Accuracy of Digital American Board of Orthodontics Discrepancy Index

BY

KRISTINA DRAGSTREM
BS, University of Illinois at Urbana-Champaign, 2006
DDS, University of Illinois at Chicago, 2011

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Oral Sciences
in the Graduate College of the
University of Illinois at Chicago, 2014

Chicago, Illinois

Committee members:

Budi Kusnoto, Chair and Advisor
Maria Therese S. Galang-Boquiren
Ales Obrez, Restorative Dentistry
Maria Grace Costa Viana

This thesis is dedicated to my husband, Ryan for his unconditional love and support.

ACKNOWLEDGEMENTS

I would like to thank my thesis committee Dr. Budi Kusnoto, Dr. Maria Therese Galang - Boquiren, Dr. Ales Obrez, and Maria Grace Costa Viana for their patience, support, and advice. I also would like to thank Dr. John Grubb for his support and knowledge of the American Board of Orthodontics Discrepancy Index system.

KGD

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D	Three-dimensional
ABO	American Board of Orthodontics
DI	Discrepancy Index
OB	Overbite
OGS	Occlusal Grading System
OJ	Overjet
mm	Millimeters

SUMMARY

The objective of the this study was to compare accuracy and reliability of the American Board of Orthodontics Discrepancy Index obtained manually on plaster casts (the gold standard) with scores obtained using their digital models in OrthoCad software and Motion View's Ortho Insight 3D software. Plaster models were obtained from 45 previous patients with varying degrees of malocclusion. Measurements were completed manually with a periodontal probe as necessary. The same casts were scanned and analyzed using Ortho Insight 3D (Motion View Software, Hixson, TN). Alginate impressions were made of the original plaster casts and sent to OrthoCad (Cadent, Carlstadt, NJ) for digital model fabrication and analysis. Total DI score and its target disorders were computed manually on plaster casts and digitally using the respective software. Intra-rater and inter-rater reliability was assessed for 15 subjects using Spearman's Rho correlation test. Accuracy DI scores and its target disorders were assessed for all 45 subjects using Wilcoxon signed ranks test to determine differences in digital calculation compared to manual calculation. Intra-rater and inter-rater reliability was high for total DI score and most target disorders (r>0.8). No significant difference was found between total DI score when measured with OrthoCad compared to manual calculation. Total DI score calculated by Ortho Insight 3D was found to be significantly larger than manual calculation by 2.71 points.

1. INTRODUCTION

1.1 Background

Digital technology has become an integral component of orthodontic offices. Its use meets the demand of multiple-doctor practices, multiple practice locations, and increase in patient base, and allows efficient and convenient storage, retrieval, and sharing of information (Redmond, 2001). Digital photography and radiography are replacing analogue systems; computer based charting and patient management software allow electronic patient record keeping (Quimby et al., 2004). For dental model analysis, use of plaster models has been the gold standard however they are subject to loss, fracture, and degradation, are difficult to retrieve and share, and require physical storage space. As a means of overcoming the above-mentioned problems, three-dimensional study models have increased in popularity. Three-dimensional imaging has undergone significant advances in recent years leading to the possibility of the “virtual orthodontic patient”, where bone, soft tissue, and teeth can be recreated in three dimensions (Fleming et al., 2011).

In addition to being an essential component of initial diagnosis and treatment planning, objective model analysis is necessary for demonstrating case complexity for those pursuing certification by The American Board of Orthodontics (Stevens et al., 2006). The ABO devised the Discrepancy Index (DI) to provide “an objective method to describe the complexity of the treatment for a patient based on measurements taken from standard [pre-treatment] orthodontic records, including [dental models] and cephalometric and panoramic radiographs”

(Cangialosi et al., 2004). In a recent survey looking at the future plans of orthodontic residents, 92.75% planned on using a digital imaging program and 81.16% planned on becoming certified by The American Board of Orthodontics (Noble et al., 2009). Thus, a digital analysis that is shown to be accurate and reliable will facilitate ease of demonstrating initial case complexity for candidates applying or reapplying for ABO certification.

1.2 Specific Aims

This study aims to determine accuracy of digital discrepancy index total score and its constituent parameters obtained by OrthoCad and Ortho Insight 3D software compared to manual calculations completed on plaster models.

1.3 Null Hypothesis

The null hypotheses were:

- There is no statistically significant difference in total DI scores obtained digitally through OrthoCad or Ortho Insight 3D when compared to total scores obtained through manual measurements made on a plaster model.
- There is no statistically significant difference in scores of overjet, overbite, open bite, crowding, occlusal relationship, crossbite presence, midline discrepancy presence, spacing, or maxillary diastema presence obtained through OrthoCad or Ortho Insight 3D when compared to scores obtained through manual measurements made on a plaster model.

2. RELATED LITERATURE

2.1 Dental Model Analysis

Successful treatment planning in orthodontics requires precise diagnostic information and an accurate diagnosis. Dental models are an important part of the initial diagnostic record. They provide a three-dimensional view of the patient's occlusion, which enable the clinician to evaluate the malocclusion in more detail than by clinical examination alone (Quimby et al., 2004). Study models are more amendable to routine measurements than are intraoral measurements. Typically tooth size and shape, crowding or spacing, overjet, overbite, and inter-arch relationships are determined by hand on plaster model (Martensson and Ryden, 1992; Santoro et al., 2003).

2.2 Occlusal Indices

One type of model analysis is the occlusal index. There are many types of occlusal indices but in general occlusal traits are scored and the results provide information that objectively describes the malocclusion. Occlusal indices have many potential benefits for an orthodontic setting. First, indices may be used for resource allocation and planning such as estimating the level of need in a community and allocating funds to those cases which demonstrate the greatest need. They also may be used for monitoring and promoting standards both on an individual level and on a larger level such as the ABO certification process. In a clinical practice, the results of an occlusal index can assist in identification and referral of potential orthodontic patients. Orthodontists may also use indices to determine level of complexity for a particular case and subsequently an appropriate fee and estimated treatment duration for that case (Shaw et al.,1995). The “ideal” occlusal index should possess the following properties (DeGuzman et al., 1995; Otuyemi and Jones, 1995; Richmond et al., 1992; Summers, 1971):

1) Reliability: ability to measure consistently when applied by different examiners and at different times
2) Validity: the index should measure what it purports to measure
3) Simplicity: the index should allow rapid application by trained examiners
4) Amenable to modification
5) Yields quantitative data

Numerous indices have been proposed over the years in attempts to quantify orthodontic need or severity of malocclusion for a specific purpose. Generally, the more complex a case is the more likely the case indicates treatment need. However, no one index is universally accepted and each has limitations. The Occlusal Index was developed by Summers in attempt to measure occlusion in an objective manner however scoring was thought to be time consuming and cumbersome (Summers, 1971). The Index of Treatment Need (IOTN), including the Dental Health Component (DHC) and Aesthetic Component (AC), was developed to determine orthodontic treatment priority based on rankings of occlusal traits but the aesthetic component is subjective (Brook and Shaw, 1989). The ABO developed its own method of determining case complexity with of purpose of selecting cases for the Phase III Examination called the Discrepancy Index (DI) (Cangialosi et al., 2004). The DI is not affected by age, sex, or time when the patients are evaluated as opposed to indices developed before it (Schafer et al., 2011).

