed (Richmond et al., 1992). The aim was to objectively record the amount of malocclusion in any stage of treatment. The PAR has been validated against the opinion of a panel of 74 dentists, and demonstrated an excellent reliability. However, there was a poor level of agreement between the dentists of the panel ‘gold standard’ in assessing the deviation of the malocclusion from normal. The poor agreement was likely related to the fact that the 74 dentists were represented five different groups of practitioners; 22 consultant orthodontists, 22 specialist practitioners, 15 general practitioners, 11 community dentists, and 4 others (Richmond et al., 1992). Consequently, the validity and the weightings of the different components of the PAR index were based on this poor agreement of the ‘gold standard’, which the PAR index was validated against. The index developers stated that the weighting of the PAR index could be changed to reflect future and current standards in other countries; advocating for the re-validation of PAR index in different countries or for future standards of orthodontics (Richmond et al., 1992).

The index of complexity, outcome, and need (ICON) was developed according to the judgments of 97 orthodontic specialists from 8 European countries and the United States (Daniels & Richmond, 2000; Richmond & Daniels, 1998a, 1998b). Therefore, it is considered to be based on the international opinion of the orthodontic community. ICON is unique as it is a multifunctional index developed to measure treatment need, complexity, and outcome (Daniels & Richmond, 2000). ICON has been shown to be a reliable and valid index for assessing orthodontic treatment need (Daniels & Richmond, 2000; Firestone, Beck, Beglin, & Vig, 2002). The validity of ICON in assessing complexity and outcome measures was examined in a separate study (Savastano, Firestone, Beck, & Vig, 2003); the authors reported moderate agreement between the ICON and the average panel opinion. Nevertheless, there was poor agreement between the members of the panel, and the authors particularly advised on reassessing the degree of improvement section of ICON with more variable subjects (Savastano et al., 2003). Similar to the IOTN and the PAR indices, it is recommended to re-validate the ICON in the geographical location of its application to optimize its cutoff points according to the national standards (Borzabadi-Farahani, 2011; Firestone et al., 2002).

Knowing that the severity of malocclusions and treatment complexity are related but not the same; The American Board of Orthodontics (ABO) in the 1990's felt the need for an objective method of assessing the complexity of the submitted board cases in the "phase III clinical examination". After comprehensive meetings, discussions, and field testing, they introduced the discrepancy index (DI) to gauge treatment complexity (Cangialosi et al., 2004). Since then, all the ABO applicants were required to measure the DI for case selection to fulfill the requirement for phase III of the certification process. The DI is the ABO's approach to objectively summarize and quantify the clinical features of a patient's condition. Yet, the psychometric properties of the DI as a complexity measure have not been investigated.

Aesthetic improvement is one of the most commonly reported reasons for seeking orthodontic treatment (Birkeland, Katle, Løvgreen, Bøe, & Wisth, 1999; Trulsson, Strandmark, Mohlin, & Berggren, 2014). As a matter of fact, psychosocial consequences of malocclusion due to unacceptable esthetics maybe as serious or even more serious than biological problems (Stricker et al., 1979). Facial aesthetic assessment is crucial in orthodontic diagnosis and treatment planning, and it should be included in the evaluation of orthodontic treatment need and outcome (Kiekens, Maltha, van't Hof, & Kuijpers-Jagtman, 2005). However, none of the occlusal indices that have been already developed incorporated soft tissue treatment needs (Sundareswaran & Ramakrishnan, 2016). It is only recently that a Facial Aesthetic Index (FAI) was developed to compliment occlusal indices to address such necessity (Sundareswaran & Ramakrishnan, 2016). The FAI showed excellent reliability and accuracy in determining different profiles into their respective group. However, no study was conducted to assess its validity in measuring orthodontic treatment need. The authors stated that the FAI was intended as an additional screening tool to complement existing indices and generate more information about the patient, with the premise of having a more comprehensive measure of orthodontic treatment need.

The most important two criteria of any diagnostic instrument are the ability to produce consistent results (reliability) and the extent to which it accurately measures what it purports to measure (validity). There are many studies assessing the reliability of different indices and comparing them to each other; nevertheless, validation studies are much less. Many authors stated that using the opinion of experts as a gold standard is an accepted technique in validating occlusal indices (Beglin et al., 2001; Brook & Shaw, 1989; Younis et al., 1997). In our opinion, the index of occlusion should be treated as any diagnostic instrument used to detect conditions that are difficult to define. In the medical field, the diagnosis of many constructs that are challenging to measure is established with an instrument validated against expert opinions, such the case of depression or dementia questionnaires and many others. Also, as mentioned earlier, various authors recommended that occlusal indices (particularly treatment of need indices) should be validated in different countries or geographical locations to establish proper national cutoff points (Richmond & Daniels, 1998a; Younis et al., 1997), with some stating that "the effect of having a local panel of experts should not be underestimated" (Beglin et al., 2001). In fact, different cutoff points of occlusal indices have been reported in different countries (Brook & Shaw, 1989; Firestone et al., 2002). Furthermore, illustrating the association between local provider's opinion and the objective classification by an occlusal index in a certain geopolitical location is considered a crucial first step in the acceptance of any index by the national clinical community (Younis et al., 1997). To the best of our knowledge, no country in our region (Near/Middle East) has validated such occlusal indices.

Despite the great numbers of occlusal indices, many disagreements are noted among the authors, with no universally accepted validated index for measuring malocclusion (Borzabadi-Farahani, 2011). Therefore, the aim of this study was to evaluate the most commonly used indices as well as the recently published Facial Aesthetic Index and compares them with each other and with different modifications of ours; the goal was to determine the optimal index in the different categories and to optimize the corresponding cutoff points to meet our standards of treatment. The evaluations and comparisons of the different indices were based on their reliability, validity, and simplicity. The indices were validated as measures of orthodontic treatment need, complexity, and outcome. Each index was assessed in the three categories, and not only in their respective category.

CHAPTER 2

LITERATURE REVIEW

2.1. Dental occlusion

2.1.1. Concepts and definitions

In dentistry, occlusion simply means the contact between the maxillary and mandibular teeth when they approach each other, and it reflects both the static and dynamic contact relation between the teeth. More technically, it is defined as the manner in which the teeth intercuspate in all mandibular positions and movements. In fact, the term ‘occlusion’ encompasses all occlusal variations from ideal occlusion to malocclusion and it is a result of the integrated components of the masticatory system: teeth, periodontal structures, maxilla, mandible, temporomandibular joints, and their associated muscles and ligaments (Ash & Ramfjord, 1982; Nelson, 2014).

The concept of ‘ideal’ occlusion refers to a theoretical perfect occlusion based on the anatomy of the teeth and is rarely found in contemporary populations (Ross, 2003; Woda, Vigneron, & Kay, 1979). Ideal occlusion can be precisely described and therefore used as the norm by which other occlusions are judged (Andrews, 1972). Unlike ‘ideal’ occlusion, the limits of the concept of ‘normal’ occlusion are not precisely specified. Normal occlusion can be defined as an acceptable deviation of ideal occlusion that does not constitute aesthetic or functional problems (Houston, Stephens, & Tulley, 1992). Nevertheless, controversies still exist in the extent of deviation that can be permitted (Proffit, Fields, & Sarver, 2014).

Any deviation in occlusion from the normal occlusion is called Malocclusion. Malocclusion is described as a normal biological variability rather than a disease and it covers a continuum from a normal occlusion to a substantial deviation from it (Borzabadi-Farahani, Borzabadi-Farahani, & Eslamipour, 2009; Bryant, 2002). Disagreements in defining the concept of the normal occlusal relationship have ultimately led to much debate regarding what is considered a 'malocclusion'. In 1987, the World Health Organization (WHO) included the term 'malocclusion' under the heading of Handicapping Dento-Facial Anomaly, and defined it as an anomaly which causes disfigurement or impedes function, and requires treatment "if the disfigurement or functional defect was likely to be an obstacle to the patient's physical or emotional well-being". Therefore, the presence of malocclusion should not be equated with the need for orthodontic treatment (Sampson & Sims, 1992).

2.1.2. Assessment and classifications

The sophistication of the masticatory system involved in occlusion, together with the confusion about the concept of 'normal' occlusion; lead to a difficulty in classifying malocclusions (Thilander & Rönning, 1995; Türp, Greene, & Strub, 2008). Various classifications of malocclusion have been proposed since the interest in this field began (Ackerman & Proffit, 1969; Angle, 1899; BSI., 1983; Miguel-Neto & Mucha, 2000). In the earlier days, studies of malocclusion did not define methods for measuring the variability of different occlusions, thus malocclusion symptoms were recorded in an all or none manner (Tang & Wei, 1993).

In 1899, Edward H. Angle developed a classification system to describe both 'normal' occlusion and 'malocclusion'. Angle's classification system has survived the test of time, as it is still the primary system employed by our specialty and the only one recognized internationally (Houston et al., 1992; Rinchuse & Rinchuse, 1989). It is based on the antero-posterior relationship between the maxillary and mandibular permanent first molars; where the ideal relationship between these first molars is classified as a class I occlusion and the deviations as either class II or class III malocclusion (**Table 2.1**).

Despite its popularity, Angle's classification has received much criticism for being too simple and ambiguous to describe the infinite variability of dental occlusion, as well as for its lack of reliability (Graber, 1972; Gravely & Johnson, 1974; Rinchuse & Rinchuse, 1989). In fact, several attempts were made to improve Angle's classification and replace it with more sensitive indicators of malocclusion rather than relying solely on the first molars relationship (Ackerman & Proffit, 1969; Björk, Krebs, & Solow, 1964; BSI., 1983; Dewey, 1915). One of these attempts is the British Standards Institute of Incisors classification (Ballard & Wayman, 1965). Although based on the same categories of angle's classifications, the system has been considered superior to it as it relies on the incisor occlusion instead of first molars (Du et al., 1998). Another example is Ackerman and Proffit classification which was developed to overcome the limitations of Angle's system by incorporating angle's method into a more comprehensive diagrammatic classification based on Venn diagram (Ackerman & Proffit, 1969) (**Figure 2.1**). However, owing to its simplicity, Angle's classification system remains the most commonly used system for classification of malocclusion (Rinchuse & Rinchuse, 1989).

Table 2.1: Angle's classification of malocclusion (Angle, 1899, 1907)

Normal occlusion	Mesiobuccal cusp of the maxillary molar occludes in the buccal groove of the mandibular molar, all teeth arranged on a smoothly curving line of occlusion
Class I malocclusion	Normal molar relationship, but line of occlusion incorrect because of malposed teeth, rotations or other causes
Class II malocclusion	Lower molar distally positioned relative to upper molar
Division 1	Class II malocclusion with protruded upper incisors
Division 1	Class II malocclusion with retruded upper incisors
Subdivision	Class II occlusion on one side, and normal occlusion the other side
Class III malocclusion	Lower molar mesially positioned relative to upper molar
Subdivision	Class III occlusion on one side, and normal occlusion the other side

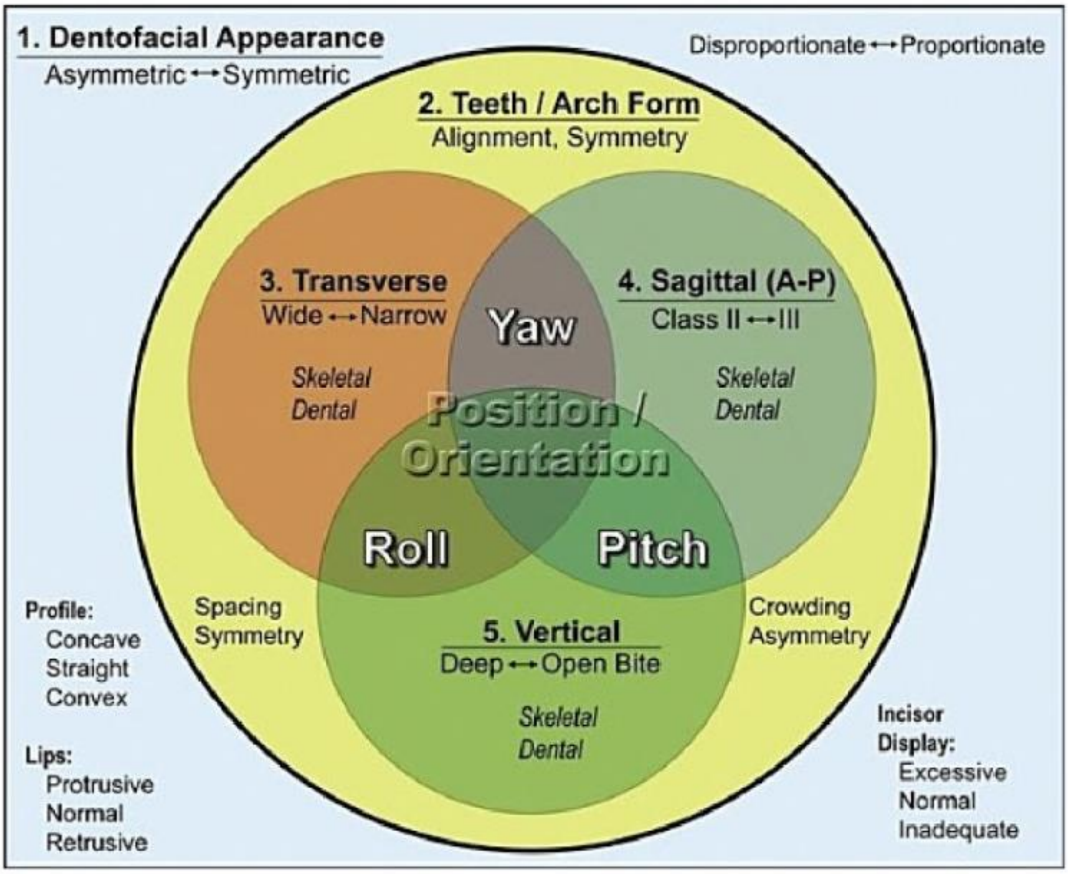

Figure 2.1: Ackerman and Proffit classification system, which present the five-major characteristic of malocclusion via a Venn diagram, adopted from (Proffit et al., 2014).

2.2. Occlusal indices

2.2.1. Background

According to the American Association of Orthodontists (AAO) Glossary of Orthodontic Terms, the term 'occlusal index' is used to describe "a rating or categorizing system that assigns a numeric score or alphanumeric label to an individual's occlusion" (AAO Glossary, 2012). Accordingly, occlusal indices include any system that attempts to assess, classify or rank occlusions (**Table 2.2**). In 1995 Shaw categorized occlusal indices according to their original purpose into 5 different categories: diagnostic, epidemiological, treatment need, treatment outcome, and treatment complexity indices (Shaw et al., 1995). It is important to note however, that indices developed for a certain category have often also been used to assess another. For example, Summer's occlusal index which is primarily designed for epidemiological purposes has also been used to assess orthodontic treatment need as well as outcome.

Table 2.2: The different types of occlusal indices

Diagnostic classification	**Angle classification system** (Angle, 1899) **Incisal categories** (Ballard & Wayman, 1965) **Five-point system** (Ackerman & Proffit, 1969)
Epidemiologic indices	**Index of Tooth Position** (Massler & Frankel, 1951) **Mal-alignment Index** (Van Kirk & Pennell, 1959) **Occlusal Feature Index** (Poulton & Aaronson, 1961) **The Bjork method** (Björk et al., 1964) **Summer's occlusal index** (Summers, 1971) **The FDI method** (Baume, 1974) **Little's irregularity index** (Little, 1975)
Treatment need indices	**Handicapping Labio-lingual Deviation Index (HLD)** (Draker, 1960) **Treatment Priority Index (TPI)** (Grainger, 1967) **Swedish Medical Board Index (SMBI)** (Linder-Aronson, 1973) **Dental Aesthetic index (DAI)** (Cons, Jenny, & Kohout, 1986) **Index of Orthodontic Treatment Need (IOTN)** (Brook & Shaw, 1989) **Index of Complexity, Outcome and Need (ICON)** (Daniels & Richmond, 2000)
Treatment outcome indices	**Peer Assessment Rating index (PAR)** (Richmond et al., 1992) **Index of Complexity, Outcome and Need (ICON)** (Daniels & Richmond, 2000) **Objective grading system (OGS)** (Casko et al., 1998)
Treatment complexity indices	**Index of Complexity, Outcome and Need (ICON)** (Daniels & Richmond, 2000) **Discrepancy index (DI)** (Cangialosi et al., 2004) **Index of Orthodontic Treatment Complexity (IOTC)** (Llewellyn, Hamdan, & Rock, 2007)

Diagnostic indices or classifications, such as angles or Ackerman and proffit classification are used to accurately describe occlusions, allowing for ease of communication between orthodontists (Ackerman & Proffit, 1969; Angle, 1899). Epidemiological indices on the other hand record different occlusal traits in a precise way to allow the estimation of malocclusion prevalence in a given population. An example of epidemiological indices is the Bjork registration method, which is widely used to record the prevalence of malocclusion in different countries (Björk et al., 1964). Both diagnostic and epidemiological indices are used to describe occlusion either at the individual level or the population level. In fact, a lot of diagnostic indices have been used for epidemiological studies or incorporated in different epidemiological indices, since both categories are extremely related.

Unlike the diagnostic and epidemiological indices, orthodontic treatment indices (need, complexity, and outcome) are instruments developed to measure a construct that is not directly observable. Because these constructs are not directly measurable, their outcomes should be validated. Numerous indices been developed to measure the constructs of treatment need, complexity, and outcome (Brook & Shaw, 1989; Daniels & Richmond, 2000; Richmond et al., 1992). However, only few studies have attempted to appropriately validate the outcome of such indices. These indices usually assign a score for each trait or component of the malocclusion, and these are subsequently weighted to calculate the overall score of the index. Orthodontic need indices were developed to prioritize malocclusions according to the need for orthodontic treatment. They classify a list of conditions or traits in ordinal categories that denote the extent to which treatment is considered necessary. Similarly, treatment complexity indices are developed to specifically measure the complexity of the treatment needed, and outcome indices to assess the changes or the end results of treatment.

2.2.2. Requirements for an ideal index

The World Health Organization report in 1966 outlined the requirement for an ideal occlusal index in 9 points (WHO, 1966) (**Table 2.3**). Those requirements can be summed into 4 main criteria; the ideal index should be an objective, reliable, valid, and simple scale. As with any diagnostic test or system of measurements, the most fundamental requirements of an occlusal index are validity and reliability. Validity is the extent to which a conclusion or measurement is accurately measuring what it is intended to measure, whereas reliability (reproducibility) is the ability of an assessment tool to produce consistent results; therefore, a reliable occlusal index should produce the same result or score for a particular malocclusion. There are many studies assessing the reliability of different indices and comparing them to each other; nevertheless, validation studies are much less.

Table2.3: Requirements for an ideal index (WHO, 1966)

1	Classification is expressed by a finite *scale* with definite upper and lower limits
2	The index should be equally *sensitive* throughout the scale.
3	The score should *correspond* closely with the clinical importance of the disease stage it represents.
4	Index value should be amendable to statistical *analysis*.
5	The index must be *reproducible*.
6	The index should also be *simple*, accurate and yield itself to modification for the collection of data.
7	The examination procedure should require a *minimum* of *judgment*.
8	The index should be simple enough to permit the study of a large population without undue cost in *time* or *energy*.
9	The examination required should be performed quickly, to evidence a group variation.

Many authors have advocated that using experts' opinion as a gold standard is an accepted technique in validating occlusal indices (Beglin et al., 2001; Brook & Shaw, 1989; Younis et al., 1997). In fact, the primary method used in validation studies is to compare the occlusal index scores with the quantitative assessment by orthodontists using Likert scales (Bagozzi & Phillips, 1982; Richmond, Daniels, Fox, & Wright, 1997). Illustrating the association between local provider's opinion and the objective classification by an occlusal index in a certain geopolitical location is considered a crucial first step in the acceptance of any index by the national clinical community (Younis et al., 1997). Researchers recommend that occlusal indices (particularly treatment need indices) should be validated in different countries or geographical locations to establish proper national cutoff points (Richmond & Daniels, 1998a; Younis et al., 1997). It has even been stated that "the effect of having a local panel of experts should not be underestimated" (Beglin et al., 2001). Different cutoff points of occlusal indices have indeed been reported in different countries (Firestone et al., 2002). It is important to note however, that to the best of our knowledge, no country in our region has validated occlusal indices.

Concerning simplicity, the index should be easy to learn and rapid to administer. Numerous indices like the IOTN, PAR, ICON, and others were reported to take short time for completion (Beglin et al., 2001; Cardoso et al., 2011). However, the majority of these statements were based on experience rather than on objective measures, with no actual study recording the time it takes for the most commonly used indices to be administered and comparing them to each other. Such information is of extreme importance in epidemiological settings where minimal time differences can have an impact.

2.3. Orthodontic treatment need indices

The American Association of Orthodontists (AAO) in 2015 re-defined Medically Necessary Orthodontic Care (MNOC) as "the treatment of a malocclusion that compromises the patient's physical, emotional, or dental health. This treatment should be based on a comprehensive assessment and diagnosis done by an orthodontist, in consultation with other health care providers when indicated" (J. G. Ghafari, 2016). This definition can be viewed from two perspectives; first is that treatment need is not only based on dental health, but also on physical and emotional health; second is that the assessment of treatment need should be done by an orthodontist. Accordingly, an ideal need index should reflect the expert opinions on the assessment of orthodontic treatment need that is based on physical, emotional, and dental health.

Treatment need indices are used to prioritize malocclusions according to the need for orthodontic treatment. Typically, there is a certain cutoff point for each index; which is the lowest possible score that allows the identification of patients in need for treatment (Grippaudo et al., 2008). These cutoff scores are determined by orthodontic experts using sophisticated psychometric methods. The perception of treatment need can be different between different countries, thus affecting the cutoff point of which the treatment is indicated (Richmond & Daniels, 1998a). For that reason, need indices are recommended to be validated in the geographical location of its application to optimize its cutoff points according to the national standards (Borzabadi-Farahani, 2011; Firestone et al., 2002; Younis et al., 1997).

2.3.1. Index of orthodontic treatment need (IOTN)

While numerous indices have been developed (**Table 2.1**), IOTN is probably the most common index for assessing the need for orthodontic treatment (Borzabadi-Farahani, 2011). This index ranks malocclusion according to the consequences of various occlusal traits on both dental health and perceived esthetics. The Dental Health Component (DHC) of IOTN is loosely based on the Swedish Medical Board index and can be applied clinically or to study models, albeit different protocols are required for each application (Linder-Aronson, 1973; Shaw et al., 1995). Compared to the clinical protocol, the dental cast protocol is used when there is no clinical information, and it always assumes the worst scenario (Richmond et al., 1992b). DHC has five grades, ranging from grade 1 (no need) to grade 5 (very great need), and the average time required to record it is approximately 1 minute (Cardoso et al., 2011). All occlusal traits which are assumed to contribute to the satisfactory functioning of the dentition are graded; subsequently, the treatment need score of the malocclusion is given based on the grade of the worst occlusal trait (worst grade recorded).

The aesthetic Component (AC) of ITON is composed of a 10- point scale, illustrated by a series of clinical intra-oral photographs that have been rated according to the attractiveness of the dental malocclusion by lay people. In fact, it is based on the SCAN scale, where the photographs of 12-year-olds were used in a large multidisciplinary survey (Evans & Shaw, 1987). The rating for the AC is allocated based on the overall dental attractiveness rather than specific similarity in the malocclusion. It gives an esthetic indication of the patient's treatment need, and it attempts to reflect the psychosocial need for orthodontic treatment. Interestingly, only a moderate agreement is reported between the AC and DHC of ITON with kappa statistic for diagnostic agreement of 0.55 (Borzabadi-Farahani & Borzabadi-Farahani, 2011a). This moderate agreement is present, because the AC assesses the malocclusion only in the frontal view; therefore, limiting the evaluation of sagittal problems. Additionally, some occlusal problems have more of dental health implications rather than aesthetic impairments, such as cross-bites, missing posterior teeth, deep overbites, and others (Fields et al., 1982).

The reliability and validity of both the DHC and AC as treatment need indices have been verified (Richmond et al., 1993; Burden & Holmes 1994; Burden et al., 1994). The IOTN was validated originally against a panel of 74 dentists in Great Britain. The panel were extremely divided on what constitute dental health need for orthodontic treatment (Brook & Shaw, 1989; Shaw et al., 1995). Despite this low level of agreement, IOTN scores for both DHC and AC were significantly correlated with the average panel scores (Spearman's correlation coefficient of 0.64 and 0.86 respectively) (Richmond et al., 1995). In the same study, cutoff points of grade 8 for the AC and grade 4 for DHC were proposed to reflect definitive need for treatment. Nevertheless, different cutoff points have been suggested by other studies. Two studies assessed the validity of the IOTN by a group of American orthodontic specialists without including general dentists, and both suggested to optimize the cutoff points of DHC to 3 and AC to 5 (Beglin et al., 2001; Younis et al., 1997). The IOTN was assessed several times against different indices and was validated in several countries like the United States, United Kingdom, Brazil, Slovenia, and Sweden (Cardoso et al., 2011; Firestone et al., 2002; Fox et al., 2002; Johansson & Follin, 2009; Johnson et al., 2000; Ovsenik & Primožič, 2007; Younis et al., 1997).

2.3.2. Modified IOTN

The main purpose of treatment need indices is to identify those individuals classified as having a definite need for orthodontic treatment. Correspondingly, since the IOTN was specifically developed to assess treatment need rather than complexity or outcome, there is no benefit of placing the malocclusion into different need categories. Therefore, Burden et al modified IOTN from a five-grade scale into a two-grade scale (1 = definite need for treatment; 0 = no definite need for treatment) (Burden, Pine, & Burnside, 2001). They did this by combining the borderline need and no need groups of the DHC into one group (no definite need for treatment), and by grouping the remaining two needs group into another group (definite need for treatment). Similarly, grade 8,9, and 10 of the AC of IOTN were combined into one group of definite need for treatment, while the remaining grades were considered to have no definitive need for orthodontic treatment (**Table 2.4**).

Interestingly, the modified IOTN did not only simplify IOTN, but also improved its reliability for both the dental health and aesthetic components (Burden et al., 2001). This has made it more suitable for epidemiological studies. Nevertheless, the modified IOTN should be considered in light of its limitation, unlike the IOTN it cannot prioritize treatment access or financial support on a continuous scale across the whole spectrum of malocclusions.