Other indices have been proposed for evaluating treatment outcome. The Index of Complexity, Outcome, and Need (ICON) was developed for the assessment of treatment need, complexity, treatment improvement, and outcome (Daniels and Richmond, 2000). Results showed high agreement of index of need with decisions of experts (Firestone et al., 2002), moderate agreement for complexity and outcome, and only fair for improvement (Savastano et al., 2003). The Peer Assessment Rating (PAR) Index measures clinical outcome as the change in total score between pre-treatment and post-treatment models (Richmond et al., 1992) however it has been criticized for its heavy weight given to a single component (Fox, 1993). Also, claims that PAR's measuring system was not precise enough to discriminate between minor inadequacies of tooth position in ABO case reports lead to the formation of an ABO committee designated to field testing precise methods of objectively evaluating post-treatment casts and radiographs (Casko et al., 1998). In 1999, ABO developed a grading system for model analysis as a means of evaluating final occlusion called the Occlusal Grading System (OGS) (Casko et al., 1998).

2.3 <u>The Discrepancy Index</u>

The Discrepancy Index (DI) was developed by the ABO to provide an objective evaluation of case complexity that may lead to a better understanding of difficulty. The DI total score is determined from observations and measurements taken from standard orthodontic pretreatment records. It attempts to summarize the clinical features of a patient's condition with a quantifiable, objective list of target disorders that represent most conditions that orthodontists treat. The target disorders that make up the index were also chosen because all could be related to deviations from generally accepted norms and be calculated relatively quickly and simply. The DI was initially developed in 1998 by ABO present and past directors and then subjected to field-testing, data analysis, and modification for a period of five years. In general, the greater the DI score, the relatively more complex a case is determined to be. The target disorder elements included in the DI are measurements of overjet, overbite, anterior open bite, lateral open bite, crowding, occlusion, lingual posterior crossbite, buccal posterior crossbite, and "other" anomalies such as the presence of midline deviation, maxillary diastema, generalized spacing, missing teeth, tooth anomalies, and tooth transposition. The target disorders also include cephalometric ANB angle, IMPA, and SN-GoGn angle (Cangialosi et al., 2004).

Despite its initial purpose for determining case complexity when selecting cases in preparation for ABO Phase III examination, the use of DI has also been evaluated with respect to predicting treatment duration and clinical outcomes. One study found significantly longer mean treatment duration for patients with a DI score greater than 20 compared to DI scores less than 20 however no significant difference in final treatment outcome measured by ABO OGS scoring was detected among the classes defined by the initial DI score (Vu et al., 2008). The authors concluded a 1-unit increase in DI score increased treatment duration by 0.1 month (Vu et al., 2008). Similarly, a significant correlation was found between DI and treatment duration such that for each point increase in DI score, an average increase in treatment duration of 11 days resulted (Parrish et al., 2011). The dental components demonstrating the strongest influence in treatment duration included presence of tooth transposition, crowding, overjet, overbite, occlusion, and lateral open bite (Parrish et al., 2011). Schafer et al. found that patients with elevated DI scores were less likely to be treated by only 1 resident in a 24-month program and thus a positive correlation between treatment duration and DI score was confirmed (Schafer et al., 2011). Pulfer et al. demonstrated a weak positive association between ABO OGS scores and DI scores as well as a weak positive association between Indiana's Comprehensive Clinical Assessment (CCA) scores and DI scores indicating DI as a modest factor in determining treatment outcomes (Pulfer et al., 2009). Patients who receive accurate information about predicted treatment duration are found to be more satisfied with their treatment and have more realistic expectations (Mavreas and Athanasiou, 2008). Thus use of DI scores to more accurately predict treatment length may be advocated.

2.4 Digital Models

Traditionally, use of plaster casts has been the gold standard in completing a model analysis. Digital models were introduced commercially in 1999 by OrthoCad (Cadent, Carlstadt, NJ, USA). The results from a recent survey conducted by the Journal of Clinical Orthodontics demonstrate a significant increase in the use of digital models for pre-treatment diagnosis and treatment from 6.6% in 2002 to 18.0% in 2008 (Keim et al., 2008). Today many companies offer digital models as a replacement for plaster models. Digital models may be obtained indirectly through a dental impression that is poured in plaster or stone followed by destructive or non-destructive image processing. The destructive imaging process removes a thin layer of material, alternating with image capture to produce a stack of images rendered in 3D. Non-destructive imaging uses a laser-based system with multi-axis robot to obtain several perspectives of a plaster model that are combined to render the 3D model. Direct production of 3D models does not require an impression and includes several non-contact optical technologies to reconstruct a virtual model by post-processing single images acquired from a single perspective (Mah and Hatcher, 2003). OrthoCad uses destructive scanning with multiple scans of a plaster equivalent in thin slices (Fleming et al., 2011). Storage of the digital models also varies from company to company. OrthoCad results in a relatively large file of 3000 kilobytes and in a proprietary format (Stevens et al., 2006). Motion View Software (Hixson, TN, USA) introduced Ortho Insight 3D Scanner and Software that utilizes a robotic scanner to produce 3D renderings of impressions or plaster models. The advantage of this software is the flexible, open format of the file produced.

2.5 Accuracy and Reliability of Digital Model Analysis

Several studies in the literature have verified the accuracy of linear measurements on 3D digital models with different software and found divergent results. A recent systematic review evaluated studies assessing validity of seven digital model systems (OrthoCad, emodel, C3D-builder, ConoProbe, Easy3D Scan, Digimodels, Cecile) (Fleming et al., 2011). Authors concluded that in most studies, mean discrepancy between plaster and digital model measurements was low. They also noted that based on the findings of the systematic review, digital models may be an alternative to plaster models however the “available evidence is of variable quality” (Fleming et al., 2011).

Measurements completed on OrthoCad digital models have been compared to measurements made on plaster models in many studies. Measurements of tooth size have been found to be similar or slightly smaller in OrthoCad (Bootvong et al., 2010; Leifert et al., 2009; Quimby et al., 2004; Santoro et al., 2003; Tomassetti et al., 2001; Zilberman et al., 2003). Arch width comparisons between OrthoCad digital models and plaster models also demonstrated no significant difference (Bootvong et al., 2010; Quimby et al., 2004; Zilberman et al., 2003). Overjet measurements between OrthoCad digital models and plaster models were not significantly different in some studies however Quimby et al. found a significantly smaller overjet measurement when obtained with OrthoCad (Bootvong et al., 2010; Quimby et al., 2004; Santoro et al., 2003). Overbite measurements on OrthoCad digital models were significantly smaller than their plaster model counterparts (Santoro et al., 2003; Quimby et al., 2004). Space available or arch length to be used in estimating crowding, demonstrated to be significantly different between OrthoCad models and plaster models with differences ranging from 0.4mm to 2.88mm (Leifert et al., 2009; Quimby et al., 2004).

With respect to occlusal indices, agreement between manual and digital measurements was high for Peer Assessment Rating (PAR) and Index of Complexity, Outcome, and Need (ICON) (Mayers et al., 2005; Stevens et al., 2006; Veenema et al., 2009). Three studies have evaluated the application of the ABO Objective Grading System (OGS) to digital models. Hildebrand et al.

found significant differences when comparing total ABO scores from plaster and digital models with scores from digital models exceeding scores of plaster casts (Hildebrand et al., 2008). The authors contributed the difference in scores to significant differences in three components: alignment, occlusal contact, and overjet (Hildebrand et al., 2008). Another study also found significant differences in total scores, occlusal contacts, and occlusal relationship (Okunami et al., 2007). Costalos et al. demonstrated no significant difference in total scores between digital and plaster casts however they did find significant differences in alignment and buccolingual inclination (Costalos et al., 2005). All three studies concluded a digital ABO OGS analysis could not substitute a manual calculation (Costalos et al., 2005; Hildebrand et al., 2008; Okunami et al., 2007).

Recently, a study evaluated the accuracy of Motion View Ortho Insight 3D software for assessing tooth width, arch width, and arch length (Kim et al., 2013). Authors found significant differences in mesio-distal widths of maxillary molars, mandibular premolars and molars, arch widths of maxillary premolars, and arch lengths between Ortho Insight 3D and plaster model measurements however 90% of mean differences were less than 0.20mm.

To date no literature exists on the accuracy or reliability of digital Discrepancy Index calculations when applied to pretreatment digital models. The potential advantages of digital models for the quantification of initial orthodontic problems and case complexity would be negated if the accuracy digital discrepancy index scores were not comparable to scores and measurements obtained on plaster models, the current gold standard in orthodontic practice.

3. MATERIALS AND METHODS

3.1 Study Design

This study evaluated manual and digital DI scores obtained from plaster models of varying degrees of complexity. OrthoCad was chosen based on popularity in literature while Ortho Insight 3D was chosen due to its unique, almost fully automated Discrepancy Index module. Cephalometric parameters were not included in total DI scores for the purpose of this study. Calculations of DI score manually were obtained by following instructions set forth by the ABO clinical examination guide (American Board of Orthodontics, 2013). Further clarification is available in the article by Cangialosi et al. (Cangialosi et al., 2004). Total DI scores and scores for its target disorders were compared between digital calculation by OrthoCad, digital calculation by Ortho Insight 3D, and the current gold standard, manual measurement using plaster casts.

3.2 Institutional Review Board Approval

A “Determination of Whether an Activity Represents Human Subjects Research” application was submitted to the UIC Office for the Protection of Research Subjects (UIC OPRS). The UIC OPRS determined on January 28, 2013 that the present study (protocol #2013-0090) did not meet the definition of human subject research and therefore permission to conduct research without further submission to the IRB was granted.

3.3 Methodology

The study assessed 45 pre-treatment plaster study models with 15 models in each category set forth by the case report category specifications of the ABO clinical examination guide (American Board of Orthodontics, 2013): 1) demonstrating a DI score less than 10 (would not qualify for use in examination); 2) DI score greater than 10 but less than 20; 3) DI score greater than 20.

Study models were obtained from the University of Illinois Department of Orthodontics storage. Models contained no personal identification of the patients they represented. Three hundred fifty-six models were assessed for selection criteria and subsequently scored by the primary investigator until 15 sets of study models fit in each category set forth by the DI score while meeting inclusion and exclusion criteria. Inclusion criteria were pretreatment models trimmed so that when placed on their heels, the teeth occlude in centric occlusion. Exclusion criteria stated no supernumeraries, no missing teeth (except third molars), no deciduous teeth, no appliances, and no fractures, voids, or positive bubbles of plaster.

Alginate impressions (Kromopan 100, Kromopan USA, Des Plaines, Ill) and a wax bite registration were taken of each set of plaster models using plastic trays, wrapped in moist towels, sealed in plastic bags, and stored in individual boxes at 37 degrees Fahrenheit overnight until they were sent the next day to OrthoCad for digitization. OrthoCad returned the digital 3D models 1 week later to be viewed and manipulated using the proprietary OrthoCad software Version 3.4 (Cadent, Carlstadt, NJ). Models were visually verified for accuracy of occlusion and if necessary were manipulated using the jaw alignment function. OrthoCad's ABO DI module computes points and total score based on user detected interproximal contacts, arch form, overbite and overjet measurements, occlusal relationships, buccal and lingual crossbite presence, diastema presence, and spacing presence.

The same plaster models were then scanned by the primary investigator using the Ortho Insight 3D scanner at slowest speed (for greatest detail capture) and uploaded to the Ortho Insight 3D software Version 4.0.6 (Motion View Software, Hixson, TN). Because the plaster models are already trimmed so that when placed on their heels, the teeth occlude in maximum intercuspation, the software will automatically detect occlusion when the "auto-align" function is chosen. Once the examiner selects all the teeth to be analyzed, Ortho Insight 3D software will automatically detect occlusal landmarks, interproximal contacts, and long axis of the teeth for DI calculation and appropriately apply points to provide a total DI score. However per Motion View's recommendation, the examiner reviewed automatically detected occlusal landmarks, interproximal contacts, and long axes and manually manipulated each point for improved accuracy. Table I defines the methods used to calculate discrepancy index scores.