Table 2.4: Modified IOTN grades

MODIFIED IOTN	DHC OF IOTN	AC OF ITON
NO DEFINITIVE NEED	Grades :1, 2, 3	Grades: 1, 2, 3, 4, 5, 6,7
DEFINITIVE NEED	Grades: 4, 5	Grades: 8,9,10

2.3.3. Index of Complexity, Outcome, and Need (ICON) as a need measure

The development of the ICON was based on the judgments of 97 orthodontic specialists from 8 European countries and the United States (Daniels & Richmond, 2000; Richmond & Daniels, 1998a, 1998b). The orthodontists evaluated 240 pre-treatment dental casts to assess the amount of dental need, and 98 pairs of pre- and post-treatment casts to assess the treatment outcome. The ICON is the only index that incorporates an aesthetic score in its overall evaluation, rather than a separate component as in the AC of IOTN. It consists of five components: aesthetic component (the same AC of IOTN), crowding/spacing component, cross-bite component, over-bite component, and anteroposterior buccal component (Daniels & Richmond, 2000). The various components are measured, scored, weighted, and summed to produce the final ICON score.

The ICON is unique for being a multifunctional index developed to measure treatment need, complexity, and outcome. Several studies assessed the ICON and concluded that it is a reliable and valid index for assessing orthodontic treatment need (Daniels & Richmond, 2000; Firestone et al., 2002; Liao et al., 2012). The recommended cutoff point for assessing need is 43; therefore, a score of 44 or greater indicates need for treatment (Daniels & Richmond, 2000). When compared to other need indices, the ICON treatment need threshold is lower than the DHC and AC of IOTN (Farahani & Eslamipour, 2010; Theis, Huang, King, & Omnell, 2005). Similar to the IOTN, different cutoff points for the ICON have been suggested in various countries; for example, Dutch orthodontists advised on changing the cut-off point from 43 to 52, whereas Chinese orthodontists recommended lowering it to 29 (Liao et al., 2012; Louwerse, Aartman, Kramer, & Prahl-Andersen, 2006).

The ICON and the DHC of IOTN are highly correlated and have good agreement in estimating treatment need (K = 0.78) (Borzabadi-Farahani & Borzabadi-Farahani, 2011). In fact, it has been stated that the ICON is a good substitute for the IOTN and the DAI in assessing need (Farahani & Eslamipour, 2010; Onyeaso & Begole, 2007). Owing to its extra ability in assessing the orthodontic treatment complexity and outcome of malocclusions, the ICON does offer some advantages over the IOTN and the DAI. At the same time, it does not suffer from the DAI deficiencies in the assessment of posterior cross-bites, impacted teeth, and deep overbite (Danyluk, Lavelle, & Hassard, 1999; Onyeaso & Begole, 2007). Moreover, recording of the ICON is significantly easier and takes less time to calculate than the DAI (Daniels & Richmond, 2000). Similar to other indices there are potential limitations with using the ICON; the index is heavily weighted for the aesthetic component (x7) which relies on subjective rating/matching of photographs rather than on objective measures. This can possibly reduce the reliability of the index by affecting both intra and inter-examiner agreements (Koochek, Shue-Te Yeh, Rolfe, & Richmond, 2001; Savastano et al., 2003).

2.3.4. Facial Aesthetics Index (FAI)

Aesthetic improvement is one of the most commonly reported reasons for seeking orthodontic treatment (Birkeland et al., 1999; Trulsson et al., 2014). As a matter of fact, psychosocial consequences of malocclusion due to unacceptable esthetics maybe as serious or even more serious than biological problems (Stricker et al., 1979). In view of the previously mentioned AAO definition of medically necessary orthodontic treatment, a true treatment need index should not only address the dental health component, but also the dental and facial aesthetic components.

Facial aesthetic assessment is crucial in orthodontic diagnosis and treatment planning, and it should be included in the evaluation of orthodontic treatment need and outcome (Kiekens et al., 2005). Albeit, none of the present occlusal indices incorporate soft tissue measurements or take into account facial aesthetic needs of treatment (Sundareswaran & Ramakrishnan, 2016). It is only recently that a Facial Aesthetic Index (FAI) was developed to complement occlusal indices in addressing such necessity (Sundareswaran & Ramakrishnan, 2016). The FAI is intended as an additional screening tool to complement existing occlusal need indices in generating more information about the patient; thus, resulting in a more comprehensive measure of orthodontic treatment need. It is based on two sets (female and male) of eight soft tissue profile variations generated by altering: profile angle, lips relation, and lip competence. It has demonstrated excellent reliability and accuracy in determining different profiles into their prospective group. However, no study was done to assess its validity in measuring orthodontic treatment need.

2.4. Orthodontic treatment complexity indices

It is important to note the subjectivity of the construct of complexity, since it invites personal perception and experience. In orthodontics, any differences in treatment philosophies, appliances, or training can lead to different perceptions of treatment difficulty: thus, some malocclusions that are considered difficult to treat by some practitioners may be considered easy by others. For those reasons, treatment complexity is an elusive entity that is hard to measure quantitatively. In fact, several definitions for treatment complexity or difficulty have been proposed. For instance, Richmond defined orthodontic treatment complexity as the factors that reduce treatment success (Richmond et al., 1997). Bergström and Halling defined treatment difficulty as the effort needed to establish proper tooth relationships (Bergström, Halling, Huggare, & Johansson, 1998), whereas Pae defined difficulty as the probability of attaining an ideal occlusion when all treatment options are available (Pae et al., 2001). It has been suggested that complexity and difficulty in orthodontics are synonymous; both are a measurement of effort and skill, whereas complexity and severity are not.

While an occlusal index of treatment complexity offers many potential benefits, there are also several potential challenges. Such assessments should be amenable to modifications, as treatment complexity changes with advances in technology (Holman, Hans, Nelson, & Powers, 1998). Moreover, it should segregate the treatment complexity of the malocclusion itself from patient (compliance, growth potential, general and oral health) or provider (training and expertise) determinants of treatment complexity (Otuyemi & Jones, 1995). There are also the concerns whether complexity index should be directed to reflect the provision of ideal or optimal (most appropriate realistic) services for a patient.

2.4.1. Discrepancy Index (DI)

Knowing that the severity of malocclusions and treatment complexity are related but not the same; the ABO formed a committee in 1998 to find a method of evaluating the treatment complexity levels of cases presented for the board clinical examination. After comprehensive meetings, discussions, and field testing, the discrepancy index (DI) was developed as a measure of treatment complexity (Cangialosi et al., 2004). Since then, all applicants are required to measure the DI on their board cases to fulfill the cases requirement for phase III of the certification process. Orthodontic models and lateral cephalometric radiographs are needed to determine the Discrepancy Index. In fact, the DI is the only index that incorporates cephalometric evaluation of the underlying skeletal structures in its occlusal assessment. It measures overjet, overbite, anterior open bite, lateral open bite, crowding, molar occlusion, lingual posterior cross-bite, buccal posterior cross-bite, ANB angle, IMPA, and SN-GoGn angle (Cangialosi et al., 2004). Even though the DI has been used since 2004 as the ABO's official measure of complexity, to the best of our knowledge no published study has examined the psychometric properties of the DI as a complexity measure.

2.4.2. Index of Complexity, Outcome, and Need (ICON) as a complexity measure

As stated earlier in this chapter, the ICON was designed not only to measure the need, but also to measure the complexity and the outcome of orthodontic treatment. It measures complexity by using a 5-point grading scale based on pre-treatment models scores: simple, mild, moderate, difficult, and very difficult (Daniels & Richmond, 2000) (**Table 2.5**).

Only few studies aimed at assessing the ICON as a complexity measure. The first attempt to validate the ICON as a measure of complexity was done by Savastano et al in 2003. Their study sample consisted of 100 pairs of pre- and post-treatment study models, which were rated by a panel of orthodontists practicing in Ohio. Even though the authors reported that the ICON is a valid measure of complexity, their conclusion should be taken with caution. Only a moderate agreement was reported between the ICON and the average panel opinion (k= 0.52), as well as between the different members of the panel (k= 0.50), which decreases the validity of the ICON as a complexity measure (Savastano et al., 2003).

Another validation study was done on an Iranian population, with the assessments of five experts used as the gold standard (Torkan et al., 2015). Unlike the previous study, complete orthodontic diagnostic records of each participant were provided for the orthodontists' assessment. Albeit, they also reported a moderate inter-examiner agreement of the panel members in assessing complexity (k= 0.51). Importantly, they reported that the recommended ICON ranges for the different complexity grades are not applicable to their society. Therefore, the ICON was not a valid measure of complexity in their sample, and they recommended that further studies are required to set new cutoff points or ranges for assessing their treatment complexity standards (Torkan et al., 2015). Both studies could not provide enough evidence to support the validity of the ICON as a measure of complexity for the required orthodontic treatment.

Table 2.5: ICON complexity measure grades

ICON score	**ICON complexity grade**
<29	Easy
29-50	Mild
51-63	Moderate
64-77	Difficult
>77	Very difficult

2.5. Orthodontic treatment outcome indices

Reliable objective assessment of orthodontic treatment outcome is a challenging endeavor that can be done either by directly measuring the post-treatment end results, or by measuring the occlusal changes during orthodontic treatment. Even though several indices have been used to assess treatment outcome, only few were validated. The American Board of Orthodontists developed the Objective Grading System (OGS) to accurately describe the quality of final results and determine whether the post-treatment results are sufficient for passing the board certification process (Casko et al., 1998). When compared to other indices, the OGS application requires much more experience and time, and it does not take into consideration the severity of the pretreatment malocclusion. As a result, its use has been limited in epidemiological studies or national surveys. Two of the most commonly used treatment outcome indices are the Peer Assessment Rating (PAR) and the Index of Complexity, Outcome and Need (ICON).

2.5.1. Peer Assessment Rating (PAR)

In addressing the need for an objective method of measuring the amount of malocclusion improvement and treatment efficacy, a series of meetings between a group of 10 experienced orthodontists called the British Orthodontic Standards Working Party were arranged. Consequently, the peer assessment rating (PAR) index was developed to objectively record the amount of malocclusion at any stage of treatment (Richmond et al., 1992). The PAR index is currently one of the most widely used occlusal indices to quantify the severity of a malocclusion and treatment outcome. Similar to other indices, it is a scoring system, where individual scores of different occlusal traits are summed to calculate an overall score (**Table 3.2**). Subsequently, the changes between the pre- and post-treatment scores are calculated to reflect the amount of success and improvement of the orthodontic treatment. The PAR index can be measured both clinically and on study models, and has demonstrated a high reliability with an interclass correlation coefficient of 0.83. The developers reported that it takes approximately 2 minutes for each set of models to be analyzed by the PAR while other authors reported more time (Onyeaso & Begole, 2007; Richmond et al., 1992). Nevertheless, no actual study recorded the exact time needed to measure PAR.

The PAR index was originally validated against the opinion of a panel of 74 British dentists (same panel that validated IOTN as need measure). Although the PAR index demonstrated excellent reliability, there was a poor level of agreement between the dentists of the panel (gold standard) in validating the assessment of malocclusion deviation from normal occlusion. The poor agreement was related to the fact that the 74 dentists were representing five different groups of practitioners; 22 consultant orthodontists, 22 specialist practitioners, 15 general practitioners, 11 community dentists, and 4 others (Richmond et al., 1992). Consequently, the validity of the present weightings of different components of the PAR index is affected by this poor agreement of the gold standard of which the PAR index was developed and validated against. Nevertheless, the agreement between the average panel opinion and the PAR index was high and by weighting the individual components of the index, it was made even better. It should be noted that only dental casts were given to the panel for assessment, thus the success of treatment was measured without taking into consideration any facial or cephalometric changes. The developers of the index stated that the weightings could be changed to reflect treatment standards in other countries; thus, advocating for the re-validation of the PAR index in different countries to reflect their current national standards or future standards of orthodontic treatment (Richmond et al., 1992).

2.5.2. Index of Complexity, Outcome, and Need (ICON) as an outcome measure

In assessing the ICON as a measure of orthodontic treatment outcome, the developers found that a post-treatment total score of 30 or less indicates an acceptable end treatment occlusion (Daniels & Richmond, 2000). Moreover, they suggested a certain formula to measure the outcome by assessing the amount of malocclusion improvement: Improvement grade = pre-treatment score – (4 x post-treatment score). However, in validating the ICON as an outcome measure, the authors reported that the improvement section could not be validated, and advised on reassessing this particular section with more variable subjects (Savastano et al., 2003). It has been estimated that the application of the ICON is faster than PAR with approximately 3 minutes spent to analyze each set of models (Onyeaso & Begole, 2007).

When compared to other outcome indices, the ICON was reported to be more stringent than the PAR and less than the OGS in assessing the end results of orthodontic treatment (Onyeaso & Begole, 2007). In the same study, the ICON post-treatment scores were poorly correlated with the PAR and OGS scores ($r = 0.35$, 0.27 respectively). However, its pre-treatment scores were highly correlated with the PAR ($r = 0.68$). This variation in the strength of correlation between the pre- and post-treatment scores was explained by the fact that treatment outcomes were consistently good. Therefore, this little variation in the post-treatment group lead to poor correlation results. Moreover, they concluded that the ICON was as appropriate as the PAR in assessing outcome, and that it can be a reasonable alternative for the PAR and OGS

(Onyeaso & Begole, 2007). It should be noted that these conclusions were based on correlations rather than validation studies and therefore do not provide enough information on which is the better index in assessing the construct of treatment outcome.

2.6. Significance

Despite the great numbers of occlusal indices, there have been many disagreements among authors, with no universally accepted index for measuring malocclusion (Borzabadi-Farahani, 2011). The following conclusions may be drawn regarding misuse or misinterpretation of existing indices:

- Not all indices have been validated for their use in orthodontics. In fact, numerous indices that have been developed and validated for one specific category have also been used for another without validation. Thus, various occlusal indices are being used without realizing the degree to which they measure the stated objective.
- Even if an occlusal index, in particular a treatment need index, was validated in a certain country, it should be re-validated in the country of application to optimize its cutoff point.
- Most of the reported validation studies of occlusal indices were carried only on dental models and thus do not reflect orthodontic treatment standards on the basis of facial aesthetics, skeletal discrepancy, or periodontal requirements. In fact, it has been recommended that the Aesthetic Component of the IOTN and the ICON should not be applied to dental models (Buchanan, Downing, & Stirrups, 1994), and that it can be replaced by a short version of soft and hard tissue measurements.

Despite the wide agreement on the DI as a measurement of treatment complexity, published studies are not available that examined its psychometric properties as a complexity measure. Similarly, the newly introduced Facial Aesthetic index has not been validated yet as a supplement to treatment need indices. Additionally, no country in the Mediterranean/Middle Eastern region has re-validated need indices to establish proper national cut-off points, despite their use in national and regional studies.

For all the aforementioned reasons, we aimed to evaluate and re-validate the most commonly used indices (PAR, ICON, IOTN, DI) as well as the recently published Facial Aesthetic Index (FAI), and compare among with each other and with different modifications of ours. The goal is to find the optimal index for the different categories (constructs) and to optimize their cutoff points to meet the current standards of treatment. The evaluations and comparisons of the different indices will be based on their reliability, validity, and simplicity. The indices will be validated separately as orthodontic treatment need measures, complexity measures, and outcome measures. Since these indices have been used for different purposes, each index will be assessed in the three categories, and not only in their original category. Accordingly, two overriding panels of significance may be defined:

- by determining if any of the evaluated indices can be used across the original objective it was designed for, recommendations for application of the pertinent index will emerge for proper utilization in epidemiologic or clinical assessments.

- by slicing out the various components, dental, facial (esthetic) and cephalometric, we could determine which instrument(s) better reflect the treatment need, complexity, or outcome of a malocclusion.

2.7. Specific objectives

The aim of this study is to evaluate the validity, and reliability of the: PAR, IOTN, ICON, and DI, with and without the FAI, in assessing orthodontic treatment need, complexity, and outcome.

2.7.1. Objectives of the treatment need study

- Validate the IOTN and the ICON (with and without FAI) as measures of orthodontic treatment need, against the assessment of a national panel of experts.
- Assess the validity of the DI and the PAR (with and without FAI) if used as measures of orthodontic treatment need, against the assessment of a national panel of experts.
- Explore the possibility and the validity of indices modifications as well as adding soft and hard tissue measurements.
- Establish proper national cutoff points for treatment need indices, and assess the effect of the geographical location of orthodontist on need assessment.
- Assess the reliability of the DI, the IOTN, the ICON, the PAR, and the FAI, along with the associations between the different indices.

2.7.2. Objectives of the treatment complexity study

- Validate the DI and the ICON as measures of orthodontic treatment complexity, against the assessment of a panel of experts with recognized credentials.

- Assess the validity of the IOTN and the PAR if used as measures of orthodontic treatment complexity, against the assessment of a panel of experts with recognized credentials.
- Explore the possibility and the validity of modifying the studied indices as well as the value of adding the FAI.
- Explore the association of different cephalometric measurements with the complexity of the malocclusion.
- Measure the association of the different indices with the panel assessment of treatment duration as a secondary measure of complexity.

2.7.3. Objectives of the treatment outcome study

- Validate the ICON, and the PAR as measures of orthodontic treatment outcome (improvement section), against the assessment of a panel of experts.
- Validate the ICON, and PAR as measures of orthodontic treatment outcome (outcome success section), against the judgment of a panel of experts.
- Assess the validity of the DI and the IOTN, if used as measures of orthodontic treatment outcome (improvement section), against the assessment of a panel of experts.
- Assess the validity of the DI and IOTN, if used as measures of orthodontic treatment outcome (outcome success section), against the judgment of a panel of experts.
- Explore the possibility and the validity of modifying the studied indices as well as the value of adding the FAI, in assessing both sections (improvement and outcome success sections).

CHAPTER 3

METHODOLOGY

The research methodology is detailed in this chapter, including sample selection and recruitment process, the procedure used in collecting the data, statistical procedures used to analyze the data, and the ethical considerations in this research.

3.1. Research design

This is a construct validation study that targeted the validation of the most commonly used occlusal indices as measures of orthodontic treatment need, complexity, and outcome. Accordingly, data were collected from the pre- and post-treatment records and files of the participants. The occlusal indices scores were assessed against the evaluations of a panel of orthodontists with recognized credentials.

3.2. Sample selection

Participants were recruited from the patient population of the Division of Orthodontics and Dentofacial Orthopedics at the American University of Beirut Medical Center AUBMC. A total of 101 participants were randomly selected, from a list containing all patients whose orthodontic treatment was completed and appliances removed by the orthodontic residents who graduated in 2015. Participants with craniofacial anomalies or missing records (x-rays or dental models) were excluded from the study.

3.3. Data collection

The sampling process involved the participant's pre-treatment and post-treatment dental casts, intraoral and extraoral photographs, and x-rays, as well as the treatment files. The sample was collected and de-identified with serial numbers assigned to the subjects enrolled in the study, as dictated by the IRB approval. The coded records were then provided to the investigator who proceeded with data collection. The data collection procedures were conducted in two stages. The first stage included medical records assessment to collect different clinical information, as well as radiographical and dental measurements to compute different occlusal indices. The second stage of data collection included the case assessments by a panel of 4 experts. The pre- and post-treatment dental models, x-rays, photographs and relevant clinical information were provided to the panel. Panel recruitment was based on strict inclusion criteria:

- Graduates of an orthodontic program accredited by the American Dental Association
- At least 10 years of experience as an orthodontist
- At least 10 consecutive years of practicing in Lebanon
- Involvement in clinical teaching in a postgraduate orthodontic program

3.4. Measures

Several indices were used to assess orthodontic treatment need, complexity and outcome. These indices were validated against the panel assessment measures (gold standard). Moreover, cephalometric measurements were recorded for skeletal and soft tissue analysis. Other related clinical variables were also collected to further describe the level of complexity of the needed orthodontic treatment. The occlusal indices were measured on both the pretreatment and posttreatment dental casts. After measuring the dental cast and cephalograms for the first time, 10% of the sample were re-measured two months later to assess the reliability of different indices.

Four of the most commonly used indices were chosen to be validated in this study: IOTN, PAR, ICON, DI. They were chosen according to the following sequence and criteria: First, a list was tabulated of all the need, complexity, and outcome indices that have been used in the literature within the last two decades. According to their citation number, the top three indices were selected (IOTN: 833, PAR: 597, ICON: 363; these numbers were adopted from Google scholar). The Discrepancy Index DI was later added to the list because it is the only occlusal index that incorporates cephalometric measurements, and currently used in the American board certification process. By following the mentioned criteria, we actually ended up with one index from each purpose category: treatment need category (IOTN), complexity category (DI), outcome category (PAR), and the ICON which is the only multifunctional index that is developed to assess the three categories. The Facial Index is the only index that assesses facial aesthesis; therefore, it was chosen to as a supplemental index to represent the facial aesthetics component, which is missing from all other indices. Prior to the initiation of the study, the examiner was calibrated to ensure precise assessment of the occlusal indices.

3.4.1. Index of Orthodontic Treatment Need

Using the photographs of the AC scale of IOTN (**Figure 3.1**), both the pre- and post-treatment intra-oral photographs of participants were rated for dental attractiveness from 1-10 (Brook & Shaw, 1989). The ratings were given according to the attractiveness level of the dental occlusion without attempting to closely match the malocclusion of the participant to the picture.

Fig 3.1: Aesthetic Component scale of IOTN index (Brook & Shaw, 1989)

The DHC of IOTN was assessed on the initial and final occlusions. Five occlusion features were assessed on each study cast: missing teeth; overjet; cross-bite; displacement of contact points (largest displacement only); and overbite (Brook & Shaw, 1989). Since all the required clinical information was available through the participant files and x-rays, the regular clinical protocol was used for assessing the DHC rather than the dental cast protocol which always assumes the worst-case scenario on missing information. For the DHC only the highest scoring occlusal trait is recorded (**Table 3.1**). Please note, that the measurement of contact point displacement was based on the maximum displacement between any teeth only, rather than the sum of all present displacements.

Table 3.1: Orthodontic treatment need defined by IOTN-DHC

DHC	**Grade 5** (Very great need)	**Grade 4** (great need)	**Grade 3** (boarder line need)	**Grade 2** (little need)
Missing teeth	Extensive hypodontia with restorative implications (more than 1 tooth missing in any quadrant) requiring pre-restorative orthodontics. Submerged deciduous teeth.	Less extensive hypodontia requiring prerestorative orthodontics or orthodontic space closure to obviate the need for prosthesis. Presence of supernumerary teeth.		
Over-jet	Increased overjet greater than 9mm. Reverse overjet greater than 3.5mm with reported masticatory and speech difficulties.	Increased overjet greater than 6mm but less than or equal to 9mm. Reverse overjet greater than 3.5mm with no masticatory or speech difficulties Reverse overjet greater than 1mm but less than 3.5 with recorded masticatory and speech difficulties.	Increased overjet greater than 3.5mm but less than or equal to 6mm with incompetent lips. Reverse overjet greater than 1mm but less than or equal to 3.5mm.	Increased overjet greater than 3.5mm but less than or equal to 6mm with competent lips. Reverse overjet greater than 0mm but less than or equal to 1mm. Pre-normal or post-normal occlusions with no other anomalies (includes up to half a unit discrepancy.
Cross-bite		Anterior or posterior crossbites with greater than 2mm discrepancy between retruded contact position and intercuspal position. Posterior lingual crossbite with no functional occlusal contact in one or both buccal segments.	Anterior or posterior crossbites with greater than 1mm but less than or equal to 2mm discrepancy between retruded contact position and intercuspal position.	Anterior or posterior crossbite with less than or equal to 1mm discrepancy between retruded contact position and intercuspal position.
Displacement of contact points	Impeded eruption of teeth (except for third molars) due to crowding, displacement, the presence of supernumerary teeth, retained deciduous teeth and any pathological cause.	Displacements greater than 4mm. Partially erupted teeth tipped and impacted against adjacent teeth.	Displacements greater than 2mm but less than or equal to 4mm	Displacements greater than 1mm but less than or equal to 2mm.
Over-bite		Extreme lateral or anterior open bites greater than 4mm. Complete overbite with gingival or palatal trauma.	Lateral or anterior open bite greater than 2mm but less than or equal to 4mm. Deep overbite, complete on gingival or palatal tissues but no trauma.	Anterior or posterior openbite greater than 1mm but less than or equal to 2mm. Increased overbite greater than or equal 3.5mm without gingival contact.

3.4.2. Peer Assessment Rating

Five components were considered in the PAR assessment of pre- and post-treatment dental casts (Richmond et al., 1992). Details of the scoring of these components relative to the measured records are presented in *Table 3.2.*

- Anterior Segment:

The contact point displacement scores of the upper and lower anterior segments, from the mesial contact point of the canine on one side to the canine on the other side were recorded. Contact point displacements including crowding and spacing were measured as the shortest distance between the contact points of adjacent teeth parallel to the occlusal plane, and then scored from 1-4.

- Buccal Segment:

The right and left occlusion were scored in the antero-posterior plan, vertical plan and transverse planes. The buccal segment recording zone was from the distal of the canine to the last standing molar. In assessing antero-posterior relationship the quality of the interdigitation of cusp to embrasure was considered. At least two teeth with greater than 2mm of lateral open bite were needed for the vertical plane to be scored 1, otherwise the score is 0. In the transverse plane assessment, the existence of a cross-bite, tendency for cross-bite, or scissors bite was recorded.

- Over-jet:

Positive and negative overjet were recorded for all incisor teeth through the most prominent distance between the labial aspects of the upper and lower incisor edges. For recording the overjet, the periodontal probe was held parallel to the occlusal plane. If both anterior cross bite and overjet existed the scores were added together.

- Overbite:

Both deep bite and open bite were assessed in this section: over-bite by the amount of the deepest coverage of the upper incisors over the lower incisors, and open-bite by the distance from the mid-incisal edge of the most deviant upper tooth to the lower incisors. If both deep bite and open bite were present, the higher score of either one was recorded.

- Center-line:

If the difference between the upper and lower dental midlines was greater than one quarter of a lower incisor width, the center line discrepancy was scored as 1, and if it was greater than the width of a half lower incisor, it was scored as 2.