TABLE I

KEY DIFFERENCES IN METHODS OF DISCREPANCY INDEX CALCULATION

Method	**Model**	**Landmark Identification**	**Scoring**
MANUAL	Plaster	Manually detected by user	Manual calculation
ORTHOCAD	Digital (obtained from alginate impression)	Digital points detected by user	Points calculated digitally following required user input with respect to tooth selection, contact points, crossbite presence, and molar relationship
ORTHO INSIGHT 3D	Digital (obtained from scanning plaster model)	Digital points detected by computer and modified by user	No user input

3.3.1 Data Measurements

Table II describes the target disorders evaluated and scored set forth by the ABO clinical examination guide (American Board of Orthodontics, 2013). To evaluate target disorders more thoroughly, the category of anterior open bite was separated into edge-to-edge open bite and open bite (greater than 0mm), the occlusal relationship category was separated into right and left occlusal relationship, and the other category was separated into midline discrepancy, spacing, and maxillary diastema. All raw data measurements were entered into Microsoft Excel for Mac 2011 (Version 14.2.3, Microsoft Corporation, Redmond, WA). In addition to point value awarded by target disorders, overjet, overbite, and crowding were recorded in millimeters rounded to the nearest 0.5mm for manual measurements and to the nearest 0.1mm for OrthoCad and Ortho Insight 3D. For all other target disorders point values were collected only.

TABLE II

DEFINITIONS AND SCORING OF DISCREPANCY INDEX TARGET DISORDERS

Target Disorder	**Definition**	**Point Allocation**
OVERJET	Measurement between two teeth (laterals or central incisors) demonstrating greatest overjet from the labial surface of most lingual mandibular tooth to middle of incisal edge of the more labial positioned maxillary tooth	≥0 to <1mm = 1 point ≥1 to ≤3mm = 0 points >3 to ≤5mm = 2 points >5 to ≤7mm = 3 points >7 to ≤9mm = 4 points >9mm = 5 points
NEGATIVE OVERJET	If any anterior teeth (canine to canine) in anterior crossbite greater than 0mm, measure from facial surface of maxillary tooth to middle of incisal edge of mandibular tooth (Round to next full mm)	1 point per mm per tooth in crossbite
OVERBITE	Measurement between two antagonistic anterior teeth (laterals or centrals) comprising greatest vertical overbite from incisal edge to incisal edge	>0 to ≤3mm = 0 points >3 to ≤5mm = 2 points >5 to ≤7mm = 3 points If lower incisors impinge palatal tissue OR 100% overbite = 5 points
ANTERIOR OPEN BITE	Measurement between anterior teeth (canine to canine) in open bite relationship from incisal edge to incisal edge (round to next full mm) * Points are not scored for teeth blocked out of the arch or not fully erupted	>0mm = 1 point per mm per tooth
EDGE-TO-EDGE OPEN BITE	Measurement between anterior teeth (canine to canine) in an incisal edge to incisal edge relationship with zero overbite	1 point per tooth
LATERAL OPEN BITE	Each maxillary posterior tooth (1st premolar to 2nd molar) in an open bite relationship ≥0.5mm with its opposing tooth (round to next full mm)	2 points per mm per tooth
CROWDING	Measure of the most crowded arch from mesial contact of right first molar to mesial contact of left first molar irrespective of arch position	≥0 to ≤1mm = 0 points >1 to ≤3mm = 1 point <3 to ≤5mm = 2 points >5 to <7mm = 4 points >7mm = 7 points

TABLE II (continued)

DEFINITIONS AND SCORING OF DISCREPANCY INDEX TARGET DISORDERS

Target Disorder	**Definition**	**Point Allocation**
OCCLUSAL RELATIONSHIP	Angle molar classification is applied to right and left molars in maximum intercuspation and classification is based on the mesio-buccal cusp of maxillary first molar with respect to the lower first molar's buccal cusp tips and interproximal contacts	Class I to End on = 0 points End on Class II /III = 2 points per side Class II/III Full Step = 4 points per side Mm beyond Full Step Class II/III = 4 points plus 1 point per mm per side
LINGUAL POSTERIOR CROSSBITE	For each maxillary posterior tooth (1^{st} premolar to 2^{nd} molar) where the maxillary buccal cusps are >0mm lingual to buccal cusps of opposing tooth	1 point per tooth in lingual crossbite
BUCCAL POSTERIOR CROSSBITE	For each maxillary posterior tooth (1^{st} premolar to 2^{nd} molar) where maxillary palatal cusps are >0mm buccal to buccal cusps of opposing tooth	2 points per tooth in buccal crossbite
MIDLINE DISCREPANCY	The mid-point between maxillary incisors and mandibular incisors demonstrated by vertical reference lines where the difference between them is ≥3mm	2 points
SPACING	Generalized spacing as defined where there is ≥0.5mm of space on both sides of any 4 or more teeth	2 points per arch
MAXILLARY DIASTEMA	Space existing between mesial of upper incisors ≥2mm	2 points

Manual measurements were completed on the original plaster casts with a periodontal probe marked in millimeters. Target disorders were measured for OrthoCad digital models using the ABO DI module. The 3D models were manipulated through zoom or rotation at the examiners discretion. When determining tooth widths for measurements of crowding, the mesio-distal, occlusal-gingival, and buccal-lingual planes were all verified. Teeth for evaluation of overjet, negative overjet, overbite, and open bite, are selected by the examiner and visually determined for buccal and lingual crossbite presence and molar relationship. DI scores were determined for digital models in the Ortho Insight 3D software using the automated DI score module. Tooth width for the purpose of crowding measurement was automatically detected; however per company recommendations, the primary investigator verified tooth widths and occlusal landmarks in three planes and adjusted contact points accordingly. Other parameters for calculation of the DI score were automatically detected by the Ortho Insight 3D software and not manually adjusted.

3.3.2 Reliability

Fifteen casts were selected at random (5 chosen from subjects with DI <10, 5 chosen from subjects with DI ≥10 and <20, and 5 chosen from subjects with DI≥20) for tests of reliability. The primary investigator measured each set of plaster casts twice for all target disorders and summed the points awarded for total DI scores. Measurements were separated by one week apart. The same 15 casts were sent to a second examiner and two scoring sheets were provided for each set of casts with instructions to grade each set twice, one week apart. Digital models corresponding to the same 15 subjects were measured and scored twice by the primary investigator using the OrthoCad software and Ortho Insight3D software. Inter-rater reliability was achieved by using a second examiner with ABO DI calculation experience for manual measurements and digital measurements. OrthoCad files were sent to the second examiner with the ABO clinical examination guide for clarification of target disorder definitions. Due to the semi-automated nature of Ortho Insight 3D, inter-rater reliability was not assessed.

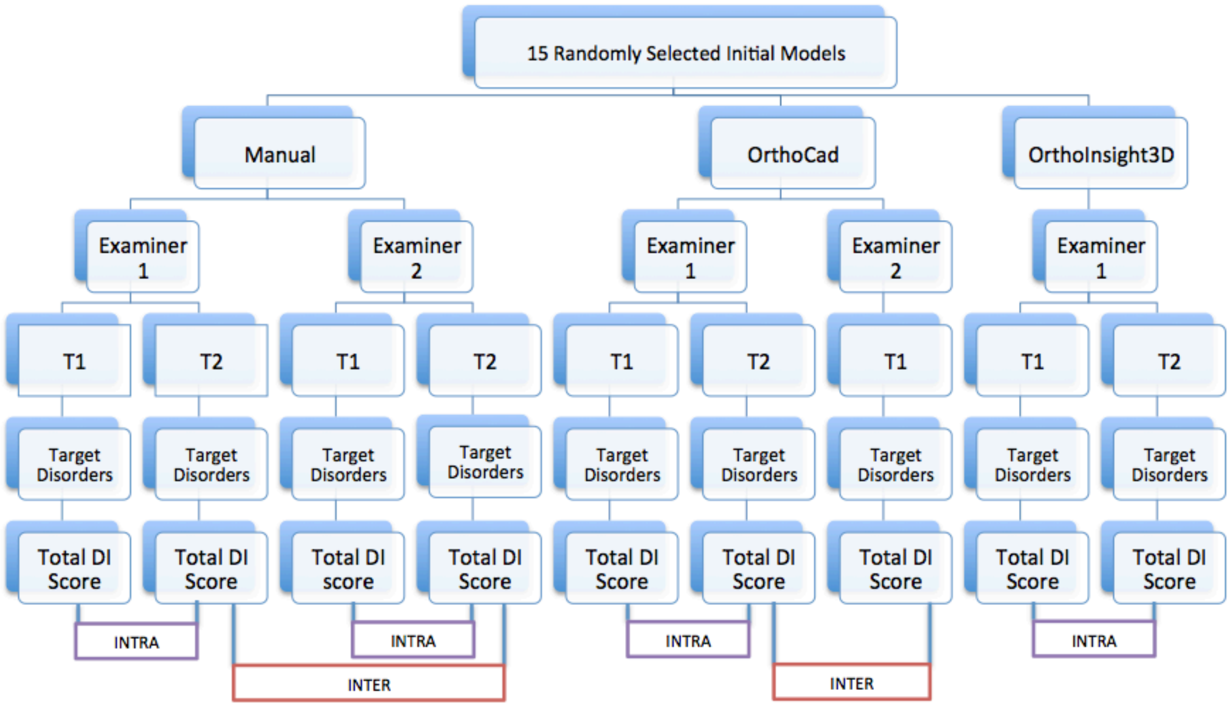

Figure 1. Intra-rater and inter-rater reliability assessed with Spearman's Rho correlation

3.3.3 Accuracy

The primary investigator measured 45 sets of initial study models manually, digitally with OrthoCad software, and digitally with Ortho Insight 3D software according to the same guidelines as described in the initial reliability testing. Each target disorder was evaluated and recorded and the total DI scores were computed by each method.

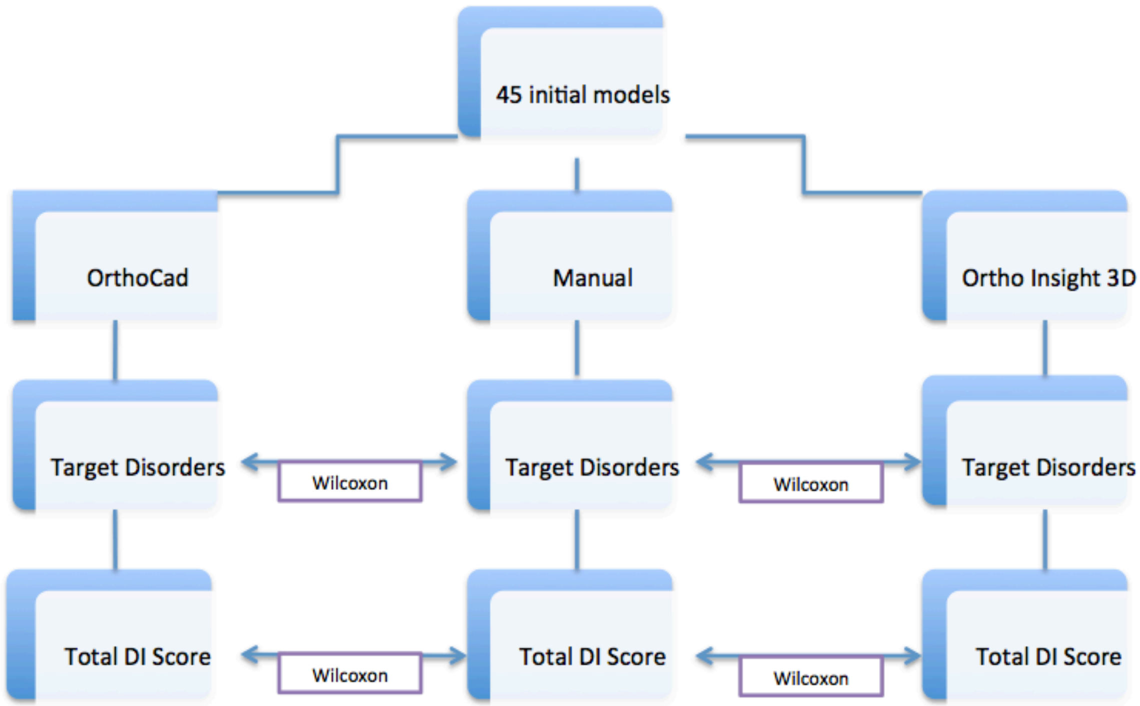

Figure 2. Accuracy of target disorders and total DI scores assessed with Wilcoxon signed ranks test

3.3.4 Statistical Analysis

Statistical analysis was conducted with SPSS vs.20 (Chicago, IL). Exam of the raw set with Shapiro-Wilk tests of normality reveals that nonparametric tests are more appropriate to analyze the data set due to non-normal distribution of all variables. A Spearman Rho (r) correlation was computed to assess test-retest reliability for all the variables involved in this study. Wilcoxon signed ranks tests were used to test the differences between mean ranks of the total DI Scores and its target disorder scores obtained manually, digitally with OrthoCad, and digitally with Ortho Insight 3D.