Table 3.2: The PAR scoring method and its components

PAR Components		**Scores**					
		0	**1**	**2**	**3**	**4**	**5**
Contact point displacement for anterior segments		0 – 1 mm	1.1 – 2.1 mm	2.1 – 4 mm	4.1 – 8 mm	> 8 mm	Impacted teeth
Buccal segment	Antero-posterior	Cusp to embrasure only	Any cusp relation up to cusp to cusp	Cusp to cusp			
	Transverse	No cross-bite	Cross-bite tendency	Single tooth in cross-bite	> 1 tooth in cross-bite	> 1 tooth in scissor bite	
	Vertical	No lateral open bite	Lateral open bite on at least 2 teeth >2 mm				
Over-jet	Increased	0 – 3 mm	3.1 – 5 mm	5.1 – 7 mm	7.1 – 9 mm	> 9 mm	
	Reversed	No cross-bite	>= 1 teeth edge to edge	1 tooth in cross-bite	2 teeth in cross-bite	> 2 teeth in cross-bite	
Overbite	Deep bite	<1/3 of lower incisor coverage	1/3 – 2/3 of coverage	2/3 - up to fully covered	Fully covered		
	Open bite	Edge to Edge	< 1mm	1 – 2 mm	2.1 – 3 mm	>= 4 mm	
Centre line		< 1/4 of deviation	1/4 – 1/2 of deviation	> 1/2 of deviation			

3.4.3. Index of complexity outcome and need

Similar to the other indices, the ICON was assessed on both pre- and post-treatment dental models. The index contains five components, all of which were scored using the criteria described by the developers (Daniels & Richmond, 2000) (**Table 3.3**).

- Dental Aesthetic component

The AC of ICON is the same AC of IOTN (**Figure 3.1**); therefore, the same scores were used for both. Once the score was obtained, it was multiplied by the weighting of 7, and subsequently added to other components.

- Crowding / Spacing component

Arch length discrepancies are measured only for the maxillary arch of each study cast. The sum of the mesio-distal crown diameters was compared to the available arch circumference, mesial to the last standing tooth on either side (not including third molars). Once the discrepancy had been calculated in millimeter, the difference was reduced to a 5-point score (**Table 3.3**). The presence of impacted teeth in either arch was directly scored 5 (the maximum score for this component), with no need for further measurement.

- Cross-bite component

The buccal segments of the study casts were examined for the presence of cross-bites, which are deemed to be present if a transverse relation of cusp to cusp or more exists, including both buccal and lingual cross-bites of one or more teeth. In the anterior segment, a cross-bite was present, when an upper incisor or canine is in edge-to-edge or lingual occlusion. A score of 1 was given for a cross-bite present in the posterior or anterior segments or both.

- Overbite component

Similar to the PAR, overbite was measured at the deepest part covering the incisor teeth, whereas open bite was measured to the mid-incisal edge of the most deviant upper tooth. If both traits were present only the highest raw score of either deep or open bite was counted.

- Buccal segment antero-posterior relationship

The antero-posterior relationship of canine, premolar, and molar teeth was assessed for each side of the study cast. Similar to the PAR, the quality of the buccal interdigitation is measured rather than the Angle classification. Each side was assessed and scored in 3 of 5 scoring categories. Scores for both sides were added together to provide an overall score for this component, and then multiplied by the weighting of 3, indicating the higher impact of the sagittal malocclusion on its severity.

Table 3.3: ICON scoring method and its components (Daniels & Richmond, 2000)

ICON Components	**Scores**						**Weight**
	0	**1**	**2**	**3**	**4**	**5**	
Aesthetic assessment	AC of IOTN Score 1 to 10						**7**
Upper arch crowding	< 2mm	2.1-5mm	5.1-9mm	9.1-13mm	13.1-17mm	> 17mm	**5**
Upper arch spacing	< 2mm	2.1-5mm	5.1-9mm	> 9 mm		Impacted teeth	
Crossbite	No crossbite	Crossbite present					**5**
Incisor open bite	Edge to edge	< 1mm	1.1-2mm	2.1-4mm	> 4mm		**4**
Incisor overbite	< 1/3 lower incisor coverage	1/3 to 2/3 coverage	2/3 up to fully covered	Fully covered			
Buccal segment A-P	Cusp to embrasure only	up to but not including cusp to cusp	Cusp to cusp				**3**

3.4.4. Discrepancy Index

Unlike other indices, cephalometric x-rays are needed to measure the DI. The detailed DI measurements and work form (**Figures 2.1**; **Appendix I**) were obtained directly from the ABO website (Clinical Examination Case Report Work File). The 12 target areas that were assessed in the DI included:

- Overjet

Overjet is measured in millimeters between antagonistic central or lateral incisors from the middle of the incisal edge of the more facially positioned tooth to the facial surface of the most lingual tooth. Then, the overjet is scored according to the criteria presented in Figure 3.2.

- Overbite

Similar to other indices, the deepest overbite was measured between two antagonistic anterior teeth (lateral or central incisors), and then scored from 0 – 5.

- Anterior open bite

Unlike other indices, open bite was measured for each anterior tooth (canine to canine) that is in an open bite relationship with its opposing tooth. It was measured from the mid-incisal edge of the upper tooth to the mid-incisal edge of the lower tooth. No points were scored for any tooth that was blocked out of the arch due to space deficiency or not fully erupted.

- Lateral open bite

It was measured from cusp to cusp of each maxillary posterior tooth (except third molars), that was in an open bite relationship ≥ 0.5 mm from its opposing tooth.

- Crowding

Only the most crowded arch was measured from the mesial contact point of the right first molar to the mesial contact point of the left first molar. If a tooth was missing, fractured or decayed, crowding was measured according to the treatment objectives.

- Occlusion

Measuring buccal occlusion in the DI differs from the ICON or PAR; it actually measures the amount of deviation from Class I molar occlusion rather than the quality of interdigitation. If the mesiobuccal cusp of the maxillary first molar occludes anywhere between the buccal groove and the mesiobuccal or distobuccal cusps (Class I to End On), no points were scored. If it occluded with the mesiobuccal (Class II end-to-end) or distobuccal (Class III end-to-end) cusps, the score was 2 pts per side. If the relationship was a full Class II or III, a score of 4 pts was given per side. Additional scores per mm were given for relationship that was beyond full Class II or III.

- Lingual posterior cross-bite,

Each maxillary posterior tooth with a buccal cusp that is > 0 mm lingual to the buccal cusp tip of the opposing tooth was scored 1 pt.

- Buccal posterior cross-bite,

Each maxillary posterior tooth with a palatal cusp that is > 0 mm buccal to the buccal cusp of the opposing tooth was scored 1 pt.

- Cephalometric measurements:
    - ANB angle: If the ANB angle was >6° or <-2°, 4 points were scored, and then additional point was added for each full degree above 6° or below -2°

- SN-MP angle: If the SN-MP angle was greater than 37°, 2 points were scored for each full degree above 37°. However, if the SN-MP angle was less than 27°, only 1 point was scored for each full degree below 27°.
- IMPA: If the Lower Incisor to MP angle was greater than 98°, then 1 point was scored for each full degree above 98°.

- Others

This category allows for the allocation of points to other conditions that may contribute to the complexity of the malocclusion, such as: a diastema of > 2mm, midline discrepancy, generalized spacing of > 4mm, ankylosis, transposition, and others (**Figure 3.2**).

EXAM YEAR ________ **ABO DISCREPANCY INDEX**

ABO ID # ________ **CASE#** _______ **PATIENT** ______________________________

TOTAL D.I. SCORE [ ]

For mm measures, round up to the next full mm.
Examiners will verify measurements in each category.

OVERJET

≥ 0 to < 1 mm (edge-to-edge)	= 1 pt
≥ 1 to ≤ 3 mm	= 0 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 3 pts
> 7 to ≤ 9 mm	= 4 pts
> 9 mm	= 5 pts
Negative Overjet (x-bite): 1 pt per mm per tooth	= ___pts
Total	[ ]

OVERBITE

> 1 to ≤ 3 mm	= 0 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 3 pts
Impinging (100%)	= 5 pts
Total	[ ]

ANTERIOR OPEN BITE

0 mm (edge-to-edge), 1 pt per tooth	=____pts
then 1 pt per mm per tooth	=____pts
Total	[ ]

LATERAL OPEN BITE

≥ 0.5 mm, 2 pts per mm per tooth	
Total	[ ]

CROWDING (only one arch)

≥ 0 to ≤1 mm	= 0 pts
> 1 to ≤ 3 mm	= 1 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 4 pts
> 7 mm	= 7 pts
Total	[ ]

OCCLUSAL RELATIONSHIP

Class I to End On	= 0 pts	
End-to-End Class II or III	= 2 pts per side	____pts
Full Class II or III	= 4 pts per side	____pts
Beyond Class II or III	= 1 pt per mm additional	____pts
Total		[ ]

30130315

LINGUAL POSTERIOR X-BITE

> 0 mm, 1 pt per tooth — Total [ ]

BUCCAL POSTERIOR X-BITE

> 0 mm, 2 pts per tooth — Total [ ]

CEPHALOMETRICS (See Instructions)

ANB ≥ 6° or ≤ -2°	@4pts	= ___
Each full degree > 6°	___x 1 pt	= ___
Each full degree < -2°	___x 1 pt	= ___
SN-MP		
≥ 38°	@2pts	= ___
Each full degree > 38°	___x 2 pts	= ___
≤ 26°	@1pt	= ___
Each full degree < 26°	___x 1 pt	= ___
$\bar{1}$ to MP ≥ 99°	@1pt	= ___
Each full degree > 99°	___x 1 pt	= ___
Total		[ ]

OTHER (See Instructions)

Supernumerary teeth	___x 1 pt	= ___
Ankylosis of permanent teeth	___x 2 pts	= ___
Anomalous morphology	___x 2 pts	= ___
Impaction (except 3rd molars)	___x 2 pts	= ___
Midline discrepancy (≥3 mm)	@ 2 pts	= ___
Missing teeth (except 3rd molars)	___x 1 pt	= ___
Missing teeth, congenital	___x 2 pts	= ___
Spacing (4 or more, per arch)	___x 2 pts	= ___
Spacing (mx cent diastema ≥ 2 mm)	@ 2 pts	= ___
Tooth transposition	___x 2 pts	= ___
Skeletal asymmetry(nonsurgical tx)	@ 3 pts	= ___
Addl. treatment complexities	___x 2 pts	= ___

Identify:

Total Other [ ]

Figure 3.2: The ABO discrepancy index scoring sheet (https://www.americanboardortho.com)

3.4.5. Facial Aesthetic Index

The FAI is intended as an additional screening tool to the existing indices in assessing treatment need (Sundareswaran & Ramakrishnan, 2016). The original published Facial Aesthetic Index charts for males and females (**Figure 3.3**) were used to assess the treatment need. The participants profile pictures (before and after treatment) were correlated to the most appropriate photograph in the FAI chart, and then scored accordingly

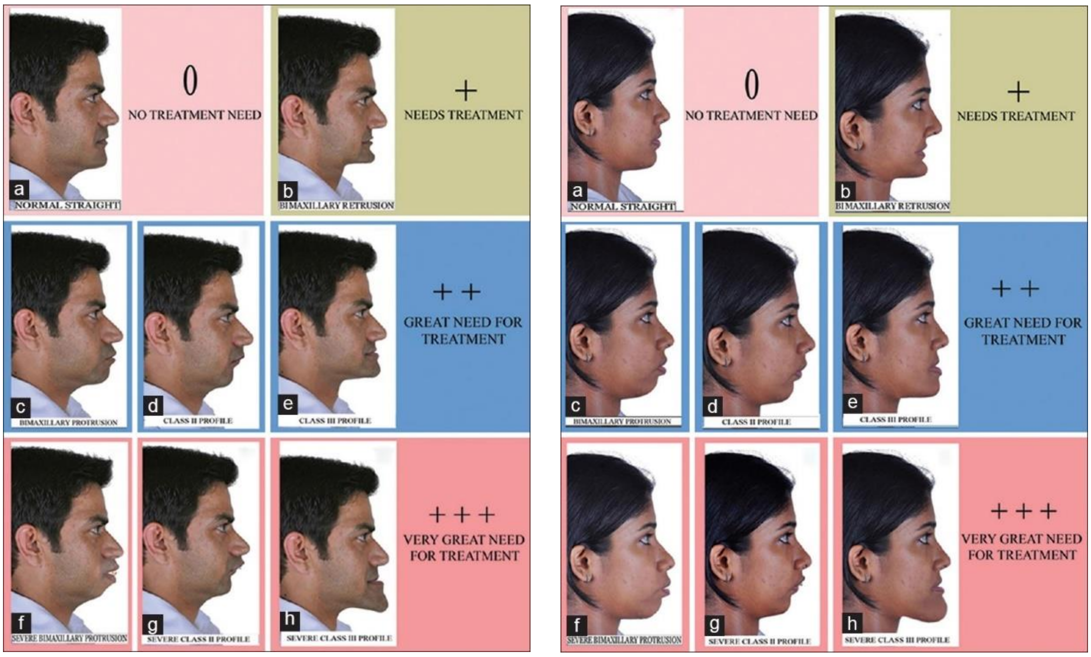

Figure 3.3: Facial Aesthetic Index charts for males and females adopted from the original article (Sundareswaran & Ramakrishnan, 2016).

3.4.6. Cephalometric measures

In addition to the cephalometric measurements mentioned earlier for the assessment of the DI (ANB, SN-Mp, IM-MP), other angles were measured and provided to the panel to ensure that all panel members relied on the same cephalometric analysis (see section 3.4.7). The pre- and post-treatment lateral cephalograms of all participants were digitized by the principal investigator using the Dolphin imaging software (Dolphin Imaging and Management Solutions, La Jolla, California). Linear and angular measurements were performed to assess the relations among cranial base, maxilla, mandible, incisors, and subnasal soft tissue analysis (**Figure 3.4, 3.5**). The measurements included were:

- ANB angle:

It measures the sagittal skeletal relationship, representing the anteroposterior discrepancy between the maxilla and the mandible.

- SNA and SNB angles:

These angles represent the relative anterioposterior position of the maxilla and mandible relative to the cranial base.

- MP-SN angle:

The angle measures the vertical skeletal relationship, representing the steepness of the mandibular plane in relation to the cranial base.

- LI- MP and LI-NB angles

These angles measure the axial inclination of the mandibular incisors relative to the mandibular border and the anterior projection of the mandible, respectively.

- UI-SN and UI-NA angles

These angles measure the axial inclination of maxillary incisors relative to the anterior cranial base and the anterior projection of the maxilla, respectively.

- Holdaway and Esthetic lines analysis

Both are soft tissue analyses to evaluate the harmony of the subnasal profile and have been used to describe the anteroposterior lip profile changes **(Figure 3.5, Table 3.4)**.

The lips are measured to the E-line, which extends from the tip of the nose to the soft tissue pogonion (Ricketts, 1968). Nose extension, the lips and the chin are measured to the H-line, which is drawn from the soft tissue pogonion to the vermillion border of the upper lip (Holdaway, 1983).

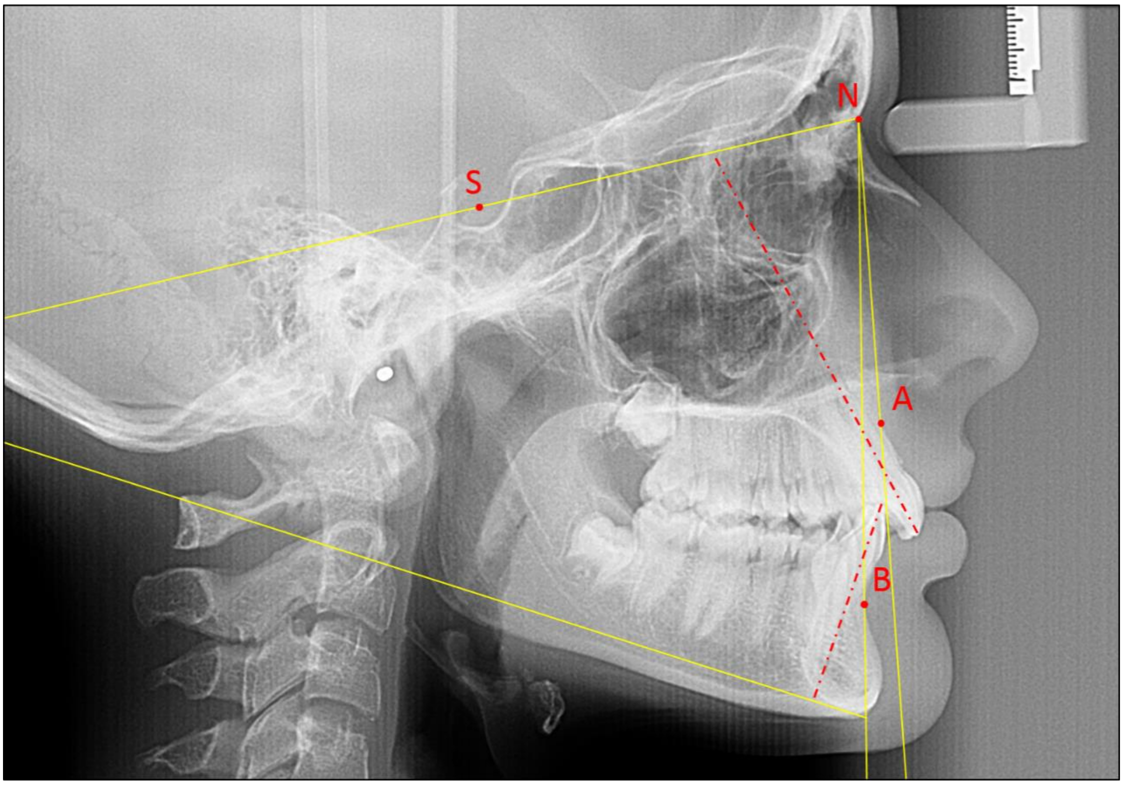

Figure 3.4: Later cephalometric radiograph with location of the cephalometric landmarks and reference lines.

Table 3.4: Subnasal soft tissue profile measurements of E- and H-lines

E-line	Norms (mm)
Upper lip	-4
Lower lip	-2
Holdaway-line	**Norms (mm)**
Nose	9
Subnasale	-5
Upper lip	0
Lower lip	0
Soft tissue B point	-5
Chin	0

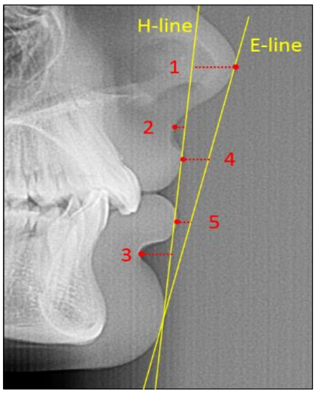

Figure 3.5: Cephalometric radiograph and soft tissue analysis

3.4.7. Panel measures

The members of the chosen panel were asked to answer specific questions on each participant. They were provided with all existing orthodontic records (dental casts, x-rays, and photographs) and relevant clinical information. Using the seven-point Likert scale, the panel members rated: the need of orthodontic treatment, the complexity of the treatment needed, and the amount of improvement after treatment (**Appendix II**). Moreover, they were required to assess if the orthodontic treatment outcome is acceptable or not, as well as estimating the required treatment time in months. It took each panel members around three sessions (30 – 60 minutes each) to complete the required tasks.

After the end of all rating sessions each expert was asked at which point of the seven-point Likert scale they would suggest that orthodontic treatment is truly indicated; this response is termed the indicated treatment need point (ITNP). Based on the average ITNP score, the participants were grouped either into the definitive need for treatment group (average need score ≥ ITNP), or the no need for treatment group (average need score ≤ ITNP).

3.4.8. Other measures

Other related variables were also collected to further describe the complexity of the malocclusion and the orthodontic treatment provided to the participants such as: pre-treatment age of the participants, treatment duration, total number of orthodontic appointments, and the total number of missed appointments.

3.5. Calibration

Prior to data collection phase, the principal investigator was calibrated against experienced orthodontists to ensure the reliability and accuracy in measuring occlusal indices, as well as reducing systematic bias and variation in assessing occlusal indices. The principal investigator measured and recorded both the PAR and the IOTN on 12 orthodontic study models displaying a various range of malocclusion severity, including both pre- and post-treatment dental casts. Separately, two experienced orthodontists measured the PAR (RF) and the IOTN (RH) as well. Two-way mixed model of intra-class correlation coefficient (ICC) were computed to test for the consistency in the scales measure. The values of the intra-class correlation coefficients were 0.94 for IOTN and 0.98 for the PAR.

3.6. Statistical analyses

After appropriate data entry, check-ups were conducted to detect entry mistakes. Subsequently, frequency distributions were generated for all variables to assess data distribution and the presence of outliers. Decisions on the need to regroup variables were taken when needed.

Data analyses were performed using Stata/SE ™ statistical software (version 11.1) and SPSS (version 19.0 Chicago, II). All statistical tests were two-tailed and P-values less than 0.05 were considered statistically significant.

Descriptive statistics were generated for all variables. The occlusal indices were compared to each other and to the gold standard (panel measures) before and after treatment. The correlations between the scales were computed using Spearman correlation and their test-retest reliability was computed using Kappa statistics and inter-class correlation coefficients.

The occlusal indices were compared to the panel measures for validation of treatment need; accordingly, a receiver operating curve (ROC) was derived for each scale, and the area under the curve (AUC) was used to find the best cutoff points of the different indices in identifying definitive need for orthodontic treatment. In the same way, the occlusal indices at the end of treatment were validated in identifying acceptable outcome. The assessment of treatment complexity by the indices was validated against both treatment time needed and panel opinion of treatment complexity; using the Spearman correlation. Similarly, Spearman correlations were used to validate the assessment of improvement after orthodontic treatment.

3.7. Ethical Considerations

The Institutional Review Board's (IRB) approval was obtained prior to the initiation of this study. Thus, all regulations and rules of confidentiality were followed according to the IRB protocols of the American University of Beirut (AUB). To ensure confidentiality and protect privacy, all the records of the enrolled participants were coded using serial numbers as identifiers. The coded records were provided to the study investigator who proceeded with data collection and computing. Upon completion of data collection, the records were returned and filed back.

Even though no direct benefits to the participants are expected in this study, the general benefit was expected in the research outcome. The major advantage is to identify the better occlusal index for the assessment of an orthodontic malocclusion and treatment outcome. This result would be of significant use in determining orthodontic needs and/or assessment of quality assurance through objective measures (see 1.4- Significance). A potentially improved index, with possible sub-combinations, can be used for research purposes, ranging from epidemiology of malocclusion to self-study evaluations in academic accreditation. Additional benefits are related directly to the potential use (regionally and globally) by insurance companies in assessment of need and coverage of orthodontics treatment, as well as setting public policies for oral health.

CHAPTER 4

RESULTS

The results were generated from the different occlusal indices, cephalometric analysis and panel's assessment. This project included 4 different constructs validation: need, complexity, improvement, and outcome. Each construct validation process is presented in a separate section. Frequency distributions and descriptive statistics were generated for all variables. The descriptive statistics were generated for occlusal indices before and after treatment (**Table 4.1**). Not only the conventional weighted occlusal indices were used in the validation assessment, but also their raw scores.

To validate the occlusal indices in gauging treatment need and outcome, the indices were compared to the panel measures (gold standard); accordingly, a receiver operating curve (ROC) was derived for each scale, and the area under the curve (AUC) was used to find the optimal cutoff points (**Table 4.5-13**). This was followed by a multivariable analysis to assess the weights of different components of each index in measuring need (**Table 4.28-31**).

The measurements of treatment complexity and improvement by the studied indices were assessed against the gold standard (average panel's assessment) using Spearman's correlation (**Table 4.13,14**). Spearman correlations were used because some of the panel assessments and occlusal indices were not normally distributed. This was followed by multivariable analysis, to assess the weight of pre- and post-treatment scores in measuring improvement (**Table 4.15**).

4.1. Descriptive statistics

The results are presented by the main categorization of the variables: participant's characteristics, occlusal indices, cephalometric measurements and panel assessments.

4.1.1. Participant's characteristics

The total number of participants included in this study was 101, with an average pre-treatment age of 17.2 ± 8 years. All the participants (39% males, 61% females) had completed orthodontic treatment assigned to the residents of the orthodontic Class of 2015: 43% by (resident MK), 30% by (SA), and 27% by (EZ). The progress notes of three participants could not be retrieved from the storage facility; therefore, the variables 'number of appointment' and 'missing appointments' were calculated for 98 participants rather than 101 subjects. The total number of appointments per participant ranged from 9 to 95 with an average of 35 ± 16, and the mean number of missing appointments were 2.5 ± 5 and ranged from 0 to 30 (**Appendix III Table 1**).

4.1.2. Occlusal indices

The average pre-treatment score of the DHC of IOTN was 3.5 ± 0.9, indicating need for orthodontic treatment, on the other hand the average pre-treatment score of the AC of IOTN indicated no need for orthodontic treatment (3.6 ± 1.9). Similar to the AC of IOTN, ICON pre-treatment mean score indicated no definitive need for orthodontic treatment (41.2 ± 18.8). All occlusal indices were significantly lower after orthodontic treatment (**Table 4.1a**). The average post-treatment scores of the PAR were 3.0 ± 4.9 and ICON were 12.3 ± 6.9, both indicating acceptable outcome of orthodontic treatment.

4.1.3. Cephalometric measurements

The participants had an average pre-treatment ANB angle of $4.7^{o} \pm 2.2$, with SNA and SNB angles of 80.8^{o} and 76.1^{o}, respectively. Post-treatment ANB angle decreased to $4.2^{o} \pm 2.2$, and it was more related to changes in the SNA angle (80.4) rather than the SNB (76.2^{o}).

On average, the participants had a normal divergence pattern with a palatal to mandibular plane angle (PP-MP) of $27.6^{o} \pm 5.5$, which remained almost unchanged after orthodontic treatment (27.5^{o}). Similarly, a normal pre-treatment maxillary incisor to SN angle ($101.7^{o} \pm 8.3$) remained relatively stable after treatment (101.5^{o}). The inclination of the mandibular incisors i-MP angle increased slightly from 94.0^{o} to 94.6^{o} after treatment.

Soft tissue measurements, specifically nose tip and subnasale to Holdaway line as well as upper and lower lips to the esthetic line, exhibited more changes during treatment than the more stable skeletal measurements (**Table 4.1b**).

4.1.4. Panel assessment

Orthodontic treatment need, complexity, and improvement were measured by the different committee members using 7-points licker scales (**Table 4.2**). The average panel assessments were 4.7 ± 1.1 for need, 3.8 ± 1.2 for complexity, and 4.6 for improvement. Outcome was assessed as a binary variable, with each committee member rating treatment outcome as acceptable or not. Non-acceptable outcome assessment ranged from 10 cases for one committee member (SB) to 2 for another (RS). The panel assessment of at least one committee members reporting unacceptable outcome was 11 cases.

4.2. Analysis of reliability

To assess the reliability of the different indices, all occlusal indices were re-measured on 10% of the sample. After at least two months, both the pre- and post-treatment indices were re-measured for those participants. All indices had high levels of reliability and ranged from 0.94 for the AC up to 0.99 for the DHC. The PAR, the ICON, and the DI all had a similar intra-class correlation coefficient of 0.98 (**Table 4.3**).

Similarly, the pre- and post-treatment skeletal, dental, and soft tissue cephalometric measurements were re-digitized to assess the intra-examiner reliability, which demonstrated excellent levels of reliability for all cephalometric measurements (**Table 4.3**).