4. RESULTS

4.1 Reliability Test Results

The analysis for intra-rater and inter-rater reliability was done for 15 subjects. As with other reliability coefficients, (r) should be above 0.8 to indicate good indices of reliability.

The results for intra-rater reliability of the primary investigator assessed with the Spearman's Rho correlation show repeated measurements with excellent agreement (r>0.8) for all manual measurements except the category of edge-to-edge open bite (r=0.535), for all OrthoCad measurements except buccal crossbite, spacing, and diastema (range of correlation r=0.608-0.681) and all Ortho Insight 3D measurements except left occlusal relationship (r=0.775). Intra-rater reliability results of manual measurements for the second examiner also show excellent agreement (r>0.8) in all measurements except lateral open bite (r=0.732).

The two sets of measurements made by two independent examiners on the plaster models were found to be significantly correlated (r>0.8) except in measurements of overjet (points), overbite (points), edge-to-edge open bite, and right occlusal relationship (range of correlation r=0.608-0.762) and low correlation with respect to spacing (r=0.221). The two sets of measurements made by two independent examiners on OrthoCad digital models were found to be significantly correlated (r>0.8) in all categories except left occlusal relationship and diastema (r=0.763 and 0.535 respectfully). Furthermore, all three methods demonstrated significantly high intra-rater and inter-rater reliability (r>0.9) with respect to the total DI score despite mild to low correlations in a few categories.

4.2 Accuracy Results

Results of the Wilcoxon signed ranks test are shown in Table III. When compared to the gold standard of DI calculation using plaster models, there was a significant difference ($p<0.05$) in Ortho Insight 3D measurements of overjet (millimeters and points), crowding (millimeters and points), edge-to-edge open bite, right occlusal relationship, left occlusal relationship, spacing, and total DI score. OrthoCad differed significantly ($p<0.05$) from manual calculation only in target disorders of overjet (points) and overbite (mm). There was no significant difference in calculation of total DI score when obtained digitally with OrthoCad compared to manually ($p=0.563$).

TABLE III

WILCOXON SIGNED RANKS TESTS OF DIFFERENCES BETWEEN METHOD PAIRS

Criterion	**OrthoCad Vs. Manual** ***(p value)***	**Ortho Insight 3D vs. Manual** ***(p value)***
OVERJET (points)	.029*	.000*
NEGATIVE OVERJET (points)	.607	.774
OVERBITE (points)	.141	.386
EDGE-TO-EDGE OPEN BITE (points)	.118	.016*
OPEN BITE (points)	.860	.527
LATERAL OPEN BITE (points)	.157	.134
CROWDING (points)	.792	.029*
R OCCLUSAL RELATIONSHIP (points)	.317	.004*
L OCCLUSAL RELATIONSHIP (points)	.059	.006*
LINGUAL CROSSBITE (points)	.554	.084
BUCCAL CROSSBITE (points)	1.000	.317
MIDLINE DISCREPANCY (points)	1.000	.564
SPACING (points)	.655	.034*
DIASTEMA (points)	.317	.317
TOTAL DI SCORE (points)	.563	.004*
OVERJET (mm)	.226	.000*
OVERBITE (mm)	.011*	.453
CROWDING (mm)	.534	.005*

* denotes significance ($p<0.05$)

For the purpose of clinical relevance, descriptive statistics were examined to depict how statistically significant disorders differed between Ortho Insight 3D and manual calculation (Table IV). Ortho Insight 3D measurements were larger than all manual and OrthoCad measurements except in the category spacing. OrthoCad significantly differed from manual in overjet (points) and overbite (mm). According to descriptive statistics, OrthoCad overjet was 0.29 points larger than manual overjet and OrthoCad overbite was 0.62 millimeters smaller than manual overbite.

TABLE IV

DESCRIPTIVE MEAN (SD) AND MEAN DIFFERENCE FOR SIGNIFICANTLY DIFFERENT ORTHO INSIGHT 3D AND MANUAL MEASUREMENTS

Criterion	**Ortho Insight 3D (OI)**	**Manual (M)**	**Mean Difference (OI-M)**
OVERJET (points)	2.76 (1.40)	1.62 (1.53)	1.14
EDGE-TO-EDGE OPEN BITE (points)	1.11 (1.79)	0.47 (0.87)	0.64
CROWDING (points)	3.36 (2.70)	2.67 (2.65)	0.69
R OCCLUSAL RELATIONSHIP (points)	2.09 (1.96)	1.53 (1.59)	0.56
L OCCLUSAL RELATIONSHIP (points)	1.73 (1.72)	1.31 (1.53)	0.42
SPACING (points)	0.18 (0.58)	0.44 (0.94)	-0.26
TOTAL DI SCORE (points)	19.13 (12.31)	16.42 (10.82)	2.71
OVERJET (mm)	5.77 (2.40)	3.82 (2.96)	1.95
CROWDING (mm)	5.07 (3.78)	4.19 (3.75)	0.88

Due to the low reliability of edge-to-edge open bite, spacing, and diastema scoring evident in the results of Spearman's rho correlation test, another Wilcoxon test was examined for total DI scores calculated without those variables. The results of that test show no significant difference between manual and OrthoCad total DI scores ($p=.850$) but a significant difference between Ortho Insight 3D and manual total DI scores ($p=.002$). The results are similar to initial Wilcoxon results despite removing variables with low reliability.

5. DISCUSSION

5.1 Discussion

Repeated measurements for most target disorders and all total DI scores were highly correlated when compared within each examiner and between examiners. Pulfer et al. evaluated DI score with respect to treatment outcome and assessed reliability of six student investigators using 20 casts. Authors reported reliability using Cohen's kappa coefficient that ranged from 0.68 to 0.94 for DI components and stated that data was collected once the correlation coefficient was greater than 0.61 (Pulfer et al., 2009). The same data set was used in two subsequent studies and authors of all three studies began data collection once investigators were within 5% of the collective mean (Pulfer et al., 2009; Parrish et al., 2011; Schafer et al., 2011). It is unclear whether intra-rater reliability was assessed. Also, authors do not explain if only the total DI score needed to be within 5% of the collected mean total DI score or if each component of the DI score needed to be within 5% of the collected mean for that component. The present study only used two examiners for each method and thus application of a 5% range around the collective mean would not be appropriate. If the same minimal correlation coefficient of 0.61 was to be applied to the reliability of components in the present study, only edge-to-edge open bite, diastema, and spacing categories had a correlation coefficient below 0.61. When those variables were removed from total DI scores, similar outcomes were obtained when compared to DI scores calculated with all variables. This study used Spearman Rho test to evaluate correlation and the previous studies used Cohen's kappa coefficient. One way to improve inter-rater reliability is have a discussion following independent scoring to resolve differences in scoring. This was not completed for the present study because authors wanted to simulate a clinical environment where a candidate would prepare for ABO Phase III examination wherein examiners and examinees cannot compare or discuss scoring. Also because there are no published studies evaluating reliability of the DI, this study serves as a baseline for intra-rater and inter-rater reliability.

The results from this study indicate that ABO DI total scores obtained from Motion View Ortho Insight 3D are significantly different than those obtained manually on plaster casts, the current gold standard. The target disorders most likely for this difference are overjet, right and left occlusal relationships, crowding, and spacing. However, total DI scores obtained with OrthoCad software were not significantly different than manual total DI scores. There are a few reasons this may have occurred.

Some may suggest that the use of alginate impressions for OrthoCad would result in distortion of the virtual models. Coleman et al. found significant changes between models poured within one hour of the impression compared to models poured 24 hours later (Coleman et al., 1979). In contrast, a more recent study found no significant differences in accuracy of measurements performed on plaster models poured within one hour and measurements performed on plaster models poured 3-5 days later (Dalstra and Melsen, 2009). Furthermore, if distortion of the alginate was a significant reason for inaccuracy of digital models, one would expect OrthoCad total DI score and target disorders to be significantly different than manual scores. The results of the present study actually show no difference between OrthoCad and manual DI score or it parameters with the exception of overjet, which may differ due to tooth selection. Ortho Insight 3D models were obtained by scanning the original models therefore any inaccuracy in measurements could not be attributed to alginate distortion and must come from landmark identification, intrinsic processing of the scanned model, or measurement algorithms of the proprietary software.

Digital landmark identification has been reported as a problem in many previous studies (Bell et al., 2003; Costalos et al., 2005; Hildebrand et al., 2008; Santoro et al., 2003; Tomassetti et al., 2001; Zilberman et al., 2003). With any computer-based system, a significant learning curve is to be expected. Digital model analysis relies on accurately identifying points of a 3D object on a 2D screen. Additionally, although the ABO initially sought to create an objective index for calculating case complexity there is some subjectivity in the selection of landmarks. For example, the category of open bite or edge-to-edge open bite excludes teeth that are blocked out due to crowding however almost all cases demonstrate crowding so there is a grey area where one might consider the canine to be in an open bite relationship and crowded and another examiner may exclude that canine due to the crowding. Also, there are a few instances where a maxillary or mandibular incisor is partially blocked out due to crowding and the instructions for the categories of overjet, negative overjet, and overbite are not clear whether those teeth may be included for scoring.

For the purpose of the study, accuracy was assessed by comparing digital DI calculation to manual calculation (our gold standard). The underlying assumption in this study is that manual calculation is the most accurate. However, it is possible that although there is a statistically significant difference between methods, Ortho Insight 3D may actually be more precise in selecting landmarks when compared to landmarks and occlusal relationships visualized on plaster models. For example, in the category of spacing, presence of at least 0.5mm of space on mesial and distal aspects of 4 separate teeth is necessary to obtain points. The computer software may be more precise in measuring space less than 0.5mm millimeters because measurements are recorded to the nearest 0.1mm whereas the manual visualization of space was most likely rounded up to 0.5mm when spacing below 0.5mm was present.

For the categories of overjet, overbite, and crowding, both the actual measurement in mm and the points allocated for that measurement were recorded and analyzed for reliability and accuracy. Both millimetric measurements and point allocation for the three categories were well correlated ($r>0.7$) within examiners and between examiners for manual, OrthoCad, and Ortho Insight 3D methods. Accuracy of overjet was significantly different between OrthoCad and manual calculation when recorded in points but not significantly different in millimeters. This implies there is a defect in the software's allocation of points for that category. The opposite was true for OrthoCad overbite with a significant difference in millimetric measurement but points awarded for overbite was not significantly different from manual. This is most likely due to the points allocated for a broad range of overbite measurements. When compared to previous studies evaluating overjet between digital and manual calculation, the results of millimetric overjet obtained by OrthoCad in this study demonstrate similar findings (Quimby et al., 2004; Otuyemi and Jones, 1995; Santoro et al., 2003; Stevens et al., 2006;

Watanabe-Kanno et al., 2009). The smaller millimetric overbite measurements obtained by OrthoCad when compared to manual overbite measurements found in this study are consistent with some previous studies who found significantly smaller overbite measurements in digital models (Quimby et al., 2004; Santoro et al., 2003) and contrary to others where no significant difference was noted between manual and digital overbite (Bootvong et al., 2010). In Ortho Insight 3D, both millimetric and point allocation for overjet and crowding were statistically different than manual calculation, which may be attributed to landmark identification or measurement algorithms of the software.

Although the results may be statistically significant, the question of clinical significance is still unclear. For example, the difference between manual and Ortho Insight 3D mean total DI scores was 2.71 points. This may not impact a candidate applying for ABO‘s phase III clinical examination if the score of the subject is in the middle of the range accepted by ABO. If however the subject’s total DI score was borderline with respect to the minimal DI score that ABO requires, then 2.71 points may either allow the subject to be accepted or rejected. Based on the results of the present study, authors suggest checking the DI manually for those planning to use digital models in application to Phase III of the ABO certification and to keep in mind that the digital DI obtained from Ortho Insight 3D is an overestimation of the case complexity. If the intent of calculating the DI score is to assist the clinician in predicting treatment duration, then 2.71 points may lead the clinician to over-estimate treatment duration. However, estimated treatment duration is usually expressed as a range and that range may not change based on the relatively small change in DI score. Another clinical application of a digital Discrepancy Index may be for state funding for patients demonstrating need. A slightly higher threshold may be used to compensate for the overestimation in case complexity but the digital DI score may be an objective method for allocating funds.