4.3. Validation of occlusal indices as measures of need

4.3.1. Panel assessment of need

The mean score of need on the 7-point Likert scale was 4.7 ± 1.1. The 4 raters exhibited an excellent level of inter-rater reliability in assessing need; with a reliability coefficient of 0.85. The mean *indicated treatment point* (ITP) for the 4 raters was 4.25 (**Table 4.2**).

Participants with mean need scores $\geq$ 4.25 were assigned to the *definitive need for treatment* category, and represented 69% (n=70) of the sample. The remaining participants (n = 31; 31%) with scores below 4.25 were in the *no need for treatment* category.

4.3.2. Correlation analysis against panel assessment

The Spearman correlation coefficient computed to assess the relationship between the pre-treatment occlusal index scales and the panel assessment of orthodontic treatment need, indicated that all occlusal indices were significantly correlated with the panel assessment (**Table 4.4**).

Both the AC and the DHC of IOTN showed a highly statistically significant correlation with the need for orthodontic treatment ($r = 0.69$ and 0.72, respectively). When both components were added into a combined index score (cIOTN), a stronger association emerged ($r = 0.79$, p-value <0.0001) (**Table 4.4**).

The ICON scores were almost equally correlated with need as the PAR scores ($r = 0.67$ and 0.64, respectively). The correlation magnitude between the PAR and the need for treatment was similar for both weighted and unweighted (raw) PAR scores with correlation coefficients of 0.64 and 0.62, respectively. Similarly, no significant difference was found in the correlations magnitude of the weighted and unweighted ICON scores ($r = 0.67$ and 0.63, respectively).

The DI scores compared to other occlusal indices had the weakest correlation with the need of orthodontic treatment ($r = 0.35$, p-value $= 0.0004$). When the conventional DI was divided into different components: the cephalometric and the dental DI components. The dental DI was substantially better correlated with treatment need than the original DI ($r = 0.63$, p-value <0.0001), whereas the cephalometric DI was not associated with treatment need ($r = 0.09$, p-value $= 0.3476$).

The Facial Aesthetic Index (FAI) scores alone were not correlated with treatment need ($r = 0.08$, p-value $= 0.4080$). Even when the FAI scores were added to all occlusal indices (DHC, AC, PAR, ICON, DI, dDI) as a supplemental tool, it did not enhance their correlation with treatment need. In fact, all indices had either equal or weaker correlation with the panel assessment of need (**Table 4.4**).

4.3.3. Receiver operating characteristic (ROC) analysis

According to the panel assessments (the gold standard in this study), 70 participants (69%) were in need for orthodontic treatment. Using the original cutoff point for the DHC of IOTN (grade 4), the number of participants in need for orthodontic treatment were 85 patients. The ≥4 cutoff value gave the best psychometric properties with sensitivity of 71% and specificity of 90% (**Table 4.5**). The area under the ROC curve (AUC) was calculated at 0.85 (95% CI 0.76, 0.93) indicating a good ability of the DHC to discriminate between participants with definitive need for orthodontic treatment and with no need for treatment.

Similar to the DHC, the AC of IOTN demonstrated good discriminating abilities with AUC equal to 0.84 (95% CI 0.75, 0.92). The original cutoff value of ≥ 8 had very poor sensitivity of only 7% and specificity of 100% (**Table 4.6**). The best balance between sensitivity and specificity was found at the cutoff point of ≥ 3 (86%, 65%, respectively). Adding both the AC and the DHC scores together into a single overall score enhanced the psychometric properties of the IOTN, which demonstrated a better AUC value of 0.88 (95% CI 0.81, 0.96). The best cutoff score for the combined IOTN was ≥ 6, which had a sensitivity of 91%, and specificity of 71% (**Table 4.7**).

The ICON demonstrated similar psychometric properties to the AC and the DHC (AUC 0.83). The cutoff value of ≥ 33 had the best combination of sensitivity at

86% and specificity at 71% (**Table 4.8**). The raw scores of the ICON (without weightings) demonstrated a comparable AUC value to the weighted ICON (0.81, 0.83 respectively). The raw ICON (rICON) scores at the chosen cutoff value of ≥ 7 had a lower sensitivity (73%) and higher specificity (74%) than the weighted ICON scores (**Table 4.8**).

Similar to the ICON, both weighted and raw PAR scores had almost the same discriminating abilities, with AUC calculated at 0.82 and 0.81, respectively. The best cutoff score for the weighted PAR was ≥ 17, which had a sensitivity of 77% and a specificity of 61% (**Table 4.10**). When compared to the weighted scores, the raw scores demonstrated lower sensitivity (73%) and higher specificity (68%) at the best cutoff point of ≥ 12 (**Table 4.11**).

The ABO discrepancy index (DI) had poor discriminating abilities with an AUC calculated at 0.69 (95% CI 0.57, 0.80), along with 63% sensitivity and 61% specificity (**Table 4.12**). When only the dental component of the DI was used (dDI), the area under the curve significantly increased and became comparable to other indices (AUC 0.81; 95% CI 0.73, 0.90). Both the ≥ 8 and the ≥ 9 cutoff points gave good sensitivity of 76% and 70%, and specificity of 74% and 84% respectively (**Table 4.13**).

4.3.4. Multivariable analysis

Multiple logistic regressions were used to generate adjusted occlusal indices in estimating the need for orthodontic treatment. Unlike the AC and the DHC, indices with multiple components like the proposed cIOTN, the rPAR, the rICON, and the dDI were included in the multivariable analysis (**Table 4.14-18**). Even though the DHC had a higher regression coefficient than the AC (β = 0.57 and 0.27, respectively), the cIOTN had similar discriminating abilities whether both components were weighted or not (AUC = 0.88 and 0.89, respectively).

Understandably, the raw PAR and ICON component scores were used in the multivariable analysis. The ICON is composed of dental health components as well as a dental aesthetic component. Although some of the ICON dental health components were not significantly associated with need, removing them from the index did not enhance its psychometric properties (**Table 16, 17**). While keeping the dental health component the same (without weights), giving the aesthetic component a weight of 3 increased the ICON discriminating abilities from an AUC of 0.81 (rICON) to 0.85 (nICON) (**Table 4.20**).

The lower anterior segment displacement, the buccal segment transverse relationship, and the anterior over-bite components of the PAR index were not associated with the need for orthodontic treatment (**Table 4.15**). In fact, when a new total PAR score (nPAR) was calculated without these components, it demonstrated more accuracy in identifying participants with treatment need with an AUC of 0.88 (**Table 4.20**). The AUC value of the nPAR was higher than the rPAR and the conventional weighted PAR (0.88, 0.81, and 0.82, respectively).

The original DI, which encompasses cephalometric measurements was poor in discriminating participants with need for orthodontic treatment (AUC = 0.69). However, its dental component was significantly better in discriminating participants with definitive treatment need (AUC = 0.82). The dental component was further enhanced by weighting over-jet (x2) and removing lateral open bite (**Table 4.18**). This modified dental DI (nDId) had an AUC of 0.85 (95% CI 0.78, 0.92) (**Table 4.21**).

4.4. Validation of occlusal indices as measures of complexity

4.4.1. Panel assessment of complexity

The average panel assessment of complexity of the required orthodontic treatment for all the participants was 3.9 ± 0.9, and ranged from 1.8 to 6.0 (**Table 4.2**). The panel showed a good inter-rater reliability in this assessment (ICC = 0.80), the lowest average between the different raters measured at 3.5 (RS) and the highest at 4.7 (FR).

4.4.2. Correlation analysis against panel assessment

To explore the validity of the different occlusal indices in assessing orthodontic treatment complexity, Spearman's correlations were conducted between the different indices and the average panel assessment of complexity (**Tables 4.22**). The DHC of IOTN was the most significantly associated with the treatment complexity (r = 0.72, p-value <0.0001). However, the AC component of IOTN had a low positive correlation with the treatment complexity (r = 0.43, p-value <0.0001). Accordingly, the proposed combined cIOTN index had a stronger association than the AC, and weaker than the DHC (r = 0.62, p-value <0.0001).

Both the PAR index and the ICON were moderately associated with treatment complexity (r = 0.47 and 0.48, respectively); no significant difference was observed between the weighted and the raw scores (**Table 4.22**).

The DI had one of the lowest correlation coefficients (r = 0.43), but when only the dental component of the DI was used, the correlation increased (r = 0.61; **Table 4.22**).

The cephalometric component (cDI) had a negligible positive correlation with the complexity of the orthodontic treatment (r = 0.20, p-value <0.0417).

Except for the DHC and the dDI, the correlation between the indices and the treatment complexity increased when the supplemental FAI was added (**Table 4.22**). This increase was more significant in some indices (AC, cIOTN, rICON) than others (rPAR, PAR, ICON, DI).

4.4.3. Correlations against panel assessment of treatment duration

The estimated treatment duration was strongly associated with the complexity of the needed treatment. Therefore, the correlation between the occlusal indices and the estimated treatment time was comparable to the correlations reported for the complexity assessment (**Table 4.23**).

The DHC of IOTN was the most significantly associated with the estimated treatment time (r = 0.69, p-value <0.0001), whereas the DI was the least (r = 0.38, p-value = 0.0001).

The cephalometric component of the DI was not associated with the estimated treatment time (r = 0.11, p-value = 0.2939). Except for the DHC and the dDI, the correlation between the indices and the panel's average of estimated treatment time increased when the supplemental FAI score was added (**Table 4.23**). As previously mentioned for the complexity assessment, this increase was more significant in some indices (AC, cIOTN, rPAR, rICON) than others (PAR, ICON, DI).

4.5. Validation of occlusal indices as measures of improvement

4.5.1. Panel assessment of improvement

The average panel assessment score of the malocclusion improvement for all the participants was 4.8 ± 1.0, and ranged from 1.5 up to 7 (**Table 4.2**). The different panel members demonstrated an excellent level of inter-rater reliability in assessing improvement of the malocclusion after orthodontic treatment (ICC 0.90). The lowest improvement average score of a panel member was 3.7 (SB), whereas the highest average score was 5.5 (RS and FR).

4.5.2. Correlations analysis against panel assessment

The pre- and post-treatment changes in the occlusal indices scores were correlated with the assessment of improvement by the panel (**Tables 4.24**). The changes in both the AC and the DHC scores had a low positive correlation with the improvement of the malocclusion (r = 0.47, r = 0.48; p-value <0.0001, <0.0001 respectively). The proposed combined IOTN (cIOTN) demonstrated a stronger association than its aesthetic and dental components separately (r = 0.54, p-value <0.0001).

The PAR and the ICON were moderately associated with improvement (r = 0.52, r = 0.53; p-value <0.0001, <0.0001 respectively); both had no significant differences in the correlations reported for their weighted and raw scores (**Tables 4.24**).

Among all occlusal indices, the changes in the DI were the most significantly correlated with the improvement of malocclusion (r = 0.60, p-value <0.0001). When the DI dental component was assessed alone, the correlation coefficient decreased to r = 0.56 (**Table 4.24**). The cephalometric component of the DI alone had a weaker correlation with the improvement of the malocclusion after treatment ($r = 0.30$, p-value <0.0028).

The changes in the pre- and post-treatment FAI scores were added to the different occlusal indices to further assess their correlation with improvement. No major changes were depicted in the correlation coefficients of the indices with or without the FAI (**Table 4.24**). The correlation coefficients of the AC and the cIOTN were slightly enhanced with the addition of the FAI scores ($r = 0.52$ and 0.57, respectively). However, the FAI effect on the remaining indices was negligible (**Table 4.24**).

4.5.3. Multivariable analysis

Improvement assessment by occlusal indices is based on the changes between pre- and post-treatment scores; thus, it encompasses two variables. Accordingly, multiple linear regressions were made to assess the weight of each of these two variables in measuring improvement. Series of multiple regressions including pre- and post-treatment scores were made for the different indices: AC, DHC, cIOTN, rPAR, PAR, rICON, ICON, DI, dDI, and cDI (**Table 4.25 a**). The regression coefficient of the second variable (post-treatment score) was divided by the coefficient of the first (pre-treatment score), to calculate the weight that should be given for the post-treatment scores (**Table 4.25 b**). On average, the post treatment scores had 3 times the weight of the pre-treatment scores in predicting the degree of improvement of the malocclusion after orthodontic treatment.

Different weights were demonstrated by the various indices, and ranged from 1.1 for the post-treatment cephalometric DI scores to 5.5 for the AC scores. The regression coefficient of the post-treatment cephalometric DI scores (cDI) was similar to the pre-treatment scores (β = 0.09 and 0.10, respectively), indicating equal weights for both scores in assessing improvement with a calculated weight of 1.1, whereas the regression coefficient of the post-treatment AC scores was more than 5 times higher than the pre-treatment scores, in assessing malocclusion improvement (β = 0.17 and 0.94, respectively) (**Table 4.25, 26**).

4.5.4. Correlations analysis of the modified indices

After computing the weights for the occlusal indices, the newly weighted indices were assessed against the panel measurement of improvement (**Table 4.26 a, b**). For each index two different weights of the post-treatment scores were used; one that is specifically calculated for the index using linear regressions (exclusive weight), and the second was the average weight calculated for all post-treatment scores which was equal to 3 (average weight). The association between the improvement scores of the occlusal indices and the panel assessment of improvement were all statistically significant (**Table 4.26 a**).

When compared to the unweighted AC, the weighted AC demonstrated a stronger correlation with improvement (**Table 4.27 a**). The highest correlation coefficient was exhibited when a weight of 3 was given to the post-treatment scores (r = 0.60, p-value <0.0001). The DHC exclusive weight was calculated at 1.5 for its post-treatment scores. When this computed weight (x1.5) was used, the association with improvement after orthodontic treatment was equal to the unweighted DHC (r = 0.48 for both).

The weighted rPAR demonstrated a stronger association with improvement than the unweighted rPAR. Both the exclusive weight computed for rPAR (x3.5) and the average weight (x3) exhibited the same correlation with improvement ($r = 0.59$ for both). Similar to the rPAR, the weighted PAR had a higher correlation with improvement than the unweighted PAR, with a coefficient of 0.57 for the average weight of 3. The same trend but to a bigger extent was demonstrated by both rICON and ICON, where the weighted post-treatment scores were significantly higher than the unweighted scores (**Table 4.26 a**).

The computed weight for the DI was 1.5, hence, the post-treatment scores had a similar weight for predicting improvement as the pre-treatment scores. As a result, the correlation with improvement remained the same for both weighted and unweighted scores ($r = 0.60$ for both). To further explore possible weightings of the DI and their association with malocclusion improvement, the dental component only was weighted by 3 and then added to the unweighted cephalometric component (**Table 4.26 b**). This adapted DI exhibited a stronger association with malocclusion improvement than the unweighted scores ($r = 0.64$ and 0.60, respectively). Different weights were used to assess the proposed combined IOTN; both the average weight and the exclusive weight exhibited better association with improvement ($r = 0.61$, 0.60 and 0.54, respectively). This association was even further enhanced by using differential weights for its components; the average weight (x3) for the AC and the unweighted DHC ($r = 0.64$, p-value <0.0001).

4.6. Validation of occlusal indices as measures of outcome

4.6.1. Panel assessment of outcome

The 4 panel members assessed and rated the orthodontic treatment outcome as either acceptable or unacceptable. All four rated 90 participants as having an acceptable outcome, whereas 5 participants were rated as having unacceptable outcome by one panel member and 6 by two or more members. For the participant to be assigned to the unacceptable outcome category, at least one panel member had to rate the outcome as unacceptable. Therefore, according to these criteria, 11 participants were in the unacceptable outcome group and 90 in the acceptable group (**Table 4.2**).

4.6.2. Receiver operating characteristic (ROC) analysis

When evaluated against the panel opinion, the AC scores were significantly better than those of the DHC in discriminating unacceptable from acceptable outcomes (AUC = 0.81 and 0.71, respectively). Using the cutoff value ≥ 2, the AC demonstrated fair sensitivity of 73% and good specificity of 86% (**Table 4.27**). When both the AC and the DHC scores were added in one score (cIOTN) a more accurate index resulted with an AUC equal to 0.86 (95% CI 0.72, 0.98). The best cutoff score for this cIOTN was ≥ 4, which gave a sensitivity of 73% and specificity of 87% (**Table 4.29**).

The raw PAR scores and the conventional weighted PAR scores exhibited the same fair abilities in discriminating between unacceptable and acceptable outcomes, with identical AUC of 0.71 (95% CI 0.51, 0.92). Both demonstrated poor to fair psychometric properties with sensitivity of 55% and specificity of 86% for the rPAR at ≥ 3 cutoff point, and 64% sensitivity and 81% specificity for the PAR at ≥ 4 cutoff point (**Table 4.30, 31**).

Similar to the rPAR, the rICON scores had a fair accuracy in identifying unacceptable cases with AUC of 0.76 (95% CI 0.59, 0.93) (**Table 4.32**). The weighted ICON had better psychometric properties with higher AUC of 0.83 (95% CI 0.68, 0.98). Using the original cutoff point for the ICON (> 31), the number of participants with unacceptable outcome were 3 and gave 92% agreement with the panel decision. Nevertheless, the ≥14 cutoff point gave the best balance between sensitivity (73%) and specificity (81%) (**Table 4.33**). Changing the cutoff point from 31 to 14 decreased the percent agreement with panel assessment from 92% to 80%.

The AUC of the DI was calculated at 0.71 (95% CI 0.59, 0.93), and gave a poor sensitivity of 46% and good specificity of 84% (**Table 4.34**). The dental component of DI alone (dDI) had significantly better psychometric abilities than the conventional DI. The dDI had the highest AUC of 0.87 (95% CI 0.75, 0.99), and demonstrated an excellent sensitivity of 91% and fair specificity of 71% at ≥ 2 cutoff value (**Table 4.35**).

4.7. Cephalometric Analysis

Spearman correlations were computed to assess the relationship of the different pre-treatment cephalometric measurements with the average panel assessment of both orthodontic treatment need and complexity. None of the cephalometric measurements were associated with either the need for orthodontic treatment or the complexity of the treatment needed (**Table 36**).

The association between the post-treatment cephalometric measurements and the panel assessment of treatment outcome was not statistically significant, except for the distance between the Holdaway's line and Subnasale (**Table 37**). Nevertheless, almost all post-treatment cephalometric measurements were lower for participants with acceptable outcome when compared to others.

CHAPTER 5

DISCUSSION

5.1. Introduction

One of the major challenges in orthodontics is the availability of a suitable objective method for recording the different facets of an orthodontic problem. Numerous authors attempted to measure occlusion objectively, which lead to substantial numbers of occlusal indices. Despite the presence of many, there is no universally accepted single index or group of indices for measuring the different facets of a malocclusion. This limitation can be related to several reasons:

- The variations in terminology, concepts, and methodology that are present to date in these indices.
- Several of the commonly used indices have not been validated in assessing the different distinctive constructs defining the need, complexity and outcome of treatment.
- Most of the reported validation studies were done solely on dental models, and consequently did not reflect treatment standards on the basis of facial aesthetics or skeletal discrepancy.
- Despite the numerous recommendations on re-validating need indices in the country of application, no country in our region has validated such occlusal indices.

To the best of our knowledge, in addition to being the first attempt in validating four of the most commonly used indices (IOTN, ICON, PAR, DI) as measures of treatment need, complexity, improvement, and outcome, this study has the following characteristics:

- it differs from the majority of other published investigations in providing all the clinical information to the panel leading to a more needed accurate assessment.
- it is the first to assess the effect of facial esthetics on the proposed concepts by adding the recently published FAI to supplement the indices of interest
- it is the first in validating treatment need indices in this region.

The overarching goal was to find the better occlusal index for each concept category, and to assess the added value of including facial and cephalometric measurements.

5.2. Reliability of the occlusal indices

As with any diagnostic test, one of the most fundamental requirements of an occlusal index is reliability. A reliable occlusal index should produce consistent results or scores for a particular malocclusion. Therefore, reliability requires that repeated measures by the same or different examiners yield the same result. After no less than 8 weeks, occlusal indices in the present study were re-measured to assess their reliability. As anticipated and in agreement with the literature, excellent levels of intra-examiner reliability were achieved with no significant differences between the different occlusal indices (Beglin et al., 2001; Brook & Shaw, 1989; Cardoso et al., 2011; Younis et al., 1997).

The inter-examiner reliability of the studied indices was not in the scope of our investigation. Indeed, given the fact that the examiner was calibrated (with high correspondence) against orthodontists who had worked with the respective index, the inclusion of other examiners who would have warranted equal training in all the investigated indices and would have represented an objective for an independent study. In addition, in studies found in the literature regarding one or more indices, no major differences between intra- and inter-examiner reliability of the occlusal indices were reported (Buchanan, Shaw, Richmond, O'Brien, & Andrews, 1993; Firestone et al., 2002; Ovsenik & Primožič, 2007; Younis et al., 1997).

5.3. Occlusal indices as measures of treatment need

As anticipated all occlusal indices in this study were significantly associated with the need for orthodontic treatment. The DHC of IOTN demonstrated a stronger correlation with need than the other occlusal indices. In agreement with the literature, our assessment revealed that the recommended original cutoff point of the DHC (≥ grade 4) gave the best psychometric properties (Beglin et al., 2001; Younis et al., 1997). Unlike the DHC, the original cutoff point of the AC (≥ 8) and the later proposed point (≥ 5) both had poor sensitivity in this study, and the best balance between sensitivity and specificity was found at the cutoff point of ≥ 3.

A possible explanation could be that previously recommended cutoff points did not reflect the current emphasis on the importance of facial and dental aesthetics in relation to the malocclusion and the need for orthodontic treatment. Over the past decade, social pressures for perfect dental esthetics have reached the highest in human history, as reflected by the huge growing demands for dental veneers ('Hollywood smile'), and by the increasing numbers of older adults seeking orthodontic treatment, as well as the numbers of patients seeking treatment for minor occlusal problems.

Both the AC and the DHC demonstrated good accuracy as psychometric instruments, yet less than the previously reported accuracy in the literature (Beglin et al., 2001; Younis et al., 1997). The main factor for this difference is that the previous validation research included dental models only; thus, the panel assessment was solely based on dental casts rather than a comprehensive assessment of the participants. This way of validation dismisses valuable information that orthodontists routinely seek in order to diagnose and treat malocclusions. To overcome this shortcoming, we aimed to assess the validity of the occlusal indices against the most accurate panel assessment, which was achieved by providing all the required clinical information to the panel, and not only the dental dimension of the malocclusion that is captured by the dental casts. Accordingly, a comparative decrease in the accuracy of the occlusal indices resulted, compared to previous research.

The ICON index is heavily based on the AC of IOTN, thus drastically affected by the psychometric properties of the AC. Consequently, the best cutoff point for identifying cases with definitive need in this study (≥ 33) was significantly lower than the originally recommended (≥ 44). In fact, this difference between the ICON cutoff points (11 points) is probably attributed to the difference found in the AC cutoff points (2 points) multiplied by their weight in the ICON (x7).

The PAR index was developed to measure the success of the orthodontic treatment, and became one of the most commonly used indices in assessing the improvement from the original malocclusion and the end result of orthodontic treatment. Despite the original views that the PAR Index was neither designed nor validated as an index of treatment need, it was reported later to be an excellent predictor of treatment need when compared with expert orthodontic opinion.

Firestone et al. (2002) concluded that both the UK and the US PAR indices are as valid as the other common indices in measuring orthodontic treatment need, and recommended the same cutoff point (≥ 17) for both. Similarly, in the present study, the PAR has demonstrated a good accuracy in measuring the need of orthodontic treatment at the cutoff point ≥ 17 and was comparable to other indices that specifically developed to measure need. Nevertheless, the calculated area under the curve for the PAR index in this study (82%) was much less than that calculated by Firestone et al (96%). This discrepancy may be related to the aforementioned methodological differences: the panel examiners in the latter study based their assessment on dental casts only and did not account for any aesthetical or skeletal reasons of treatment need (Firestone et al., 2002).

The raw scores for both the ICON and the PAR demonstrated comparable accuracy and psychometric properties to the weighted scores. Thus, in the present study there was no added value for using the complicated weightings in assessing need. In fact, both indices being weighted or not demonstrated a similar psychometric and discriminating properties to the AC and the DHC. In this context, one would question the value of using these more complicated indices (ICON, PAR) to identify treatment need, while it can be captured with the same efficiency if not better using simpler indices such as the DHC of IOTN.

The DI was originally developed to measure complexity. However, because of its cephalometric component, we were interested in assessing its psychometric properties as a measure of treatment need. To the best of our knowledge, the DI was not validated as a measure of need against a gold standard of panel assessment. Among the different occlusal indices used in this study, the DI had unexpectedly the weakest correlation with the need for orthodontic treatment. This weak association was related to the cephalometric section, that was not associated with treatment need. On other hand, the dental component of the DI was associated with need and its psychometric properties were comparable to other indices. Therefore, the dental component of the DI alone without the cephalometric component can be considered a valid index for identifying definitive need with a cutoff value of ≥ 8. This finding may be attributed to the fact that skeletal correction in orthodontic treatment is limited, particularly past the patient's growth spurt, reflecting both the constitutional and the therapeutic limitations of orthodontic treatment (J. G. Ghafari, 2004). Accordingly, a patient who has an acceptable dental occlusion with steep mandibular plane or slightly increased ANB angle is not considered to have a *definitive need* for orthodontic treatment. This inference underscores the importance of dental health and aesthetics over skeletal measurements in assessing orthodontic treatment need. Even if a patient presented with severe skeletal abnormalities that necessitate combined orthodontic and orthognathic treatment, it is extremely rare for these skeletal abnormalities not to be accompanied by severe occlusal abnormalities that are captured by the dental assessments of the different occlusal indices.

The fact that psychosocial consequences of malocclusions due to unacceptable esthetics maybe as serious or even more serious than biological problems, warranted the use of soft tissue measurements in identifying need. Remarkably, none of the occlusal indices incorporated facial or soft tissue measurements in its overall score. A Facial Aesthetic Index (FAI) was recently introduced to compliment occlusal indices in generating more comprehensive measure of treatment need (Sundareswaran & Ramakrishnan, 2016). Nevertheless, it has not been validated as a supplemental need measure. In the present research, the FAI by itself was not associated with the need of orthodontic treatment. Even when the FAI scores were added to the different occlusal indices (DHC, AC, PAR, ICON, DI, dDI), it did not enhance their association with orthodontic treatment need. On the contrary, all indices had either equal or weaker correlations with need. This finding likely resulted from the same reasons mentioned above for the cephalometric measurements, but also because that the FAI index is greatly biased by the chin form.

The findings of the treatment need study suggest that the dental malocclusion actually is, at least on average, a proxy of both the underlying skeletal relationships and associated soft tissue profile. Moreover, orthodontists primarily focus their judgment of treatment need, by education and practice, on the initial assessment of dental relations. Skeletal relations and soft tissue appearance often represent boundary decisions not to violate or that would correspond to the "dental" decision. Departures from these premises would reflect an essential need to steer to orthognathic surgery or at best to compromised occlusion. In essence, these results validate the very concept of indexing malocclusions, as well as the primacy of the dental relations in this undertaking. This primacy is further illustrated by findings in the literature. The reliance on dental relationships by orthodontists to plan treatment has been documented by the prevalence of practitioners' opinion (diagnosis and treatment plan) based on only dental casts, even when shown some time later corresponding additional records such as a lateral cephalograph (Han, Vig, Weintraub, Vig, & Kowalski, 1991). Also buttressing the dominance of dental relations is the finding that the PAR index is highly correlated with orthodontists' subjective assessment of treatment need (McGorray et al., 1999).

Comparing the different indices together, the DHC, despite its simplicity, had slightly better psychometric properties than the other conventional indices. Interestingly however, when both scores of the AC and the DHC were added together into one simple score (cIOTN), the combined instrument yielded a better accuracy and psychometric properties than both of its components and all other indices included in this study. Remarkably, even after the multi-regression analysis of the different indices (cIOTN, rPAR, rICON, and dDI), that generated optimized weights for their components, the cIOTN remained on the top of the list and had similar discriminating abilities whether its components were weighted or not. Moreover, when the pre-treatment indices were correlated with each other, the cIOTN had higher correlations with most of the studied indices, making it a viable candidate to substitute them (**Appendix III Table 2**). Therefore, due to its simplicity and high psychometric abilities, the proposed combined cIOTN has all the necessary requirements to be the index of choice for measuring treatment need.

5.4. Occlusal indices as measures of treatment complexity

The two separate components of the IOTN were specifically developed to measure orthodontic treatment need. Neither the DHC nor the AC were ever used as measures of treatment complexity probably because of their simplicity in measuring occlusal traits. Surprisingly, among all the indices validated in this study, the DHC had the strongest correlation with treatment complexity. This finding indicates that the complexity of a malocclusion can in fact be gauged by its single most difficult occlusal trait. For instance, a malocclusion with multiple mild abnormalities will be labeled as relatively easy to treat, whereas a malocclusion with only one difficult abnormality will be considered more complex to treat. Compared to the DHC, the AC had a significantly weaker correlation with the complexity of the required treatment. As a result, the proposed combined cIOTN had a stronger association with complexity than the AC, but weaker than the DHC.

Regardless of the differences in their original purpose, both the PAR and the ICON were moderately associated with complexity, and there was no significant difference between weighted scores and raw scores. The ICON is a multifunctional index, designed to measure complexity, outcome, and need. Two of the previously published studies aimed at validating the ICON as a complexity measure (Savastano et al., 2003; Torkan et al., 2015). Our results are consistent with the Savastano et al. research that reported only a moderate agreement between ICON and the average opinion of the evaluating panel (n=15). In fact, our findings along with the findings from the previous two studies could not provide enough evidence to support the ICON as a measure of treatment complexity (Savastano et al., 2003; Torkan et al., 2015).

Only one study, to our knowledge, aimed at validating the PAR index as a measure of complexity (DeGuzman et al., 1995). In that study, the estimated duration of treatment was selected as a measure of difficulty, which the PAR index was validated against. Using only dental casts, the authors reported a moderate correlation (r = 0.57) between unweighted PAR scores and the ratings for treatment difficulty by a panel of orthodontists (n=11). Subsequently, this correlation was increased to r = 0.65 by applying their computed set of weightings, thus defining the US PAR (DeGuzman et al., 1995). The unweighted PAR scores in our study demonstrated a comparable correlation with the estimated *treatment duration* to the previous study (r = 0.52), and it was significantly increased by applying their recommended weighting system (US PAR) in our sample (r = 0.63), which also was significantly higher than the conventional (UK) PAR (r = 0.54). Nevertheless, the panel assessment of complexity rather than estimated duration was used in our study as the indicator of treatment difficulty. Compared to duration, *treatment complexity* had significantly lower correlations with the rPAR, the PAR and the US PAR (r = 0.44, 0.47, and 0.53, respectively). These findings highlight the differences between the estimated treatment duration and the complexity assessment, and furthermore questions the validity of the PAR index in assessing the complexity of the needed orthodontic treatment.

Despite being developed specifically to measure complexity, the DI had one of the lowest correlation coefficients with complexity assessment. In fact, the cephalometric component of the DI had a negligible correlation with the complexity of orthodontic treatment. Nevertheless, when only the dental component of the DI was computed, the correlation with complexity significantly increased. Only in a recently published study, the DI was associated with the perceived assessment of complexity, by adding a relatively small assessment section to a larger web survey that included 343 dental professionals: orthodontists, general dentists, orthodontic residents, and dental students (Heath, English, Johnson, Swearingen, & Akyalcin, 2017). Compared to our results, the authors reported a substantially higher correlation between the DI and the dental professionals. Besides their small sample size of 29 cases and employing a web survey, the cases were chosen according to their complexity as assessed by the DI: 10 mild (DI, 0-10), 9 moderate (DI, 11-20), and 10 complex (DI, .20). This methodology might explain the significant difference in the correlation magnitude between the two studies. The present study is the first, to our knowledge, to aim at validating the DI as a measure of complexity against orthodontic experts and compare the DI to the assessment of other occlusal indices. Based on our results, we suggest using the dental component of the DI as a complexity measure, without including the cephalometric component. Even when crude cephalometric measurements were used, no significant association was found with treatment complexity or need. Nevertheless, further studies targeting specifically the cephalometric measurements and their relation to other panel assessments are recommended.

Except for the DHC and the dDI, the correlation between the indices and the complexity of treatment increased by adding the supplemental FAI. Although this increase was more significant in some indices (AC, cIOTN, rICON) than others (rPAR, PAR, ICON), it was not significant enough to consider the present FAI as a discerning index. Nevertheless, further exploration is warranted because when FAI was used in combination with occlusal indices, the validity of complexity measures increased. Remarkably, the conventional DHC (without the FAI) compared to all other indices, with or without the FAI, had the strongest association with treatment complexity.

5.5. Occlusal indices as measures of treatment outcome

5.5.1 Improvement measure

The outcome of orthodontic treatment can be measured either directly without taking the initial malocclusion into consideration, or as the amount of malocclusion improvement between initial and post-treatment scores.

To assess the validity of the occlusal indices as measures of improvement, the changes between the pre- and post-treatment scores were associated with the average panel assessment of malocclusion improvement. Interestingly and as in measuring need, the proposed combined index cIOTN demonstrated a greater association with the improvement of malocclusion than its components, the AC and the DHC. It was even slightly higher than either the PAR and the ICON, and equal to the dental component of the DI. Also of great interest is the finding that, among the different occlusal indices, the changes in the DI were the most significantly correlated with improvement, with a correlation coefficient ($r = 0.60$) higher than either its dental or cephalometric component. Adding the FAI scores to the different occlusal indices did not significantly affect their validity in measuring improvement; therefore, the conventional DI remained the better measure of malocclusion improvement.

When computing multiple linear regressions for the weights of post-treatment improvement index scores, these scores were weighted, on average, 3 times higher than the pre-treatment scores. This finding underscores the importance of the final occlusion standards, even when taking the initial malocclusion into consideration. When the modified (newly weighted) indices were assessed against the panel judgment, almost all of them had a stronger correlation with improvement. This increase was more evident in the AC index, as well as the ICON as it is heavily weighted toward its AC component. The DHC on the other hand, did not require any weight for its post-treatment scores. Interestingly, when the weighted AC was combined with the unweighted DHC to create a weighted cIOTN, the latter had a significantly higher correlation with improvement than either.

The DI had the highest correlation with the improvement assessment before modifying the indices. However, after applying the weight of 3 for the post-treatment scores of the cIOTN (3xAC), the ICON (3x), and the DI (3xdDI), all had the same association with improvement, and can in fact be used as improvement measures (**Table 4.26**). We preferred using the uniformed weight of 3 (the average weight) for the following reasons: the differences between this weight and other more exclusive weights were negligible; it was simpler to use than exclusive weights with decimals; and it could be uniformly applied to the different indices.

5.5.2 Outcome success measure

Despite the fact that the AC were significantly better than the DHC in discriminating unacceptable from acceptable outcomes, the proposed cIOTN yielded a more accurate index than its components (AUC 0.85; **Table 5.1**). The cIOTN had a significantly higher discriminative power than the PAR and the DI, and it was comparable to the ICON and the dDI. The original cutoff point for the ICON (> 31) gave a very low sensitivity of 45%, whereas the best balance between sensitivity (73%) and specificity (81%) was for the cutoff point ≥14. Changing the cutoff point from 31 to 14 decreased the percent agreement with panel assessment from 92% to 80%, probably because of the low prevalence of participants with unacceptable treatment outcome in this study (11%). In fact, given the low prevalence and high cost of false positive diagnosis of unacceptable outcome, the cut-point could be chosen at higher values to maximize specificity. Nevertheless, the sensitivity, specificity and AUC are independent on the prevalence of unacceptable outcome. The dDI had the highest AUC value at 0.87, and demonstrated an excellent sensitivity (91%) and fair specificity (71%)

at a cutoff value ≥ 2. To maximize specificity (90%) at the expense of sensitivity (64%), the ≥ 3 cut-point could be used.

After comparing the results of the different indices as measures of outcome success, the cIOTN, the ICON and the dDI can be considered as valid measures of unacceptable outcome. Since the ICON is heavily weighted towards its dental aesthetic component (AC), it was very comparable to the cIOTN. In fact, when the raw ICON scores were used, the discriminating abilities were significantly decreased. Therefore, the cIOTN and the dDI are preferable measures of orthodontic treatment outcome.

5.6. Limitations and research considerations

The interpretation of findings from this validation study should be considered in light of their limitations. Using the opinion of an orthodontic panel as a gold standard is an accepted method to establish and validate occlusal indices (Brook & Shaw, 1989; Richmond et al., 1992). However, it was commonly stated that the practitioner's country of origin as well as place of practice has a significant effect on prescribing orthodontic treatment (Richmond & Daniels, 1998a), which underscores variability, particularly in determining treatment need over different geopolitical areas.

In the present study, all the panel members had completed their orthodontic training in the United States at ADA–accredited programs, and have been practicing orthodontics in Lebanon for more than 10 years. These criteria were placed to establish the best diagnostic gold standard that is attainable for the participants. The strict criteria led to the smaller number of committee members. While this number could affect the result, it is in line with at least the lower number of included reviewers in the literature (n= 3 to 15) (Beglin et al., 2001; Cardoso et al., 2011; Louwerse et al., 2006; Torkan et al., 2015). In fact, the excellent inter-rater reliability between the different raters, along with the high level of agreement between the panel's assessment and the occlusal indices, strongly indicate the efficacy of such indices in distinguishing patients with definitive need for treatment from others. To further scrutinize the adequacy of judgment by the panelists, we visually examined the individual scoring of each committee member. We determined that the ratings of one of the members (FR) were consistently slightly higher than the others. To test the potential effect of this member's rating, a sensitivity test was conducted by excluding him from the statistical analyses. The results demonstrated no changes in the different diagnostic measures used in this study (**Appendix III Table 3**). Accordingly, given the results of the sensitivity analysis and the high inter-class correlation coefficients, the ratings of this panel member were kept in all assessments.

Authors have indicated that validation studies of diagnostic instruments should include patients that mirror the prevalence of the disease in the general population (Metz, 1978; Swets, 1988). However, occlusal indices are not only used as an epidemiologic instrument; an index should also function as a clinical tool. This premise is crucially important for complexity and outcome indices. Accordingly, the participants were randomly recruited in the present study to reflect the prevalence of malocclusion in an orthodontic clinic population. Because of this random sample selection, the number of participants with unacceptable outcome was small, reflecting the norm for orthodontic residency programs. The small number of participants did not allow for a rigorous analysis of the indices as outcome measures.

We aimed at assessing indices as measures of different concepts (need, complexity, improvement, and outcome). The panel members were asked to assess the final outcome of orthodontic treatment without taking into consideration the initial malocclusion. Since the panel assessed the need, complexity, and amount of improvement before assessing the outcome success, it is possible that their final outcome judgment was biased and not totally independent from the improvement assessment. In retrospect, this presumed bias could have been avoided by providing the post-treatment records first to allow the panel to rate the final outcome before providing them with the remaining records for the other assessments.

To determine potential reasons behind the selection of the panel of “need” (71%) and “no need” for treatment, we compared the demographic and cephalometric characteristics of participants in these groups. No differences in age or gender were observed. As expected, the majority of cephalometric measurements were closer to the norm among the participants with no definitive need for treatment, albeit not statistically significant (**Appendix III Table 4**). The only statistically significant differences were noted for 2 facial esthetic measurements, the nose tip projection relative to the Holdaway line (lower than the norm in the treatment “need” group), and the upper lip to the Esthetic line (closer to the line, thus fuller upper lip in the “need” group). These findings represent expected tendencies in patients who need orthodontic treatment.

The cephalometric measurements at T1 and T2 revealed average numbers of ANB (4.7° and 4.2°, respectively) that were closer to a Class II relationship, and average numbers for SNB (76.1° and 76.2°, respectively) indicating mandibular retrognathism (**Table 4.1b**). These findings prompted the question of prevalence of Class II malocclusions in the sample. We thus separated the participants with ANB > 4.5°, a definite Class II “zone” from the total sample. With a total subsample of 55 subjects, we ran all panel assessment analyses. The findings were similar (**Appendix III Table 5**) to those observed for the entire sample (**Table 4.2**).

5.7. Strengths

To best of our knowledge, this is the first study to validate this group of commonly used indices against the same ‘gold standard’ assessment of experts. Also, the facial esthetic dimension of malocclusions was evaluated for its possible relation to treatment need, complexity, and outcome. Therefore, an attempt was made to validate the recently published FAI, as well as assessing the different skeletal and soft tissue cephalometric measurements.

More specifically, this study represented the first attempt at assessing the validity of the DI, the PAR, the ICON, and the IOTN not only in measuring their original purpose category, but also in measuring the other categories that they were not developed for (need, complexity, improvement, and outcome). This simultaneous appraisal disclosed important conclusions regarding the IOTN, the DI, and the apparent lack of added value of cephalometric and facial esthetics components (**Table 5.1**).

To accurately validate diagnostic instruments, it need to be compared to a gold standard, that should reflect the true status of the construct. Therefore, it was crucial that the gold standard against which the scores are validated had the best diagnostic levels achievable. Accordingly, the present study differed from the majority of other published investigations, not only in the inclusion of strict selection criteria for the panel members, but also in providing them with all clinical information (photographs, models, and x-rays) needed to reach the most accurate clinical diagnosis.

Finally, to reflect a truthful clinical situation, the sample was randomly selected from a list containing the names of all the patients who finished treatment by the same class of residents who entered and graduated from the program at the same time, thus exposed to similar educational experiences.

Table 5.1: Comparison of the AUC (need, outcome) and correlation coefficients (complexity, improvement) of the original indices (AD, DHC, PAR, DI, ICON), and the new indices suggested in this study (cIOTN, dDI):

INDICES	Need	Outcome	Complexity	Improvement
	AUC		Correlation coefficient	
AC	0.84	0.81	0.43	-0.60
DHC	0.85	0.71	0.72	-0.48
cIOTN	0.88	0.85	0.62	-0.64
ICON	0.83	0.83	0.48	-0.64
PAR	0.82	0.72	0.47	-0.57
DI	0.69	0.71	0.43	-0.64
dDI	0.82	0.87	0.61	-0.56

Bright yellow indicates strong relationships within the method of evaluation (AUC or correlation). Light yellow represents lower strength. cIOTN and dDI joined AC and ICON as usable for assessment of need and outcome, although cIOTN was the strongest in scoring need, and dDI in measuring outcome. The original IOTN DHC component was highest in evaluating complexity, seconded equally by cIOTN and dDI, surprisingly edging the original DI, which was specifically developed for such measure.
Note 1: colors of indices correspond to their original purpose: blue for AC and DHC (need); orange for PAR (outcome); green for complexity (DI); all colors for ICON (multifunction).
Note 2: In ROC analysis, an AUC between 0.80-0.90 is considered good, and >0.9 excellent (Fawcett, 2006)

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

In a unique study comparing common occlusal indices across treatment need, complexity, and outcome, our findings provide extensive evidence on their reliability and validity. The indices were tested in the different construct categories, and the results yielded remarkable contributions to the development and evaluation of the optimal characteristics of a comprehensive occlusal index that is based on established indices, simple, reliable and valid to measure all different constructs.

6.1.1. Conclusions of the treatment need study

A substantial agreement was found between the gold standard and the AC, the DHC, the ICON and the PAR. Thus, these indices are valid to assess the need of orthodontic treatment. However, the DHC of IOTN demonstrated a stronger correlation with need than the other conventional indices (AC, ICON, and PAR).

For better diagnostic accuracy, the cutoff point of the AC required adjustment from the recommended ≥ 5 to ≥ 3. The shift can be attributed to the increased social pressure for perfect dental esthetics. Adding both the AC and the DHC together into a newly proposed index score cIOTN gave a better accuracy and psychometric properties than either of its components and all other indices included in this study.

The raw scores for both the ICON and the PAR demonstrated comparable psychometric properties to the conventional weighted scores, and therefore are valid measures of treatment need.

The DI had the weakest correlation with the need for orthodontic treatment, and it could not be validated as a treatment need measure. Nevertheless, the dental component of the DI alone *without* the cephalometric component was significantly associated with need and had a comparable validity to other need indices. Neither the cephalometric component of the DI nor any of the different cephalometric measurements were associated with need.

The FAI did not enhance the validity of the occlusal indices in our sample. Therefore, it could not be validated as a supplemental measure of treatment need.

6.1.2. Conclusions of the treatment complexity study

The DHC had the strongest correlation with treatment complexity; it was significantly better than the AC, the ICON, the PAR, the US PAR, and even the DI, which was developed specifically to measure treatment complexity. Remarkably, the simple conventional DHC (without the FAI) compared to all other indices, with or without the FAI, had the strongest association with treatment complexity.

The (US) PAR had stronger associations with complexity assessment and estimated duration than the original (UK) PAR. Nevertheless, both as well as the ICON were moderately associated with complexity; in addition, there were no significant difference between weighted scores and raw scores. Consequently, our findings could not provide the evidence to support them as a measure of treatment complexity.

With one of the lowest correlation coefficients with the complexity assessment, the DI could not be validated as a complexity measure. The cephalometric component of the DI had a negligible correlation with the complexity of the needed orthodontic treatment. However, the dental component of the DI had a significantly stronger correlation with the complexity assessment. Hence, using the dental component of the DI alone without the cephalometric component is a valid measure of complexity.

Even though it was developed to supplement occlusal indices in measuring need, adding the FAI to the occlusal indices increased their correlation with treatment complexity. This increase was more significant in some indices than others and warrants further exploration in future studies.

Despite being closely related, one of the major findings in this study was the differences found between the estimated treatment duration and the complexity of the needed treatment when correlated with the indices and between each other. Accordingly, we do not recommend the use of the estimated treatment duration as the gold standard for assessing the complexity of the needed orthodontic treatment.

6.1.3. *Conclusions of the improvement study*

Measuring the amount of malocclusion improvement after orthodontic treatment requires the assessment of both pre- and post-treatment index scores. Based on multivariable analyses, the post-treatment scores were weighted, on average, 3 times higher than the pre-treatment scores to more accurately gauge the amount of improvement assessed by the panel. Among the different indices, only the DHC did not require an increased amount of weight for its post-treatment scores.

Without weighting the post-treatment scores, the DI had the most significant association with improvement assessment, followed by the cIOTN. However, after applying the weights, the DI, the cIOTN, and the ICON all had a significantly stronger association with improvement and were considered valid improvement measures. Accordingly, the weighted DI, cIOTN, and ICON are recommended for determining malocclusion improvement after orthodontic treatment

6.1.4. Conclusions of the outcome success study

In directly assessing the end result of orthodontic treatment, the cIOTN as well as the dental component of the DI had a "good" ability to correctly classify those with and without acceptable treatment outcome. Similarly, the ICON had good discrimination abilities; however, the cutoff point had to be substantially decreased in order to get the best balance between sensitivity and specificity.

Because the DHC, the PAR, and the DI had fair to poor diagnostic accuracy in assessing acceptable and unacceptable outcomes, they are not recommended for determining the end result of orthodontic treatment, and are considered less valid than the dDI, cIOTN, and ICON. Given the possibility of bias in the panel assessment and the small number of participants with unacceptable outcome, it would be appropriate to reassess outcome success with only the final records of a wider range of treated subjects presented to the panel.

6.1.5. Main contributions of the study

1. One of the novel findings of this investigation is that the proposed combined IOTN (cIOTN) can be an excellent multifunctional index, as it demonstrated surprising

abilities in measuring all different constructs, with only minor adjustments; like using the DHC alone in assessing complexity, and weighting the post-treatment scores (x3) in assessing improvement.

2. The dental DI (dDI) reflected need, complexity and outcome success better than the original DI, which includes cephalometric assessment.
3. Non-dental measures, specifically lateral cephalometric and facial soft-tissue did not represent an added value to the accuracy of the respective indices in assessing need and outcome.

6.2. Recommendations for future studies

Our in-depth analysis of the common occlusal indices clearly illustrates the deficiencies and limitations of some of these indices. We therefore emphasize on the importance of utilizing the right index that is more valid to measure the outcome of interest. While the cIOTN best fit this description, further studies aiming on examining its validity against other panels assessment are essential to validate our novel finding.

Since orthodontic constructs such as treatment need or complexity are not directly observable and cannot be defined with a series of well recognized symptoms, the importance of expert or professional panels in assessing these constructs cannot be overestimated. Despite the common use of such expert panels as a reference standard in validating occlusal indices, the panel composition differed substantially across studies with large variations in the certainty of assessments, type of expertise represented, and years of experience. Therefore, it is necessary to standardize the construction of these panels to better serve their purpose. Important factors to be considered in future validation studies are selection criteria for panel members, type and format of information presented to the panel, and the inter-rater reliability between panel members.

Based on the results of this study, the challenges faced, the experience accumulated, and the lack of formal guidelines, we propose the following practical guidelines for future orthodontic validation studies:

- Examiners calibration and training in measuring occlusal traits or indices.
- Selection of the cases that represent the population of interest, where the index will be used.
- Selection of panel members with accredited orthodontic education and enough experience to reflect the reference standard of experts' opinion.
- Providing the panel with all needed clinical information to enable the most accurate diagnosis.
- Evaluation of the inter-rater reliability between the different members of the panel.
- Evaluation of the different measurements of diagnostic accuracy and occlusal index reliability.

TABLES

Table 4.1a: Descriptive statistics of pre- and post-treatment indices' score:

Occlusal Indices	Pre-treatment Mean (SD)	Post-treatment Mean (SD)	Percent Change	P-value
IOTN-AC	3.6 (1.9)	1.3 (0.6)	64%	< 0.0001
IOTN-DHC	3.5 (0.9)	1.5 (0.7)	57%	< 0.0001
PAR	21.7 (10.2)	3.0 (4.9)	86%	< 0.0001
ICON	41.2 (18.8)	12.3 (6.9)	70%	< 0.0001
FAI	0.5 (1.0)	0.3 (0.7)	40%	0.0005
DI	16.7 (9.9)	9.3 (8.6)	44%	< 0.0001

Table 4.1b: Descriptive statistics of pre- and post-treatment cephalometric measurements:

Cephalometric Measurements	Pre- treatment		Post- treatment		P-value
	Mean (SD)	(Min, Max)	Mean (SD)	(Min, Max)	
SNA angle	80.8 (4.1)	(70.6, 90.8)	80.4 (4.3)	(71.4, 93.4)	0.0011*
SNB angle	76.1 (4.0)	(68.9, 84.2)	76.2 (4.3)	(68.5, 87.8)	0.2020
ANB angle	4.7 (2.2)	(-0.7, 9.9)	4.2 (2.2)	(-1.8, 9.3)	< 0.0001*
PP-MP angle	27.6 (5.5)	(15.5, 42.7)	27.5 (5.8)	(14.1, 42.3)	0.5890
MP-SN angle	38.0 (5.9)	(26.4, 52.7)	38.2 (6.0)	(26.1, 53.6)	0.2037
I-SN angle	101.7 (8.3)	(80.2, 122.1)	101.5 (7.6)	(87.2, 120.9)	0.8629
I-NA angle	20.9 (7.5)	(0.3, 37.7)	21.1 (7.1)	(6.9, 41.6)	0.8007
i-MP angle	94.0 (6.5)	(75.6, 107.5)	94.6 (7.4)	(77.1, 113.1)	0.3806
i-NB angle	28.1 (5.8)	(11.5, 42.3)	29.1 (6.0)	(12.3, 40.9)	0.1464
Nose Tip / HL	4.3 (4.7)	(-5.6, 16.1)	6.0 (4.2)	(-2.3, 15.1)	< 0.0001*
Subnasale / HL	5.4 (2.6)	(-0.8, 13.1)	4.8 (2.7)	(-0.2, 11.4)	0.0046*
Lower Lip / HL	1.3 (1.5)	(-2.3, 5.6)	1.1 (1.4)	(-1.9, 5.1)	0.1665
Mental sulcus / HL	4.1 (1.7)	(0.2, 8.8)	4.1 (1.9)	(0.1, 9.4)	0.8065
Upper lip / EL	-2.4 (2.6)	(-9.2, 3.0)	-3.3 (2.3)	(-8.3, 1.3)	< 0.0001*
Lower Lip /EL	-0.2 (2.6)	(-7.6, 5.6)	-1.0 (2.4)	(-7.3, 5.9)	< 0.0001*

I Maxillary incisors, i Mandibular incisors, HL Holdaway line, EL Esthetic line
*P-value < 0.05

Table 4.2: Summary statistics of panel assessment scores:

Variable	PA	SB	FR	RS	gold standard
Need					
Mean (SD)	4.4 (1.6)	4.2 (0.9)	5.6 (1.4)	4.6 (1.5)	4.7 (1.1)
ITP					
Value indicated	3	3	5	6	4.25*
Complexity					
Mean (SD)	3.8 (1.2)	3.7 (1.0)	4.7 (1.2)	3.5 (1.2)	3.9 (0.9)
Improvement					
Mean (SD)	4.6 (1.3)	3.6 (1.0)	5.5 (1.3)	5.5 (1.2)	4.8 (1.0)
Outcome					
Not acceptable	4 (4.0%)	10 (10.0%)	4 (4.0%)	2 (2.0%)	11 (11.0%)
Acceptable	97 (96.0%)	91 (90.0%)	96 (96.0%)	99 (98.0%)	90 (89.0%)

* This average was calculated from the different ITP values reported by each panel member

Table 4.3: Intra-class correlation coefficients for intra-examiner reliability

Occlusal Indices		Skeletal measurements		Dental measurements		Soft tissue measurements	
AC	0.941	SNA	0.996	UI-SN	0.985	H-Nose Tip	0.998
DHC	0.985	SNB	0.995	UI-NA	0.968	H-Sub	0.987
PAR	0.978	ANB	0.961	LI-MP	0.961	H-LL	0.983
ICON	0.978	MP-PP	0.974	LI-NB	0.951	H-Sulcus	0.992
DI	0.982	MP-SN	0.976			E-UL	0.997
						E-LL	0.993

Table 4.4: Spearman correlation analysis between need for orthodontic treatment and different occlusal indices

NEED	Indices without FI	P-value	Indices with FI	P-value
AC	0.6869	< 0.0001	0.6578	< 0.0001
DHC	0.7170	< 0.0001	0.5310	< 0.0001
cIOTN*	0.7890	< 0.0001	0.7521	< 0.0001
rPAR**	0.6242	< 0.0001	0.6172	< 0.0001
PAR	0.6389	< 0.0001	0.6296	< 0.0001
rICON**	0.6331	< 0.0001	0.6272	< 0.0001
ICON	0.6680	< 0.0001	0.6755	< 0.0001
DI	0.3476	0.0004	0.3510	0.0003
cDI	0.0856	0.3948	0.1143	0.2550
dDI	0.6339	< 0.0001	0.5898	< 0.0001

* combined IOTN indicates adding the AC and the DHC together into one score
**raw indicates adding the raw scores into a total score without using any weights

Table 4.5: The ROC curve analysis of the DHC as measure orthodontic treatment need

DHC Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 1	100.00%	0.00%	69.31%	1	
≥ 2	98.57%	0.00%	68.32%	0.9857	
≥ 3	95.71%	45.16%	80.20%	1.7454	0.0949
≥ 4	71.43%	90.32%	77.23%	7.381	0.3163
≥ 5	15.71%	96.77%	40.59%	4.8714	0.871
> 5	0.00%	100.00%	30.69%		1

AUC = 0.85 (95% CI = 0.75, 0,93)

Table 4.6: The ROC curve analysis of the AC as measure orthodontic treatment need:

AC Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 1	100.00%	0.00%	69.31%	1	
≥ 2	100.00%	35.48%	80.20%	1.55	0
≥ 3	85.71%	64.52%	79.21%	2.4156	0.2214
≥ 4	54.29%	90.32%	65.35%	5.6095	0.5061
≥ 5	30.00%	96.77%	50.50%	9.3	0.7233
≥ 6	22.86%	96.77%	45.54%	7.0857	0.7971
≥ 7	11.43%	96.77%	37.62%	3.5429	0.9152
≥ 8	7.14%	100.00%	35.64%		0.9286
≥ 9	2.86%	100.00%	32.67%		0.9714
≥ 10	1.43%	100.00%	31.68%		0.9857
> 10	0.00%	100.00%	30.69%		1

AUC = 0.84 (95% CI = 0.75, 0.92)

Table 4.7: The ROC curve analysis of the combined IOTN (cIOTN) as measure orthodontic treatment need:

cIOTN Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 3	100.00%	0.00%	69.31%	1	
≥ 4	100.00%	22.58%	76.24%	1.2917	0
≥ 5	98.57%	45.16%	82.18%	1.7975	0.0316
≥ 6	91.43%	70.97%	85.15%	3.1492	0.1208
≥ 7	71.43%	83.87%	75.25%	4.4286	0.3407
≥ 8	48.57%	96.77%	63.37%	15.0572	0.5314
≥ 9	32.86%	96.77%	52.48%	10.1857	0.6938
≥ 10	25.71%	96.77%	47.52%	7.9714	0.7676
≥ 11	15.71%	100.00%	41.58%		0.8429
≥12	7.14%	100.00%	35.64%		0.9286
≥ 13	4.29%	100.00%	33.66%		0.9571
≥ 14	1.43%	100.00%	31.68%		0.9857
> 14	0.00%	100.00%	30.69%		1

AUC = 0.88 (95% CI = 0.81, 0.96)

Table 4.8: The ROC curve analysis of the ICON as a measure of orthodontic treatment need:

ICON Cutoff point	Sensitivity	Specificity	Classified	LR+	LR-
≥ 7	100.00%	0.00%	69.31%	1	
≥ 11	100.00%	6.45%	71.29%	1.069	0
≥ 12	100.00%	9.68%	72.28%	1.1071	0
≥ 14	100.00%	12.90%	73.27%	1.1481	0
≥ 16	100.00%	16.13%	74.26%	1.1923	0
≥ 17	100.00%	19.35%	75.25%	1.24	0
≥ 18	100.00%	25.81%	77.23%	1.3478	0
≥ 20	97.14%	32.26%	77.23%	1.434	0.0886
≥ 23	97.14%	35.48%	78.22%	1.5057	0.0805
≥ 24	97.14%	41.94%	80.20%	1.673	0.0681
≥ 25	97.14%	45.16%	81.19%	1.7714	0.0633
≥ 26	94.29%	45.16%	79.21%	1.7193	0.1265
≥ 27	94.29%	51.61%	81.19%	1.9486	0.1107
≥ 28	91.43%	51.61%	79.21%	1.8895	0.1661
≥ 29	91.43%	58.06%	81.19%	2.1802	0.1476
≥ 30	87.14%	58.06%	78.22%	2.078	0.2214
≥ 31	87.14%	61.29%	79.21%	2.2512	0.2098
≥ 32	85.71%	64.52%	79.21%	2.4156	0.2214
≥ 33	85.71%	70.97%	81.19%	2.9524	0.2013
≥ 34	78.57%	70.97%	76.24%	2.7063	0.3019
≥ 35	75.71%	77.42%	76.24%	3.3531	0.3137
≥ 37	72.86%	80.65%	75.25%	3.7643	0.3366
≥ 38	65.71%	80.65%	70.30%	3.3952	0.4251
≥ 39	61.43%	80.65%	67.33%	3.1738	0.4783
≥ 40	58.57%	80.65%	65.35%	3.0262	0.5137
≥ 41	55.71%	83.87%	64.36%	3.4543	0.528
≥ 42	54.29%	83.87%	63.37%	3.3657	0.5451
≥ 43	45.71%	83.87%	57.43%	2.8343	0.6473
≥ 44	44.29%	87.10%	57.43%	3.4321	0.6397
≥ 45	41.43%	90.32%	56.44%	4.281	0.6485
≥ 46	40.00%	90.32%	55.45%	4.1333	0.6643
≥ 48	35.71%	93.55%	53.47%	5.5357	0.6872
≥ 49	35.71%	96.77%	54.46%	11.0715	0.6643

≥ 50	34.29%	96.77%	53.47%	10.6286	0.679
≥ 51	32.86%	96.77%	52.48%	10.1857	0.6938
≥ 53	31.43%	96.77%	51.49%	9.7429	0.7086
≥ 58	30.00%	96.77%	50.50%	9.3	0.7233
≥ 61	27.14%	96.77%	48.51%	8.4143	0.7529
≥ 62	24.29%	96.77%	46.53%	7.5286	0.7824
≥ 63	22.86%	96.77%	45.54%	7.0857	0.7971
≥ 64	20.00%	96.77%	43.56%	6.2	0.8267
≥ 65	17.14%	96.77%	41.58%	5.3143	0.8562
≥ 67	15.71%	100.00%	41.58%		0.8429
≥ 68	14.29%	100.00%	40.59%		0.8571
≥ 73	12.86%	100.00%	39.60%		0.8714
≥ 75	11.43%	100.00%	38.61%		0.8857
≥ 77	10.00%	100.00%	37.62%		0.9
≥ 79	8.57%	100.00%	36.63%		0.9143
≥ 80	7.14%	100.00%	35.64%		0.9286
≥ 82	5.71%	100.00%	34.65%		0.9429
≥ 85	4.29%	100.00%	33.66%		0.9571
≥ 87	2.86%	100.00%	32.67%		0.9714
≥ 92	1.43%	100.00%	31.68%		0.9857
> 92	0.00%	100.00%	30.69%		1

AUC = 0.83 (95% CI = 0.74, 0.92)

Table 4.9: The ROC curve analysis of the raw ICON scores (rICON) as a measure of orthodontic treatment need:

rICON Cutoff point	Sensitivity	Specificity	Classified	LR+	LR-
≥ 1	100.00%	0.00%	69.31%	1	
≥ 2	100.00%	6.45%	71.29%	1.069	0
≥ 3	100.00%	12.90%	73.27%	1.1481	0
≥ 4	97.14%	22.58%	74.26%	1.2548	0.1265
≥ 5	97.14%	38.71%	79.21%	1.585	0.0738
≥ 6	85.71%	54.84%	76.24%	1.898	0.2605
≥ 7	72.86%	74.19%	73.27%	2.8232	0.3658
≥ 8	61.43%	83.87%	68.32%	3.8086	0.4599
≥ 9	47.14%	83.87%	58.42%	2.9229	0.6302
≥ 10	37.14%	96.77%	55.45%	11.5143	0.6495
≥ 11	28.57%	96.77%	49.50%	8.8572	0.7381
≥ 12	24.29%	100.00%	47.52%		0.7571
≥ 13	15.71%	100.00%	41.58%		0.8429
≥ 14	10.00%	100.00%	37.62%		0.9
≥ 15	4.29%	100.00%	33.66%		0.9571
≥ 16	2.86%	100.00%	32.67%		0.9714
> 16	0.00%	100.00%	30.69%		1

AUC = 0.81 (95% CI = 0.72, 0.90)

Table 4.10: The ROC curve analysis of the PAR as a measure of orthodontic treatment need:

PAR Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 4	100.00%	0.00%	69.31%	1	
≥ 6	100.00%	6.45%	71.29%	1.069	0
≥ 7	100.00%	19.35%	75.25%	1.24	0
≥ 8	100.00%	25.81%	77.23%	1.3478	0
≥ 9	95.71%	35.48%	77.23%	1.4836	0.1208
≥ 10	94.29%	38.71%	77.23%	1.5383	0.1476
≥ 11	92.86%	38.71%	76.24%	1.515	0.1845
≥ 12	91.43%	41.94%	76.24%	1.5746	0.2044
≥ 13	90.00%	41.94%	75.25%	1.55	0.2385
≥ 14	88.57%	48.39%	76.24%	1.7161	0.2362
≥ 15	82.86%	51.61%	73.27%	1.7124	0.3321
≥ 16	80.00%	54.84%	72.28%	1.7714	0.3647
≥ 17	77.14%	61.29%	72.28%	1.9929	0.3729
≥ 18	72.86%	64.52%	70.30%	2.0532	0.4207
≥ 19	70.00%	67.74%	69.31%	2.17	0.4429
≥ 20	68.57%	70.97%	69.31%	2.3619	0.4429
≥ 21	67.14%	80.65%	71.29%	3.469	0.4074
≥ 22	65.71%	87.10%	72.28%	5.0929	0.3937
≥ 23	62.86%	90.32%	71.29%	6.4952	0.4112
≥ 24	60.00%	93.55%	70.30%	9.3	0.4276
≥ 25	55.71%	93.55%	67.33%	8.6357	0.4734
≥ 26	50.00%	96.77%	64.36%	15.5	0.5167
≥ 27	45.71%	96.77%	61.39%	14.1715	0.561
≥ 28	41.43%	96.77%	58.42%	12.8429	0.6052
≥ 29	37.14%	96.77%	55.45%	11.5143	0.6495
≥ 30	35.71%	96.77%	54.46%	11.0715	0.6643
≥ 31	31.43%	100.00%	52.48%		0.6857
≥ 32	28.57%	100.00%	50.50%		0.7143
≥ 33	24.29%	100.00%	47.52%		0.7571
≥ 34	21.43%	100.00%	45.54%		0.7857
≥ 35	18.57%	100.00%	43.56%		0.8143
≥ 36	17.14%	100.00%	42.57%		0.8286
≥ 37	11.43%	100.00%	38.61%		0.8857
≥ 38	10.00%	100.00%	37.62%		0.9
≥ 39	8.57%	100.00%	36.63%		0.9143
≥ 40	7.14%	100.00%	35.64%		0.9286
≥ 42	4.29%	100.00%	33.66%		0.9571
≥ 45	1.43%	100.00%	31.68%		0.9857
> 45	0.00%	100.00%	30.69%		1

AUC = 0.82 (95% CI = 0.74 0.90)

Table 4.11: The ROC curve analysis of the raw PAR scores (rPAR) as a measure of orthodontic treatment need:

rPAR Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 3	100.00%	0.00%	69.31%	1	
≥ 4	100.00%	6.45%	71.29%	1.069	0
≥ 5	100.00%	9.68%	72.28%	1.1071	0
≥ 6	100.00%	22.58%	76.24%	1.2917	0
≥ 7	97.14%	32.26%	77.23%	1.434	0.0886
≥ 8	97.14%	41.94%	80.20%	1.673	0.0681
≥ 9	90.00%	51.61%	78.22%	1.86	0.1938
≥ 10	81.43%	58.06%	74.26%	1.9418	0.3198
≥ 11	74.29%	58.06%	69.31%	1.7714	0.4429
≥ 12	72.86%	67.74%	71.29%	2.2586	0.4007
≥ 13	65.71%	77.42%	69.31%	2.9102	0.4429
≥ 14	55.71%	87.10%	65.35%	4.3179	0.5085
≥ 15	47.14%	90.32%	60.40%	4.8714	0.5852
≥ 16	34.29%	100.00%	54.46%		0.6571
≥ 17	27.14%	100.00%	49.50%		0.7286
≥ 18	24.29%	100.00%	47.52%		0.7571
≥ 19	20.00%	100.00%	44.55%		0.8
≥ 20	17.14%	100.00%	42.57%		0.8286
≥ 21	14.29%	100.00%	40.59%		0.8571
≥ 22	12.86%	100.00%	39.60%		0.8714
≥ 23	7.14%	100.00%	35.64%		0.9286
≥ 24	2.86%	100.00%	32.67%		0.9714
≥ 25	1.43%	100.00%	31.68%		0.9857
> 25	0.00%	100.00%	30.69%		1

AUC = 0.81 (95% CI = 0.73, 0.90)

Table 4.12: The ROC curve analysis of the DI as a measure of orthodontic treatment need:

DI Cutpoint	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 2	100.00%	0.00%	69.31%	1	
≥ 3	100.00%	3.23%	70.30%	1.0333	0
≥ 4	97.14%	6.45%	69.31%	1.0384	0.4429
≥ 5	95.71%	9.68%	69.31%	1.0597	0.4429
≥ 6	92.86%	22.58%	71.29%	1.1994	0.3163
≥ 7	90.00%	32.26%	72.28%	1.3286	0.31
≥ 8	88.57%	35.48%	72.28%	1.3729	0.3221
≥ 9	82.86%	45.16%	71.29%	1.5109	0.3796
≥ 11	81.43%	48.39%	71.29%	1.5777	0.3838
≥ 12	71.43%	51.61%	65.35%	1.4762	0.5536
≥ 13	68.57%	54.84%	64.36%	1.5184	0.5731
≥ 14	62.86%	61.29%	62.38%	1.6238	0.606
≥ 15	60.00%	67.74%	62.38%	1.86	0.5905
≥ 16	57.14%	67.74%	60.40%	1.7714	0.6327
≥ 17	52.86%	74.19%	59.41%	2.0482	0.6354
≥ 18	50.00%	74.19%	57.43%	1.9375	0.6739
≥ 19	47.14%	74.19%	55.45%	1.8268	0.7124
≥ 20	42.86%	77.42%	53.47%	1.898	0.7381
≥ 21	32.86%	83.87%	48.51%	2.0371	0.8005
≥ 22	32.86%	87.10%	49.50%	2.5464	0.7709
≥ 23	28.57%	90.32%	47.52%	2.9524	0.7908
≥ 24	27.14%	90.32%	46.53%	2.8048	0.8066
≥ 25	27.14%	93.55%	47.52%	4.2071	0.7788
≥ 26	27.14%	96.77%	48.51%	8.4143	0.7529
≥ 27	20.00%	96.77%	43.56%	6.2	0.8267
≥ 28	18.57%	96.77%	42.57%	5.7572	0.8414
≥ 29	17.14%	96.77%	41.58%	5.3143	0.8562
≥ 30	15.71%	96.77%	40.59%	4.8714	0.871
≥ 31	12.86%	96.77%	38.61%	3.9857	0.9005
≥ 32	11.43%	96.77%	37.62%	3.5429	0.9152
≥ 34	8.57%	96.77%	35.64%	2.6571	0.9448
≥ 37	5.71%	96.77%	33.66%	1.7714	0.9743
≥ 40	4.29%	96.77%	32.67%	1.3286	0.989
≥ 45	2.86%	100.00%	32.67%		0.9714
≥ 46	1.43%	100.00%	31.68%		0.9857
> 46	0.00%	100.00%	30.69%		1

AUC = 0.69 (95% CI = 0.57, 0.80)

Table 4.13: The ROC curve analysis of the DI dental component (dDI) as a measure of orthodontic treatment need:

dDI Cutoff point	Sensitivity	Specificity	Classified	LR+	LR-
≥ 0	100.00%	0.00%	69.31%	1	
≥ 1	100.00%	3.23%	70.30%	1.0333	0
≥ 2	100.00%	6.45%	71.29%	1.069	0
≥ 3	100.00%	16.13%	74.26%	1.1923	0
≥ 4	95.71%	19.35%	72.28%	1.1869	0.2214
≥ 5	91.43%	22.58%	70.30%	1.181	0.3796
≥ 6	87.14%	48.39%	75.25%	1.6884	0.2657
≥ 7	80.00%	64.52%	75.25%	2.2545	0.31
≥ 8	75.71%	74.19%	75.25%	2.9339	0.3273
≥ 9	70.00%	83.87%	74.26%	4.34	0.3577
≥ 10	58.57%	90.32%	68.32%	6.0524	0.4587
≥ 11	47.14%	96.77%	62.38%	14.6143	0.5462
≥ 12	34.29%	96.77%	53.47%	10.6286	0.679
≥ 13	25.71%	96.77%	47.52%	7.9714	0.7676
≥ 14	24.29%	96.77%	46.53%	7.5286	0.7824
≥ 15	18.57%	96.77%	42.57%	5.7572	0.8414
≥ 16	15.71%	96.77%	40.59%	4.8714	0.871
≥ 17	14.29%	100.00%	40.59%		0.8571
≥ 18	11.43%	100.00%	38.61%		0.8857
≥ 19	10.00%	100.00%	37.62%		0.9
≥ 20	5.71%	100.00%	34.65%		0.9429
≥ 22	2.86%	100.00%	32.67%		0.9714
≥ 26	1.43%	100.00%	31.68%		0.9857
> 26	0.00%	100.00%	30.69%		1

AUC = 0.81 (95% CI = 0.73, 0.90)

Table 4.14: Multiple linear regression analysis showing association between need for orthodontic treatment and cIOTN components:

cIOTN	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
AC	0.2653685	0.040674	6.52	<0.0001	0.1846523	0.3460847
DHC	0.5785828	0.0834718	6.93	<0.0001	0.4129357	0.7442299
_cons	1.752265	0.2636186	6.65	<0.0001	1.229122	2.275407

R-squared 0. 6377

Table 4.15: Multiple linear regression analysis showing association between need for orthodontic treatment and rPAR components:

rPAR	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
Upper ant. segment	0.2071978	0.0299944	6.91	<0.0001	0.1476347	0.2667608
Lower ant. segment	0.0004877	0.0320044	0.02	0.988	-0.0630668	0.0640422
Buccal ap. occlusion	0.1121534	0.0570301	1.97	0.052	-0.0010971	0.225404
Buccal tv. occlusion	- 0.0626303	0.0700161	-0.89	0.373	-0.2016683	0.0764078
Buccal v. occlusion	(omitted)*					
Over-jet	0.2474678	0.0759739	3.26	0.002	0.0965987	0.3983369
Over-bite	0.0070622	0.098305	0.07	0.943	-0.1881519	0.2022764
Central line	0.6196748	0.158491	3.91	<0.0001	0.304943	0.9344066
_cons	3.132841	0.2211336	14.17	<0.0001	2.693714	3.571969

R-squared 0.5762

*no variability in buccal vertical occlusal problem to calculate the regression.

Table 4.16: Multiple linear regression analysis showing association between need for orthodontic treatment and rICON components:

rICON	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
AC	0.3529777	0.0476335	7.41	<0.0001	0.2584002	0.4475553
Upper crowding	0.1619111	0.0728994	2.22	0.029	0.0171675	0.3066546
Upper spacing	-0.0233585	0.1181505	-0.2	0.844	-0.257949	0.211232
Cross-bite	0.0144821	0.1917771	0.08	0.94	-0.3662959	0.3952601
Over-bite	-0.0024878	0.1093918	-0.02	0.982	-0.2196877	0.2147122
Buccal ap. occlusion	0.1376491	0.0577998	2.38	0.019	0.0228862	0.2524119
_cons	3.079133	0.2194434	14.03	<0.0001	2.643423	3.514843

R-squared 0.5216

Table 4.17: Multiple linear regression analysis showing association between need for orthodontic treatment and rICON components:

DI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Aesthetic Component	0.3477831	0.04543	7.66	<0.0001	0.2576288	0.4379374
Dental component*	0.1189541	0.0405741	2.93	0.004	0.0384362	0.199472
_cons	2.964944	0.2004711	14.79	<0.0001	2.567116	3.362772

R-squared 0.5036

*Dental component is the sum of all raw scores of the dental occlusal components present in ICON (upper crowding and spacing, cross-bite, over-bite, and buccal occlusion)

Table 4.18: Multiple linear regression analysis showing association between need for orthodontic treatment and dDI components:

DI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Over-jet	0.2324824	0.0598388	3.89	<0.0001	0.1136199	0.3513449
Over-bite	0.1106872	0.0569302	1.94	0.055	-0.0023976	0.223772
Ant. Open-bite	0.0712350	0.0605221	1.18	0.242	-0.0489848	0.1914547
L. Open-bite	-0.0749766	0.0847474	-0.88	0.379	-0.243317	0.0933637
Crowding	0.1930331	0.052625	3.67	<0.0001	0.0885001	0.2975661
Buccal Occlusion	0.0956058	0.0407602	2.35	0.021	0.0146406	0.1765709
Lingual Cross-bite	0.0952353	0.1020484	0.93	0.353	-0.1074713	0.2979418
Buccal Cross-bite	-0.0087542	0.1705213	-0.05	0.959	-0.3474737	0.3299654
Others	0.2097342	0.0461536	4.54	<0.0001	0.1180558	0.3014127
_cons	3.3378650	0.1929482	17.3	<0.0001	2.954597	3.721133

R-squared 0.5059

Table 4.19: ROC analysis of nICON as measure orthodontic treatment need:

nICON Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 3	100.00%	0.00%	69.31%	1	
≥ 4	100.00%	6.45%	71.29%	1.069	0
≥ 5	100.00%	12.90%	73.27%	1.1481	0
≥ 6	100.00%	19.35%	75.25%	1.24	0
≥ 7	100.00%	29.03%	78.22%	1.4091	0
≥ 8	97.14%	32.26%	77.23%	1.434	0.0886
≥ 9	97.14%	41.94%	80.20%	1.673	0.0681
≥ 10	92.86%	58.06%	82.18%	2.2143	0.123
≥ 11	90.00%	58.06%	80.20%	2.1462	0.1722
≥ 12	84.29%	64.52%	78.22%	2.3753	0.2436
≥ 13	78.57%	80.65%	79.21%	4.0595	0.2657
≥ 14	67.14%	83.87%	72.28%	4.1629	0.3918
≥ 15	55.71%	83.87%	64.36%	3.4543	0.528
≥ 16	45.71%	93.55%	60.40%	7.0857	0.5803
≥ 17	40.00%	93.55%	56.44%	6.2	0.6414
≥ 18	38.57%	96.77%	56.44%	11.9572	0.6348
≥ 19	31.43%	96.77%	51.49%	9.7429	0.7086
≥ 20	28.57%	96.77%	49.50%	8.8572	0.7381
≥ 22	27.14%	96.77%	48.51%	8.4143	0.7529
≥ 23	24.29%	96.77%	46.53%	7.5286	0.7824
≥ 24	18.57%	96.77%	42.57%	5.7572	0.8414
≥ 25	12.86%	96.77%	38.61%	3.9857	0.9005
≥ 26	12.86%	100.00%	39.60%		0.8714
≥ 28	11.43%	100.00%	38.61%		0.8857
≥ 29	7.14%	100.00%	35.64%		0.9286
≥ 30	5.71%	100.00%	34.65%		0.9429
≥ 31	4.29%	100.00%	33.66%		0.9571
≥ 32	2.86%	100.00%	32.67%		0.9714
≥ 34	1.43%	100.00%	31.68%		0.9857
> 34	0.00%	100.00%	30.69%		1

AUC = 0.85 (95% CI = 0.76, 0.93)

Table 4.20: ROC analysis of nPAR as measure orthodontic treatment need:

nPAR Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	69.31%	1	
≥ 1	100.00%	3.23%	70.30%	1.0333	0
≥ 2	100.00%	9.68%	72.28%	1.1071	0
≥ 3	100.00%	19.35%	75.25%	1.24	0
≥ 4	98.57%	32.26%	78.22%	1.4551	0.0443
≥ 5	92.86%	41.94%	77.23%	1.5992	0.1703
≥ 6	85.71%	74.19%	82.18%	3.3214	0.1925
≥ 7	74.29%	83.87%	77.23%	4.6057	0.3066
≥ 8	62.86%	93.55%	72.28%	9.7429	0.397
≥ 9	55.71%	96.77%	68.32%	17.2715	0.4576
≥ 10	47.14%	100.00%	63.37%		0.5286
≥ 11	32.86%	100.00%	53.47%		0.6714
≥ 12	28.57%	100.00%	50.50%		0.7143
≥ 13	18.57%	100.00%	43.56%		0.8143
≥ 14	11.43%	100.00%	38.61%		0.8857
≥ 16	5.71%	100.00%	34.65%		0.9429
≥ 17	4.29%	100.00%	33.66%		0.9571
≥ 20	1.43%	100.00%	31.68%		0.9857
> 20	0.00%	100.00%	30.69%		1

AUC = 0.88 (95% CI = 0.81, 0.95)

Table 4.21: ROC analysis of ndDI as measure orthodontic treatment need:

nDId Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	69.31%	1	
≥ 1	100.00%	3.23%	70.30%	1.0333	0
≥ 2	100.00%	6.45%	71.29%	1.069	0
≥ 3	100.00%	12.90%	73.27%	1.1481	0
≥ 4	95.71%	22.58%	73.27%	1.2363	0.1898
≥ 5	92.86%	25.81%	72.28%	1.2516	0.2768
≥ 6	91.43%	38.71%	75.25%	1.4917	0.2214
≥ 7	88.57%	48.39%	76.24%	1.7161	0.2362
≥ 8	87.14%	61.29%	79.21%	2.2512	0.2098
≥ 9	78.57%	70.97%	76.24%	2.7063	0.3019
≥ 10	74.29%	80.65%	76.24%	3.8381	0.3189
≥ 11	68.57%	87.10%	74.26%	5.3143	0.3608
≥ 12	58.57%	96.77%	70.30%	18.1572	0.4281
≥ 13	51.43%	96.77%	65.35%	15.9429	0.5019
≥ 14	41.43%	100.00%	59.41%		0.5857
≥ 15	34.29%	100.00%	54.46%		0.6571
≥ 16	28.57%	100.00%	50.50%		0.7143
≥ 17	25.71%	100.00%	48.51%		0.7429
≥ 18	20.00%	100.00%	44.55%		0.8
≥ 19	18.57%	100.00%	43.56%		0.8143
≥ 20	17.14%	100.00%	42.57%		0.8286
≥ 21	11.43%	100.00%	38.61%		0.8857
≥ 22	7.14%	100.00%	35.64%		0.9286
≥ 23	5.71%	100.00%	34.65%		0.9429
≥ 24	4.29%	100.00%	33.66%		0.9571
≥ 27	2.86%	100.00%	32.67%		0.9714
≥ 31	1.43%	100.00%	31.68%		0.9857
> 31	0.00%	100.00%	30.69%		1

AUC = 0.85 (95% CI = 0.78, 0.92)

Table 4.22: Spearman correlation analysis between the complexity of orthodontic treatment and different occlusal indices:

Complexity	Indices without FI	P-value	Indices with FI	P-value
AC	0.4301	< 0.0001	0.5472	< 0.0001
DHC	0.7219	< 0.0001	0.6756	< 0.0001
cIOTN*	0.6190	< 0.0001	0.6728	< 0.0001
rPAR**	0.4384	< 0.0001	0.4714	< 0.0001
PAR	0.4695	< 0.0001	0.4818	< 0.0001
rICON**	0.4945	< 0.0001	0.5657	< 0.0001
ICON	0.4827	< 0.0001	0.5013	< 0.0001
DI	0.4257	< 0.0001	0.4437	< 0.0001
cDI	0.2031	0.0417	0.2497	0.0118
dDI	0.6063	< 0.0001	0.6068	< 0.0001

* combined IOTN indicates adding the AC and the DHC together into one score
**raw indicates adding the raw scores into a total score without using any weights

Table 4.23: Spearman correlation analysis between the estimated treatment duration and the different occlusal indices:

Duration	Indices without FI	P-value	Indices with FI	P-value
AC	0.5169	< 0.0001	0.6163	< 0.0001
DHC	0.6871	< 0.0001	0.6402	< 0.0001
cIOTN*	0.6553	< 0.0001	0.7077	< 0.0001
rPAR**	0.5169	< 0.0001	0.5603	< 0.0001
PAR	0.5381	< 0.0001	0.5568	< 0.0001
rICON**	0.6390	< 0.0001	0.6910	< 0.0001
ICON	0.6043	< 0.0001	0.6237	< 0.0001
DI	0.3756	0.0001	0.4000	< 0.0001
cDI	0.1023	0.2939	0.1632	0.1048
dDI	0.6554	< 0.0001	0.6510	< 0.0001

* combined IOTN indicates adding the AC and the DHC together into one score
**raw indicates adding the raw scores into a total score without using any weights

Table 4.24: Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the different occlusal indices:

Improvement	Indices without FI	P-value	Indices with FI	P-value
AC	-0.4677	< 0.0001	-0.5155	< 0.0001
DHC	-0.4822	< 0.0001	-0.4953	< 0.0001
cIOTN*	-0.5423	< 0.0001	-0.5712	< 0.0001
r.PAR**	-0.5161	< 0.0001	-0.5366	< 0.0001
PAR	-0.5328	< 0.0001	-0.5391	< 0.0001
rICON**	-0.5259	< 0.0001	-0.5532	< 0.0001
ICON	-0.5301	< 0.0001	-0.5406	< 0.0001
DI	-0.5997	< 0.0001	-0.5941	< 0.0001
cDI	-0.2946	0.0028	-0.3119	0.0015
dDI	-0.5585	< 0.0001	-0.5616	< 0.0001

* combined IOTN indicates adding the AC and the DHC together into one score

**raw indicates adding the raw scores into a total score without using any weights

Table 4.25 a: Multivariable analysis showing association between the malocclusion improvement by orthodontic treatment and the different occlusal indices:

Improvement	Coef.	Std. Err.	t	P>t	[95%Conf.	Interval]
AC						
pre	0.1735167	0.041042	4.23	< 0.0001	0.0920703	0.2549632
post	-0.9377486	0.1314312	-7.13	< 0.0001	-1.19857	-0.6769276
DHC						
pre	0.4401455	0.0910486	4.83	< 0.0001	0.2594624	0.6208285
post	-0.7317092	0.1292626	-5.66	< 0.0001	-0.9882267	-0.4751918
IOTN						
pre	0.1511606	0.0298637	5.06	< 0.0001	0.091897	0.2104242
post	-0.5658402	0.0720459	-7.85	< 0.0001	-0.7088129	-0.4228675
rPAR						
pre	0.0777689	0.0142407	5.46	< 0.0001	0.0495086	0.1060292
post	-0.268941	0.0346331	-7.77	< 0.0001	-0.3376692	-0.2002127
PAR						
pre	0.0449737	0.0075623	5.95	< 0.0001	0.0299666	0.0599808
post	-0.1106815	0.0157516	-7.03	< 0.0001	-0.1419401	-0.079423
rICON						
pre	0.101615	0.0218696	4.65	< 0.0001	0.0582155	0.1450145
post	-0.3855766	0.0503202	-7.66	< 0.0001	-0.4854353	-0.2857178
ICON						
pre	0.0181699	0.0038242	4.75	< 0.0001	0.0105808	0.0257589
post	-0.0874448	0.0104659	-8.36	< 0.0001	-0.108214	-0.0666756
DI						
pre	0.0781869	0.0100731	7.76	< 0.0001	0.0581972	0.0981767
post	-0.1095486	0.0116113	-9.43	< 0.0001	-0.1325909	-0.0865062
dDI						
pre	0.0893829	0.0147728	6.05	< 0.0001	0.0600667	0.1186991
post	-0.2184873	0.0258545	-8.45	< 0.0001	-0.2697947	-0.16718
cDI						
pre	0.085153	0.0244746	3.48	0.001	0.036584	0.1337219
post	-0.1031849	0.0233447	-4.42	< 0.0001	-0.1495116	-0.0568581

Table 4.25 b: Multivariable analysis showing association between the malocclusion improvement by orthodontic treatment and the different occlusal indices:

Occlusal index	Pre-treatment β	Post-treatment β	Weight
AC	0.17	-0.94	5.53
DHC	0.44	-0.73	1.65
cIOTN	0.15	-0.57	3.8
rPAR	0.08	-0.27	3.48
PAR	0.05	-0.11	2.20
rICON	0.10	-0.39	3.90
ICON	0.02	-0.09	4.50
DI	0.08	-0.11	1.38
dDI	0.09	-0.22	2.44
cDI	0.09	-0.10	1.11

Table 4.26 a: Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the weighted occlusal indices:

Weight	AC	DHC	PAR	wPAR	ICON	wICON
No weight	-0.4677	-0.4822	-0.5161	-0.5328	-0.5259	-0.5301
Exclusive weight *	-0.5844	-0.4754	-0.5879	-0.5578	-0.6304	-0.6547
Average weight **	-0.5992	-0.4504	-0.5904	-0.5695	-0.6258	-0.6436

* The exclusive weights used for the post-treatment scores were obtained from the multiple linear regressions and used to compute occlusal index scores: ACx5, DHCx1.5, rPARx3.5, PARx2, rICONx4, ICONx4.5, and DIx1.5
** The average weight for the post-treatment scores that was used is 3

Table 4.26 b: Spearman correlation analysis between the malocclusion improvement by orthodontic treatment and the weighted occlusal indices:

Weight	cIOTN	DI
No weight	-0.5423	-0.5997
Exclusive weight *	-0.5987	-0.6004
Average weight **	-0.6121	-0.4358
Components weight ***	-0.6383	-0.6378

* The exclusive weights used for the post-treatment scores were obtained from the multiple linear regressions and used to compute occlusal index scores: IOTNx4, DIx1.5
** The average weight for the post-treatment scores that was used is 3
*** Only one component of the index were weighted: IOTN-ACx3; DI-dDIx3

Table 4.27: The ROC curve analysis of the AC as a measure of orthodontic treatment outcome:

AC Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 1	100.00%	0.00%	10.89%	1	
≥ 2	72.73%	85.56%	84.16%	5.035	0.3188
≥ 3	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 4	18.18%	100.00%	91.09%		0.8182
> 4	0.00%	100.00%	89.11%		1

AUC = 0.81 (95% CI = 0.66, 0.96)

Table 4.28: The ROC curve analysis of the DHC as a measure orthodontic treatment outcome:

DHC Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 1	100.00%	0.00%	10.89%	1	
≥ 2	72.73%	58.89%	60.40%	1.769	0.4631
≥ 3	27.27%	97.78%	90.10%	12.2727	0.7438
≥ 4	9.09%	98.89%	89.11%	8.1818	0.9193
> 4	0.00%	100.00%	89.11%		1

AUC = 0.71 (95% CI = 0.78, 0.92)

Table 4.29: The ROC curve analysis of the cIOTN as a measure orthodontic treatment outcome:

CIOTN Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 2	100.00%	0.00%	10.89%	1	
≥ 3	90.91%	55.56%	59.41%	2.0455	0.1636
≥ 4	72.73%	86.67%	85.15%	5.4545	0.3147
≥ 5	36.36%	98.89%	92.08%	32.7272	0.6435
≥ 6	9.09%	98.89%	89.11%	8.1818	0.9193
≥ 8	9.09%	100.00%	90.10%		0.9091
> 8	0.00%	100.00%	89.11%		1

AUC = 0.85 (95% CI = 0.72, 0.99)

Table 4.30: The ROC curve analysis of the rPAR as a measure orthodontic treatment outcome:

rPAR Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	10.89%	1	
≥ 1	81.82%	26.67%	32.67%	1.1157	0.6818
≥ 2	72.73%	57.78%	59.41%	1.7225	0.472
≥ 3	54.55%	85.56%	82.18%	3.7762	0.5313
≥ 4	45.45%	90.00%	85.15%	4.5455	0.6061
≥ 5	36.36%	94.44%	88.12%	6.5455	0.6738
≥ 6	27.27%	96.67%	89.11%	8.1818	0.7524
≥ 7	27.27%	97.78%	90.10%	12.2727	0.7438
≥ 9	18.18%	98.89%	90.10%	16.3636	0.8274
≥ 10	18.18%	100.00%	91.09%		0.8182
≥ 11	9.09%	100.00%	90.10%		0.9091
> 11	0.00%	100.00%	89.11%		1

AUC = 0.71 (95% CI = 0.51, 0.92)

Table 4.31: The ROC curve analysis of the PAR as a measure of orthodontic treatment outcome:

PAR Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	10.89%	1	
≥ 1	81.82%	26.67%	32.67%	1.1157	0.6818
≥ 2	72.73%	55.56%	57.43%	1.6364	0.4909
≥ 3	63.64%	76.67%	75.25%	2.7273	0.4743
≥ 4	63.64%	81.11%	79.21%	3.369	0.4483
≥ 5	54.55%	84.44%	81.19%	3.5065	0.5383
≥ 6	45.45%	84.44%	80.20%	2.9221	0.6459
≥ 7	45.45%	87.78%	83.17%	3.719	0.6214
≥ 8	36.36%	93.33%	87.13%	5.4545	0.6818
≥ 9	36.36%	94.44%	88.12%	6.5455	0.6738
≥ 10	36.36%	95.56%	89.11%	8.1818	0.666
≥ 13	27.27%	97.78%	90.10%	12.2727	0.7438
≥ 16	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 17	18.18%	98.89%	90.10%	16.3636	0.8274
≥ 23	9.09%	98.89%	89.11%	8.1818	0.9193
≥ 32	9.09%	100.00%	90.10%		0.9091
> 32	0.00%	100.00%	89.11%		1

AUC = 0.71 (95% CI = 0.51, 0.92)

Table 4.32: The ROC curve analysis of the rICON as a measure of orthodontic treatment outcome:

rICON Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 1	100.00%	0.00%	10.89%	1	
≥ 2	90.91%	31.11%	37.62%	1.3196	0.2922
≥ 3	63.64%	73.33%	72.28%	2.3864	0.4959
≥ 4	45.45%	95.56%	90.10%	10.2273	0.5708
≥ 5	45.45%	97.78%	92.08%	20.4546	0.5579
≥ 6	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 7	9.09%	98.89%	89.11%	8.1818	0.9193
≥ 11	9.09%	100.00%	90.10%		0.9091
> 11	0.00%	100.00%	89.11%		1

AUC = 0.76 (95% CI = 0.59, 0.93)

Table 4.33: The ROC curve analysis of the ICON as a measure of orthodontic treatment outcome:

ICON Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 7	100.00%	0.00%	10.89%	1	
≥ 10	90.91%	31.11%	37.62%	1.3196	0.2922
≥ 11	90.91%	63.33%	66.34%	2.4793	0.1435
≥ 12	90.91%	66.67%	69.31%	2.7273	0.1364
≥ 13	90.91%	68.89%	71.29%	2.9221	0.132
≥ 14	72.73%	81.11%	80.20%	3.8503	0.3362
≥ 15	45.45%	84.44%	80.20%	2.9221	0.6459
≥ 16	45.45%	85.56%	81.19%	3.1469	0.6375
≥ 17	45.45%	86.67%	82.18%	3.4091	0.6294
≥ 18	45.45%	91.11%	86.14%	5.1136	0.5987
≥ 19	45.45%	93.33%	88.12%	6.8182	0.5844
≥ 20	45.45%	95.56%	90.10%	10.2273	0.5708
≥ 21	45.45%	96.67%	91.09%	13.6364	0.5643
≥ 22	45.45%	97.78%	92.08%	20.4546	0.5579
≥ 26	45.45%	98.89%	93.07%	40.909	0.5516
≥ 30	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 31	27.27%	100.00%	92.08%		0.7273
≥ 33	18.18%	100.00%	91.09%		0.8182
≥ 54	9.09%	100.00%	90.10%		0.9091
> 54	0.00%	100.00%	89.11%		1

AUC = 0.83 (95% CI = 0.68, 0.98)

Table 4.34: The ROC curve analysis of the DI as a measure of orthodontic treatment outcome:

DI Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	10.89%	1	
≥ 1	90.91%	15.56%	23.76%	1.0766	0.5844
≥ 2	90.91%	18.89%	26.73%	1.1208	0.4813
≥ 3	90.91%	27.78%	34.65%	1.2587	0.3273
≥ 4	90.91%	30.00%	36.63%	1.2987	0.303
≥ 5	81.82%	37.78%	42.57%	1.3149	0.4813
≥ 6	81.82%	40.00%	44.55%	1.3636	0.4545
≥ 7	81.82%	47.78%	51.49%	1.5667	0.3805
≥ 8	81.82%	51.11%	54.46%	1.6736	0.3557
≥ 9	81.82%	56.67%	59.41%	1.8881	0.3209
≥ 10	63.64%	61.11%	61.39%	1.6364	0.595
≥ 11	45.45%	68.89%	66.34%	1.461	0.7918
≥ 12	45.45%	73.33%	70.30%	1.7045	0.7438
≥ 13	45.45%	77.78%	74.26%	2.0455	0.7013
≥ 14	45.45%	80.00%	76.24%	2.2727	0.6818
≥ 16	45.45%	82.22%	78.22%	2.5568	0.6634
≥ 17	45.45%	86.67%	82.18%	3.4091	0.6294
≥ 18	45.45%	90.00%	85.15%	4.5455	0.6061
≥ 20	45.45%	91.11%	86.14%	5.1136	0.5987
≥ 22	45.45%	94.44%	89.11%	8.1818	0.5775
≥ 24	36.36%	95.56%	89.11%	8.1818	0.666
≥ 26	36.36%	97.78%	91.09%	16.3636	0.6508
≥ 28	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 31	18.18%	98.89%	90.10%	16.3636	0.8274
≥ 37	9.09%	98.89%	89.11%	8.1818	0.9193
≥ 44	9.09%	100.00%	90.10%		0.9091
> 44	0.00%	100.00%	89.11%		1

AUC = 0.71 (95% CI = 0.52, 0.90)

Table 4.35: The ROC curve analysis of the dDI as a measure of orthodontic treatment outcome:

DI Cutoff point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
≥ 0	100.00%	0.00%	10.89%	1	
≥ 1	90.91%	65.56%	68.32%	2.6393	0.1387
≥ 2	90.91%	73.33%	75.25%	3.4091	0.124
≥ 3	63.64%	90.00%	87.13%	6.3636	0.404
≥ 4	54.55%	92.22%	88.12%	7.013	0.4929
≥ 5	36.36%	96.67%	90.10%	10.9091	0.6583
≥ 7	36.36%	98.89%	92.08%	32.7272	0.6435
≥ 8	27.27%	98.89%	91.09%	24.5454	0.7354
≥ 9	18.18%	98.89%	90.10%	16.3636	0.8274
≥ 10	18.18%	100.00%	91.09%		0.8182
≥ 21	9.09%	100.00%	90.10%		0.9091
> 21	0.00%	100.00%	89.11%		1

AUC = 0.87 (95% CI = 0.75, 0.99)

Table 4.36: Spearman correlation analysis of the need and complexity of orthodontic treatment with different cephalometric measurements:

Cephalometric measurements	Need	P-value	Complexity	P-value
ANB angle	0.0420	0.6764	0.0592	0.5565
PP-MP angle	0.0619	0.5385	0.0850	0.3983
MP-SN angle	0.0119	0.9061	0.0828	0.4102
I-SN angle	0.0964	0.3377	0.0314	0.7549
I-NA angle	0.1817	0.0690	0.1523	0.1284
i-MP angle	-0.0578	0.5661	0.0026	0.9793
i-NB angle	-0.1654	0.0983	-0.1006	0.3166
Nose Tip / HL	-0.1137	0.2576	-0.0530	0.5989
Subnasale / HL	0.0164	0.8708	0.0420	0.6764
Lower Lip / HL	-0.0504	0.6169	-0.0871	0.3864
Mental sulcus / HL	0.1221	0.2237	0.0403	0.6893
Upper lip / EL	0.1223	0.2229	0.0556	0.5811
Lower Lip /EL	0.0813	0.4189	0.0299	0.7667

Table 4.37: The association between the different cephalometric measurements and the outcome of the orthodontic treatment as assessed by the panel:

Cephalometric measurements	Unacceptable N=11 Mean (SD)	Acceptable N=90 Mean (SD)	P-value
ANB angle	4.4 (3.0)	4.2 (2.1)	0.7726
PP-MP angle	29.0 (7.0)	27.3 (5.6)	0.4261
MP-SN angle	39.2 (7.7)	38.1 (5.8)	0.5819
I-SN angle	103.4 (10.0)	101.3 (7.3)	0.6707
I-NA angle	22.9 (10.6)	20.9 (6.6)	0.6588
i-MP angle	95.3 (7.3)	94.5 (7.5)	0.7602
i-NB angle	30.6 (8.3)	28.9 (5.7)	0.2159
Nose Tip / HL	4.7 (3.3)	6.1 (4.2)	0.4357
Subnasale / HL	6.4 (1.6)	4.6 (2.7)	0.0157*
Lower Lip / HL	1.4 (2.2)	1.1 (1.3)	0.7600
Mental sulcus / HL	4.8 (2.2)	4.0 (1.9)	0.1228
Upper lip / EL	-2.7 (1.9)	-3.4 (2.4)	0.5379
Lower Lip /EL	-0.3 (3.2)	-1.0 (2.3)	0.7643

*P-value < 0.05

FIGURES

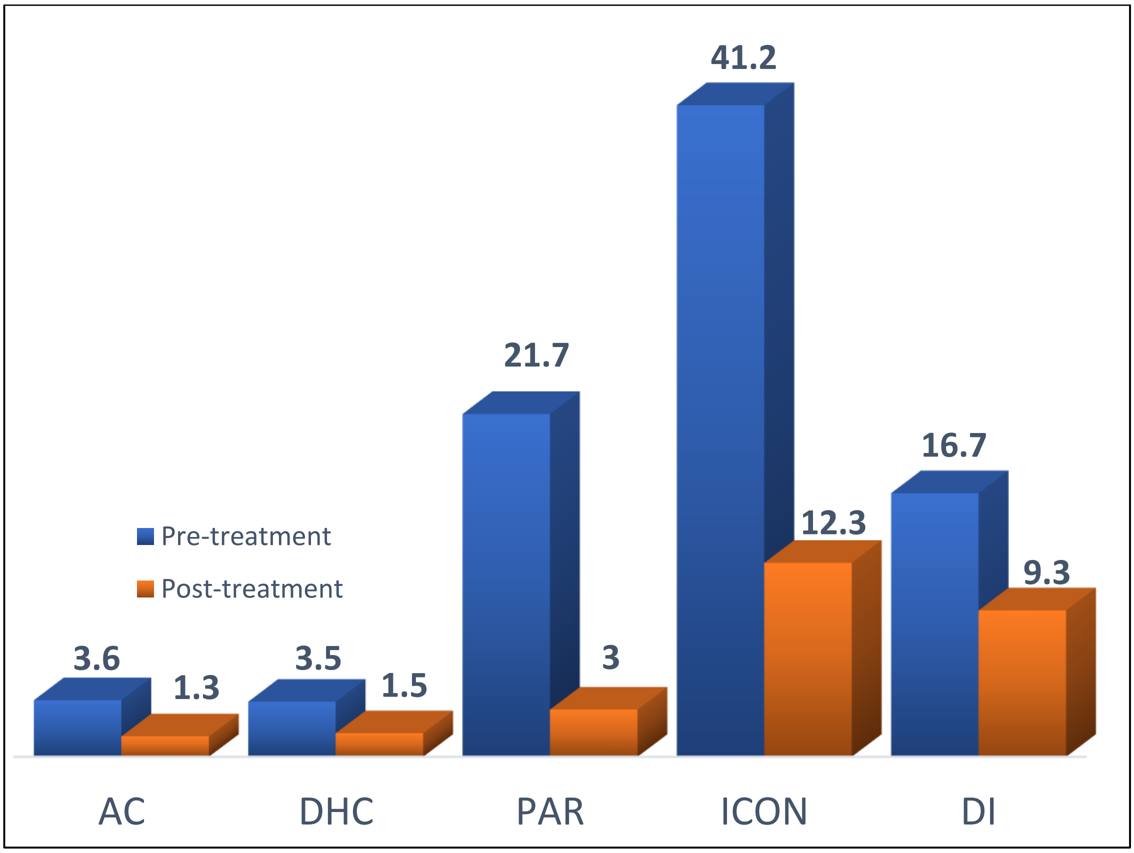

Figure 4.1: Pre- and post-treatment scores of the different indices.

REFERENCES

Ackerman, J. L., & Proffit, W. R. (1969). The characteristics of malocclusion: a modern approach to classification and diagnosis. *American journal of orthodontics, 56*(5), 443-454.

Andrews, L. F. (1972). The six keys to normal occlusion. *American journal of orthodontics*, *62*(3), 296-309.

Angle, E. H. (1899). Classification of malocclusion.

Angle, E. H. (1907). *Treatment of malocclusion of the teeth: Angle's system*: White Dental Manufacturing Company.

Ash, M. M., & Ramfjord, S. P. (1982). *An introduction to functional occlusion*: WB Saunders Company.

Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly*, 459-489.

Ballard, C., & Wayman, J. (1965). A report on a survey of the orthodontic requirements of 310 army apprentices. *The Dental practitioner and dental record, 15*, 221-226.

Baume, L. J. (1974). *A method for measuring occlusal traits*: Fédération Dentaire Internationale.

Beglin, F. M., Firestone, A. R., Vig, K. W., Beck, F. M., Kuthy, R. A., & Wade, D. (2001). A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics, 120*(3), 240-246.

Bergström, K., Halling, A., Huggare, J., & Johansson, L. (1998). Treatment difficulty and treatment outcome in orthodontic care. *The European Journal of Orthodontics, 20*(2), 145-157.

Birkeland, K., Katle, A., Løvgreen, S., Bøe, O. E., & Wisth, P. J. (1999). Factors influencing the decision about orthodontic treatment. *Journal of Orofacial Orthopedics/Fortschritte der Kieferorthopädie, 60*(5), 292-307.

Björk, A., Krebs, A., & Solow, B. (1964). A Method for Epidemiological Registration of Malocculusion. *Acta Odontologica Scandinavica, 22*(1), 27-41.

Borzabadi-Farahani, A. (2011). *An overview of selected orthodontic treatment need indices*: In Tech.

Borzabadi-Farahani, A., & Borzabadi-Farahani, A. (2011). Agreement between the index of complexity, outcome, and need and the dental and aesthetic components of the index of orthodontic treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics, 140*(2), 233-238.

Borzabadi-Farahani, A., Borzabadi-Farahani, A., & Eslamipour, F. (2009). Malocclusion and occlusal traits in an urban Iranian population. An epidemiological study of 11-to 14-year-old children. *The European Journal of Orthodontics*, cjp031.

Brook, P. H., & Shaw, W. C. (1989). The development of an index of orthodontic treatment priority. *The European Journal of Orthodontics, 11*(3), 309-320.

Bryant, S. (2002). The rationale for management of morphologic variations and nonphysiologic occlusion in the young dentition. *The International journal of prosthodontics, 16*, 75-77; discussion 89-90.

BSI., B. S. I. (1983). *British Standard Incisor Classification. Glossary of Dental Terms BS 4492*: British Standard Institute London, UK.

Buchanan, I., Downing, A., & Stirrups, D. (1994). A comparison of the Index of Orthodontic Treatment Need applied clinically and to diagnostic records. *British journal of orthodontics, 21*(2), 185-188.

Buchanan, I., Shaw, W., Richmond, S., O'Brien, K., & Andrews, M. (1993). A comparison of the reliability and validity of the PAR Index and Summers' Occlusal Index. *The European Journal of Orthodontics, 15*(1), 27-31.

Burden, D. J., Pine, C. M., & Burnside, G. (2001). Modified IOTN: an orthodontic treatment need index for use in oral health surveys. *Community dentistry and oral epidemiology, 29*(3), 220-225.

Cangialosi, T. J., Riolo, M. L., Owens, S. E., Dykhouse, V. J., Moffitt, A. H., Grubb, J. E., . . . James, R. D. (2004). The ABO discrepancy index: a measure of case complexity. *American Journal of Orthodontics and Dentofacial Orthopedics, 125*(3), 270-278.

Cardoso, C. F., Drummond, A. F., Lages, E., Pretti, H., Ferreira, E. F., & Abreu, M. H. N. (2011). The Dental Aesthetic Index and dental health component of the Index of Orthodontic Treatment Need as tools in epidemiological studies. *International journal of environmental research and public health, 8*(8), 3277-3286.

Casko, J. S., Vaden, J. L., Kokich, V. G., Damone, J., James, R. D., Cangialosi, T. J., . . . Bills, E. D. (1998). Objective grading system for dental casts and panoramic radiographs. *American Journal of Orthodontics and Dentofacial Orthopedics, 114*(5), 589-599.

Cons, N. C., Jenny, J., & Kohout, F. J. (1986). *Dental Aesthetic Index*: Health Quest.

Daniels, C., & Richmond, S. (2000). The Development of the Index of Complexity, Outcome and Need (ICON). *Journal of Orthodontics, 27*(2), 149.

Danyluk, K., Lavelle, C., & Hassard, T. (1999). Potential application of the dental aesthetic index to prioritize the orthodontic service needs in a publicly funded dental program. *American Journal of Orthodontics and Dentofacial Orthopedics, 116*(3), 279-286.

DeGuzman, L., Bahiraei, D., Vig, K., Vig, P., Weyant, R., & O'brien, K. (1995). The validation of the Peer Assessment Rating index for malocclusion severity and treatment difficulty. *American Journal of Orthodontics and Dentofacial Orthopedics, 107*(2), 172-176.

Dewey, M. (1915). Classification of malocclusion. *International Journal of Orthodontia, 1*(3), 133-147.

Draker, H. L. (1958). HANDICAPPING LABIO-LINGUAL DEVIATIONS: A PROPOSED INDEX FOR PUBLIC HEALTH PURPOSES. *Journal of Public Health Dentistry, 18*(4), 1-7.

Draker, H. L. (1960). Handicapping labio-lingual deviations: a proposed index for public health purposes. *American journal of orthodontics, 46*(4), 295-305.

Farahani, A., & Eslamipour, F. (2010). The relationship between ICON index and Dental and Aesthetic components of IOTN index. *World J Orthod, 11*, 43-48.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters, 27*(8), 861-874.

Firestone, A. R., Beck, F. M., Beglin, F. M., & Vig, K. W. (2002). Validity of the Index of Complexity, Outcome, and Need (ICON) in determining orthodontic treatment need. *The Angle Orthodontist, 72*(1), 15-20.

Fox, N., Daniels, C., & Gilgrass, T. (2002). A comparison of the index of complexity outcome and need (ICON) with the peer assessment rating (PAR) and the index of orthodontic treatment need (IOTN). *British dental journal, 193*(4), 225-230.

Ghafari, J., Locke, S. A., & Bentley, J. M. (1989). Longitudinal evaluation of the treatment priority index (TPI). *American Journal of Orthodontics and Dentofacial Orthopedics, 96*(5), 382-389.

Ghafari, J. G. (2004). Therapeutic and developmental maxillary orthopedics: evaluation of effects and limitations. *Biological Mechanisms of Tooth Eruption, Resorption, and Replacement by Implants, The Harvard Society for the Advancement of Orthodontics, Boston*, 167-181.

Ghafari, J. G. (2016). *"Medically necessary" orthodontic care: Challenges and applications.* Paper presented at the Seminars in Orthodontics.

Graber, T. M. (1972). *Orthodontics: Principles & Practice*: WB Saunders Company.

Grainger, R. (1967). Orthodontic Treatment Priority Index Public Health Service Publication No 1000, Series 2, No. 25. *Government Printing Office, Washington DC, USA*.

Gravely, J., & Johnson, D. (1974). Angle's classification of malocclusion: an assessment of reliability. *British journal of orthodontics, 1*(3), 79-86.

Grippaudo, C., Paolantonio, E. G., La Torre, G., Gualano, M. R., Oliva, B., & Deli, R. (2008). Comparing orthodontic treatment need indexes. *Italian Journal of Public Health, 5*(3).

Han, U. K., Vig, K. W., Weintraub, J. A., Vig, P. S., & Kowalski, C. J. (1991). Consistency of orthodontic treatment decisions relative to diagnostic records. *American Journal of Orthodontics and Dentofacial Orthopedics, 100*(3), 212-219.

Heath, E. M., English, J. D., Johnson, C. D., Swearingen, E. B., & Akyalcin, S. (2017). Perceptions of orthodontic case complexity among orthodontists, general practitioners, orthodontic residents, and dental students. *American Journal of Orthodontics and Dentofacial Orthopedics, 151*(2), 335-341.

Holdaway, R. A. (1983). A soft-tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. *American journal of orthodontics, 84*(1), 1-28.

Holman, J. K., Hans, M. G., Nelson, S., & Powers, M. P. (1998). An assessment of extraction versus nonextraction orthodontic treatment using the peer assessment rating (PAR) index. *The Angle Orthodontist, 68*(6), 527-534.

Houston, W., Stephens, C., & Tulley, W. (1992). A Textbook of Orthodontics. 2nd edn. Wright. *Oxford, 18*, 350-352.

Johansson, A. M., & Follin, M. E. (2009). Evaluation of the dental health component, of the index of orthodontic treatment need, by Swedish orthodontists. *The European Journal of Orthodontics, 31*(2), 184-188.

Johnson, M., Harkness, M., Crowther, P., & Herbison, P. (2000). A comparison of two methods of assessing orthodontic treatment need in the mixed dentition: DAI and IOTN. *Australian orthodontic journal, 16*(2), 82-87.

Kiekens, R. M., Maltha, J. C., van't Hof, M. A., & Kuijpers-Jagtman, A. M. (2005). A measuring system for facial aesthetics in Caucasian adolescents: reproducibility and validity. *The European Journal of Orthodontics, 27*(6), 579-584.

Koochek, A., Shue-Te Yeh, M., Rolfe, B., & Richmond, S. (2001). General Practice: The relationship between Index of Complexity, Outcome and Need, and patients' perceptions of malocclusion: a study in general dental practice. *British dental journal, 191*(6), 325-329.

Liao, Z.-Y., Jian, F., Long, H., Lu, Y., Wang, Y., Yang, Z., . . . Ye, N.-S. (2012). Validity assessment and determination of the cutoff value for the Index of Complexity, Outcome and Need among 12–13 year-olds in Southern Chinese. *International journal of oral science, 4*(2), 88-93.

Linder-Aronson, S. (1973). Orthodontics in the Swedish Public Dental Health Service. *Transactions. European Orthodontic Society*, 233-240.

Little, R. M. (1975). The irregularity index: a quantitative score of mandibular anterior alignment. *American journal of orthodontics, 68*(5), 554-563.

Llewellyn, S. K., Hamdan, A. M., & Rock, W. P. (2007). An index of orthodontic treatment complexity. *The European Journal of Orthodontics, 29*(2), 186-192.

Louwerse, T., Aartman, I., Kramer, G., & Prahl-Andersen, B. (2006). The reliability and validity of the Index of Complexity, Outcome and Need for determining treatment need in Dutch orthodontic practice. *The European Journal of Orthodontics, 28*(1), 58-64.

Massler, M., & Frankel, J. M. (1951). Prevalence of malocclusion in children aged 14 to 18 years. *American journal of orthodontics, 37*(10), 751-768.

McGorray, S. P., Wheeler, T. T., Keeling, S. D., Yurkiewicz, L., Taylor, M. G., & King, G. J. (1999). Evaluation of orthodontists' perception of treatment need and the peer assessment rating (PAR) index. *The Angle Orthodontist, 69*(4), 325-333.

Metz, C. E. (1978). *Basic principles of ROC analysis.* Paper presented at the Seminars in nuclear medicine.

Miguel-Neto, A., & Mucha, J. (2000). Malocclusion Classification-A Newly Proposal System for Malocclusion Classification. *Ortodontia Gaúcha, 4*(1), 41-58.

Nelson, S. J. (2014). *Wheeler's dental anatomy, physiology and occlusion*: Elsevier Health Sciences.

Onyeaso, C. O., & Begole, E. A. (2007). Relationship between index of complexity, outcome and need, dental aesthetic index, peer assessment rating index, and American Board of Orthodontics objective grading system. *American Journal of Orthodontics and Dentofacial Orthopedics, 131*(2), 248-252.

Otuyemi, O., & Jones, S. (1995). Methods of assessing and grading malocclusion: a review. *Australian orthodontic journal, 14*(1), 21.

Ovsenik, M., & Primožič, J. (2007). Evaluation of 3 occlusal indexes: Eismann index, Eismann-Farčnik index, and index of orthodontic treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics, 131*(4), 496-503.

Pae, E.-K., McKenna, G. A., Sheehan, T. J., Garcia, R., Kuhlberg, A., & Nanda, R. (2001). Role of lateral cephalograms in assessing severity and difficulty of orthodontic cases. *American Journal of Orthodontics and Dentofacial Orthopedics, 120*(3), 254-262.

Poulton, D. R., & Aaronson, S. A. (1961). The relationship between occlusion and periodontal status. *American journal of orthodontics, 47*(9), 690-699.

Proffit, W. R., Fields, H. W., & Sarver, D. M. (2014). *Contemporary Orthodontics-E-Book*: Elsevier Health Sciences.

Richmond, S., Daniels, C., Fox, N., & Wright, J. (1997). The professional perception of orthodontic treatment complexity. *British dental journal, 183*(10), 371-375; discussion 375-377.

Richmond, S., & Daniels, C. P. (1998a). International comparisons of professional assessments in orthodontics: part 1—treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics, 113*(2), 180-185.

Richmond, S., & Daniels, C. P. (1998b). International comparisons of professional assessments in orthodontics: part 2—treatment outcome. *American Journal of Orthodontics and Dentofacial Orthopedics, 113*(3), 324-328.

Richmond, S., Shaw, W., O'brien, K., Buchanan, I., Jones, R., Stephens, C., . . . Andrews, M. (1992). The development of the PAR Index (Peer Assessment Rating): reliability and validity. *The European Journal of Orthodontics, 14*(2), 125-139.

Richmond, S., Shaw, W., O'Brien, K., Buchanan, I., Stephens, C., Andrews, M., & Roberts, C. (1995). The relationship between the index of orthodontic treatment need and consensus opinion of a panel of 74 dentists. *British dental journal, 178*(10), 370-374.

Ricketts, R. M. (1968). Esthetics, environment, and the law of lip relation. *American journal of orthodontics, 54*(4), 272-289.

Rinchuse, D. J., & Rinchuse, D. J. (1989). Ambiguities of Angle’s classification. *The Angle Orthodontist, 59*(4), 295-298.

Ross, B. (2003). Satisfactory occlusal relations for the individual with a craniofacial anomaly. *The International journal of prosthodontics, 16*, 74.

Salzmann, J. (1968). Handicapping malocclusion assessment to establish treatment priority. *American journal of orthodontics, 54*(10), 749-765.

Sampson, W., & Sims, M. (1992). Variability of Occlusal Traits in Tropical Populations: Oral diseases in the Tropics. Oxford: Oxford University Press.

Savastano, N. J., Firestone, A. R., Beck, F. M., & Vig, K. W. (2003). Validation of the complexity and treatment outcome components of the index of complexity, outcome, and need (ICON). *American Journal of Orthodontics and Dentofacial Orthopedics, 124*(3), 244-248.

Shaw, W., Richmond, S., & O'Brien, K. (1995). The use of occlusal indices: a European perspective. *American Journal of Orthodontics and Dentofacial Orthopedics, 107*(1), 1-10.

Stricker, G., Clifford, E., Cohen, L. K., Giddon, D. B., Meskin, L. H., & Evans, C. A. (1979). Psychosocial aspects of craniofacial disfigurement: A “state of the art” assessment conducted by the Craniofacial Anomalies Program Branch, The National Institute of Dental Research. *American journal of orthodontics, 76*(4), 410-422.

Summers, C. J. (1966). *A system for identifying and scoring occlusal disorders: The occlusal index.* University of Michigan Ann Arbor.

Summers, C. J. (1971). The occlusal index: a system for identifying and scoring occlusal disorders. *American journal of orthodontics, 59*(6), 552-567.

Sundareswaran, S., & Ramakrishnan, R. (2016). The Facial Aesthetic index: An additional tool for assessing treatment need. *Journal of orthodontic science, 5*(2), 57.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*(4857), 1285-1293.

Tang, E. L., & Wei, S. H. (1993). Recording and measuring malocclusion: a review of the literature. *American Journal of Orthodontics and Dentofacial Orthopedics, 103*(4), 344-351.

Theis, J. E., Huang, G. J., King, G. J., & Omnell, M. L. (2005). Eligibility for publicly funded orthodontic treatment determined by the handicapping labiolingual deviation index. *American Journal of Orthodontics and Dentofacial Orthopedics, 128*(6), 708-715.

Thilander, B., & Rönning, O. (1995). *Introduction to orthodontics*: Gothia.

Torkan, S., Pakshir, H. R., Fattahi, H. R., Oshagh, M., Danaei, S. M., Salehi, P., & Hedayati, Z. (2015). An Analytical Study on an Orthodontic Index: Index of Complexity, Outcome and Need (ICON). *Journal of Dentistry, 16*(3), 149.

Trulsson, U., Strandmark, M., Mohlin, B., & Berggren, U. (2014). A qualitative study of teenagers' decisions to undergo orthodontic treatment with fixed appliance. *Journal of Orthodontics*.

Türp, J., Greene, C., & Strub, J. (2008). Dental occlusion: a critical reflection on past, present and future concepts. *Journal of oral rehabilitation, 35*(6), 446-453.

Van Kirk, L. E., & Pennell, E. H. (1959). Assessment of malocclusion in population groups. *American journal of orthodontics, 45*(10), 752-758.

WHO, W. H. O. (1966). *An international methodology for epidemiological studies of oral disease. Manual No. 5: Epidemiological studies of periodontal disease. First draft*.

Woda, A., Vigneron, P., & Kay, D. (1979). Nonfunctional and functional occlusal contacts: a review of the literature. *The Journal of prosthetic dentistry*, *42*(3), 335-341.

Younis, J. W., Vig, K. W., Rinchuse, D. J., & Weyant, R. J. (1997). A validation study of three indexes of orthodontic treatment need in the United States. *Community dentistry and oral epidemiology*, *25*(5), 358-362.

Appendix I

ABO DISCREPANCY INDEX INSTRUCTIONS

Updated 11/07/2013

Occlusion for plaster models is determined by placing the separated, properly trimmed study casts (Mx/Mn) on a flat surface and then bringing them together into maximum intercuspation. All measurements must be made from this position. For digital models, measurements will be made from a standard 3D orientation that is described in ABO Digital Model Requirements.

OVERJET: Overjet is a measurement between two antagonistic anterior teeth (lateral or central incisors) comprising the greatest overjet and is measured from the facial surface of the most lingual mandibular tooth to the middle of the incisal edge of the more facially positioned maxillary tooth.

- For ≥0 to <1 mm, score 1 pt (edge-to-edge)
- For ≥1 to ≤3 mm, score 0 pts
- For >3 to ≤5 mm, score 2 pts
- For >5 to ≤7 mm, score 3 pts
- For >7 to ≤9 mm, score 4 pts
- For >9 mm, score 5 pts.
- In addition, if there are anterior teeth with negative overjet (canine to canine in anterior crossbite >0 mm), measure from the facial surface of the maxillary tooth to the middle of the incisal edge of the mandibular tooth.
  - Round any fractional remainder to the next full mm,
  - Then score 1 pt per mm per anterior tooth in crossbite.

OVERBITE: Overbite is a measurement between two antagonistic anterior teeth (lateral or central incisors) comprising the greatest overbite.

- For >0 to ≤3 mm, score 0 pts
- For >3 to ≤5 mm, score 2 pts
- For >5 to ≤7 mm, score 3 pts
- If any of the lower incisors are impinging on the palatal tissues (≤ 0.5m) or there is 100% overbite (a complete vertical overlap of antagonistic incisors), score 5 pts.

ANTERIOR OPEN BITE: For each anterior tooth (canine to canine) in an open bite relationship with an opposing tooth, measure from the incisal edge of the Mx tooth to the incisal edge of the Mn tooth.

- For each anterior tooth in edge-to-edge relationship (0 mm), score 1 pt per tooth.
- For each anterior tooth in open bite (> 0 mm), round any fractional remainder to the next full mm,
- Then add 1 pt per mm per tooth in open bite.
- No points are scored for any anterior tooth that is blocked-out of the arch due to space deficiency or not fully erupted.

LATERAL OPEN BITE: For each maxillary posterior tooth (from the 1st premolar to 2nd molar) in an open bite relationship ≥ 0.5 mm from its opposing tooth, measure cusp to cusp.

- Round any fractional remainder to next full mm
- Then score 2 pts per mm of open bite for each tooth.
- No points are scored for any tooth that is blocked-out of the arch due to space deficiency or not fully erupted.

CROWDING: Measure the most crowded arch (only one arch) from the mesial contact point of the right first molar to the mesial contact point of the left first molar. If there are conditions such as missing, fractured or decayed teeth, then measure crowding consistent with your treatment objectives and be prepared to defend the score at your oral examination.

- For ≥0 to ≤1 mm, score 0 pts
- For >1 to ≤3 mm, score 1 pt
- For >3 to ≤5 mm, score 2 pts
- For >5 to ≤7 mm, score 4 pts
- For >7 mm, score 7 pts.

OCCLUSAL RELATIONSHIP: Models must exhibit the patient's maximum intercuspation. The Angle molar classification is used.

- If the mesiobuccal cusp of the maxillary first molar occludes with the buccal groove of the mandibular first molar or anywhere between the buccal groove and the mesiobuccal or distobuccal cusps (Class I to End On) - Score 0 pts.
- If the mesiobuccal cusp of the maxillary first molar occludes with the mesiobuccal (Class II end-to-end) or distobuccal (Class III end-to-end) cusps of the mandibular first molar – Score 2 pts per side.
- If the relationship is a full Class II or III - Score 4 pts per side.
- If the relationship is beyond Class II or III, measure the additional distance, round any fractional remainder to next full mm – Score 4 pts plus 1 addl. pt per mm per side.

LINGUAL POSTERIOR CROSSBITE: For each maxillary posterior tooth (from the 1st premolar to the 2nd molar) where the maxillary buccal cusp is > 0 mm lingual to the buccal cusp tip of the opposing tooth - Score 1 pt per tooth.

BUCCAL POSTERIOR CROSSBITE: For each maxillary posterior tooth (from the 1st premolar to the 2nd molar) where the maxillary palatal cusp is > 0 mm buccal to the buccal cusp of the opposing tooth - Score 2 pts per tooth.

CEPHALOMETRICS: (See Construction of Mandibular Plane)

- If the ANB angle is ≥ 6° OR ≤ -2 , score 4 pts; then, add 1 pt for each full degree > 6° OR < -2° .
- If the SN-MP angle is between 27° and 37° score 0 pts.
- If the SN-MP angle is ≥ 38°, score 2 pts; then, add 2 pts for each full degree > 38°.
- If the SN-MP angle is ≤ 26°, score 1 pt; then, add 1 pt for each full degree < 26°.
- If the Lower Incisor to MP angle is ≥ 99°, score 1 pt; then, add 1 pt for each full degree greater than 99°.

OTHER: (List number of occurrences and total points.)

- **Supernumerary** teeth – Score 1 pt for each extra tooth.
- **Ankylosis** of permanent teeth – Score 2 pts per tooth.
- **Anomalous morphology** of tooth size & shape (e.g. natural and/or iatrogenic) - Score 2 pts per tooth.
- **Impaction** of teeth (except 3rd molars) – Score 2 pts per tooth.
- **Midline discrepancy** – The midline for each arch equals the mid-point between the Mx central incisors and the Mn central incisors demonstrated by two vertical reference lines. The discrepancy is the difference between the two vertical reference lines measured in the horizontal plane – Score 2 pts for ≥ 3 mm.
- **Missing teeth** (except 3rd molars) -
  - Non-congenital – Score 1 pt per tooth.
  - Congenital – Score 2 pts per tooth.
- **Spacing** –
  - For generalized spacing per arch in which there is ≥ 0.5 mm of space on both sides of any 4 teeth or more - Score 2 pts per arch.
  - For Mx central diastema of ≥ 2 mm - Score 2 pts.
- **Tooth transposition** – Score 2 pts for each event.
- **Skeletal asymmetry** (treated nonsurgically) – Score 3 pts (appropriate diagnostic information recommended)
- **Additional treatment complexities** - Score 2 pts each and identify.

FOR ADDITIONAL VISUAL REFERENCE, SEE "DISCREPANCY INDEX SCORING SYSTEM"

20131107

Appendix II

SCORING SHEET

Case number:

Need for treatment

You are the orthodontist consultant for a private corporation for which a fund has been established to provide orthodontic treatment for personnel, you are to assess these records of personnel and answer the following question:

In your opinion, to what extent does this occlusion need orthodontic treatment?
(Please circle the appropriate number)

None or Minimal Need						Great need
1	2	3	4	5	6	7

Complexity of the treatment

When treating these patients, assume that you have a full range of orthodontic appliances and that the patient is willing to undergo 30 months of orthodontic treatment, and that an ideal result is the intended outcome. Assume that the teeth are healthy and that the patient is at an age to respond favorably to myofunctional or functional therapy:

In your opinion, how complex is the orthodontic treatment needed?
(The ranking of 'impossible' should only be given to the case if you feel that it is untreatable without orthognathic surgery)

Easy						Impossible
1	2	3	4	5	6	7

Estimated treatment time

In your opinion how many months the treatment will take: [ ] months

Outcome of the treatment

In your own clinical judgment decide:

Whether the outcome is acceptable:

Acceptable	*Not Acceptable*

The degree of improvement:

Worse / no improvement						Greatly improved
1	2	3	4	5	6	7

Appendix III

Table 1: Demographics of the sample select:

	N	Mean	Std. Dev.
Chronological age	101	17.2	8
Number of total appointments	98	35.1	15.9
Number of missing appointments	98	2.5	5.0
Gender N (%)			
Males	39 (38.6%)		
Females	62 (61.4%)		

Table 2: Spearman correlation analysis between the pre-treatment scores of the different occlusal indices:

	PAR	ICON	AC	DHC	DI	dDI	cIOTN
PAR	1						
ICON	0.56	1					
AC	0.52	0.91	1				
DHC	0.67	0.58	0.48	1			
DI	0.33	0.42	0.36	0.35	1		
dDI	0.65	0.64	0.5	0.61	0.71	1	
cIOTN	0.65	0.92	0.94	0.74	0.41	0.61	1

Table 3: Sensitivity test comparing the results of the of the summery statistics with and without FR panel member:

INDICES	NEED (AUC)		COMPLEXITY (r)		IMPROVEMENT (r)	
	without	With FR	without	With FR	without	With FR
AC	0.83	0.84	0.45	0.43	-0.48	-0.48
DHC	0.83	0.85	0.68	0.72	-0.51	-0.48
cIOTN	0.88	0.88	0.61	0.62	-0.57	-0.54
PAR	0.82	0.82	0.44	0.47	-0.56	-0.53
ICON	0.83	0.83	0.50	0.48	-0.53	-0.53
DI	0.67	0.69	0.42	0.43	-0.60	-0.60
dDI	0.80	0.81	0.60	0.61	-0.57	-0.56

Table 4: Demographics and cephalometric measurements by panel assessment of definitive need for orthodontic treatment:

	No need N=31	**Need N=70**	**P-value**
	Mean (SD)	Mean (SD)	
Demographics			
Age	18.1 (7.1)	16.7 (8.4)	0.4111
Gender N (%)			
Males	9 (29.0%)	30 (42.9%)	0.1880
Females	22 (71.0%)	40 (57.1%)	
Cephalometric Measurements			
SNA angle	81.0 (4.0)	80.7 (4.1)	0.7032
SNB angle	76.9 (4.2)	75.7 (3.9)	0.1746
ANB angle	4.1 (1.7)	5.0 (2.4)	0.0949
PP-MP angle	27.0 (5.8)	27.8 (5.4)	0.4778
MP-SN angle	38.1 (6.1)	38.0 (5.8)	0.9092
I-SN angle	100.0 (7.8)	101.7 (8.4)	0.1875
I-NA angle	19.0 (7.0)	21.7 (7.5)	0.0930
i-MP angle	93.7 (6.1)	94.1 (6.7)	0.7720
i-NB angle	28.7 (4.7)	27.8 (6.3)	0.4750
Nose Tip / HL	5.7 (4.0)	3.7 (4.8)	0.0404*
Subnasale / HL	4.9 (2.2)	5.5 (2.7)	0.2910
Lower Lip / HL	1.3 (1.2)	1.3 (1.7)	0.8777
Mental sulcus / HL	4.1 (1.8)	4.1 (1.7)	0.9230
Upper lip / EL	-3.2 (2.2)	-2.1 (2.7)	0.0369*
Lower Lip /EL	-0.8 (2.2)	0 (2.8)	0.1782

I Maxillary incisors, i Mandibular incisors, HL Holdaway line, EL Esthetic line
*P-value < 0.05

Table 5: Summary statistics of panel assessment scores in the Class II population:

N = 55	**PA**	**SB**	**FR**	**RS**	**gold standard**
Need					
Mean (SD)	4.4 (1.4)	4.2 (0.8)	5.7 (1.2)	4.5 (1.6)	4.7 (1.1)
ITP					
Value indicated	3	3	5	6	4.25*
Complexity					
Mean (SD)	3.7 (1.1)	3.7 (1.0)	4.9 (1.2)	3.5 (1.2)	3.9 (0.9)
Improvement					
Mean (SD)	4.5 (1.3)	3.6 (1.0)	5.5 (1.2)	5.5 (1.2)	4.8 (1.0)
Outcome					
Not acceptable	3 (5.45%)	4 (7.3%)	2 (3.6%)	1 (1.8%)	5 (5.45%)
Acceptable	52 (94.6%)	51 (92.7%)	53 (96.4%)	54 (98.2%)	50 (90.9%)