The efficiency of digital model analysis was not compared in the present study. The primary investigator estimated about 30 minutes for digitally calculating the DI score once the digital models were returned from OrthoCad and about 20 minutes for digitally manipulating landmarks for calculation of the DI score once the models were scanned in Ortho Insight 3D. This implies time saving for a busy orthodontic office however the most time consuming step in obtaining digital models from a laser scanner such as Ortho Insight is scanning the models. In a private practice setting, the scanning of models or impressions would most likely be delegated to staff and not affect doctor time. Ortho Insight 3D has the capability to scan impressions and bite registration but it takes considerably more time to scan impressions and bite registrations and properly occlude the digital models compared to scanning trimmed, plaster models. For time saving purposes and to eliminate variables associated with impression scanning, trimmed plaster models were scanned.

5.2 Limitations

A limitation of the present study is the small number of examiners included for inter-rater reliability. Ideally, multiple residents and multiple ABO representatives would grade the same sample manually and digitally multiple times.

5.3 Future Research

The present study aimed to evaluate accuracy of the Discrepancy Index with respect to manual calculation and digital calculation. Based on the results, there are a few categories that are not reliably measured between examiners. A future study that incorporated many more examiners, ideally ABO graders, orthodontic residents, and practicing orthodontists, may be beneficial for determining reliability on a larger scale. From those results, categories may change or instructions with respect to those categories may improve reliability. Because digital landmark identification seems to inhibit the accuracy and reproducibility of digital measurements, a study evaluating repeated landmark selection might be warranted using model superimposition. The present study utilized the original plaster casts for fabrication of Ortho Insight 3D digital models however the scanner is able to capture models from impressions (alginate and polyvinyl siloxane), untrimmed plaster models, and trimmed plaster models. A future study may compare accuracy of landmark identification in all methods of digital model fabrication and record the time it takes to scan and occlude the different types of models and impressions.

6. CONCLUSION

Based on the extent and limitations of the present study, the following conclusions were obtained with respect to manual and digital calculation of ABO Discrepancy Index:

- All methods (manual, OrthoCad, Ortho Insight 3D) of calculating total DI score were significantly correlated for intra-rater and inter-rater reliability of repeated measurements.
- No significant difference was demonstrated between total DI score or any of its target disorders scores except overjet measured with OrthoCad compared to manual calculation.
- Significant differences were found in Motion View's Ortho Insight 3D when compared to manual for the calculation of overjet (in mm and points), edge-to-edge open bite, crowding (in mm and points), right and left occlusal relationship, spacing, and total DI score.
- For categories demonstrating a significant difference between Ortho Insight 3D and manual measurements, Ortho Insight 3D scores were larger than manual scores with the exception of spacing.

LITERATURE CITED

American Board of Orthodontics: 2013. *Clinical Examination Guide* (6th Edition), Last update 12-17-13 2013 [cited 07/01/2013 2013]. Available from http://www.americanboardortho.com/professionals/downloads/clinical_examination_guide.pdf.

Bell, A., A.F. Ayoub, and P. Siebert.: Assessment of the accuracy of a three-dimensional imaging system for archiving dental study models. *J Orthod,* 30:219-23, 2003.

Bootvong, K., Z. Liu, C.McGrath, U. Hagg, , R.W. Wong, M. Bendeus, and S. Yeung.: Virtual model analysis as an alternative approach to plaster model analysis: reliability and validity. *Eur J Orthod*, 32:589-95, 2010.

Brook, P. H. and W.C. Shaw.: The development of an index of orthodontic treatment priority. *Eur J Orthod*, 11:309-20, 1989.

Cangialosi, T. J., M. L. Riolo, S. E. Owens, Jr., V. J. Dykhouse, A. H. Moffitt, J. E. Grubb, P. M. Greco, J. D. English, and R. D. James.: The ABO discrepancy index: a measure of case complexity. *Am J Orthod Dentofacial Orthop*, 125:270-8, 2004.

Casko, J. S., J. L. Vaden, V. G. Kokich, R. D. Damone, J. James, T. J. Cangialosi, M. L. Riolo, S.E. Owens, Jr., and E.D. Bills.: Objective grading system for dental casts and panoramic radiographs. American Board of Orthodontics. *Am J Orthod Dentofacial Orthop*, 114:589-99, 1998.

Coleman, R. M., J. H. Hembree, Jr., and F. N. Weber.: Dimensional stability of irreversible hydrocolloid impression material. *Am J Orthod*, 75:438-46, 1979.

Costalos, P. A., K. Sarraf, T. J. Cangialosi, and S. Efstratiadis.: Evaluation of the accuracy of digital model analysis for the American Board of Orthodontics objective grading system for dental casts. *Am J Orthod Dentofacial Orthop*, 128:624-9, 2005.

Dalstra, M., and B. Melsen.: From alginate impressions to digital virtual models: accuracy and reproducibility. *J Orthod*, 36:36-41; discussion 14, 2009.

Daniels, C., and S. Richmond.: The development of the index of complexity, outcome and need (ICON). *J Orthod*, 27:149-62, 2000.

DeGuzman, L., D. Bahiraei, K. W. Vig, P. S. Vig, R. J. Weyant, and K. O'Brien.: The validation of the Peer Assessment Rating index for malocclusion severity and treatment difficulty. *Am J Orthod Dentofacial Orthop*, 107:172-6, 1995.

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

Fleming, P. S., V. Marinho, and A. Johal.: Orthodontic measurements on digital study models compared with plaster models: a systematic review. *Orthod Craniofac Res*, 14:1-16, 2011.

Fox, N. A.: The first 100 cases: a personal audit of orthodontic treatment assessed by the PAR (peer assessment rating) index. *Br Dent J*, 174:290-7, 1993.

Hildebrand, J. C., J. M. Palomo, L. Palomo, M. Sivik, and M. Hans.: Evaluation of a software program for applying the American Board of Orthodontics objective grading system to digital casts. *Am J Orthod Dentofacial Orthop*, 133:283-9, 2008.

Keim, R. G., E. L. Gottlieb, A. H. Nelson, and D. S. Vogels, 3rd.: 2008 JCO study of orthodontic diagnosis and treatment procedures, part 1: results and trends. *J Clin Orthod*, 42:625-40, 2008.

Kim, J., G. Heo, and M. O. Lagravere.: Accuracy of laser-scanned models compared to plaster models and cone-beam computed tomography. *Angle Orthod*, 2013. doi: 10.2319/051213-365.1.

Leifert, M. F., M. M. Leifert, S. S. Efstratiadis, and T. J. Cangialosi.: Comparison of space analysis evaluations with digital models and plaster dental casts. *Am J Orthod Dentofacial Orthop*, 136:16 e1-4, 2009.

Mah, J., and D. Hatcher.: Current status and future needs in craniofacial imaging. *Orthod Craniofac Res,* 6 Suppl 1:10-6; discussion 179-82, 2003.

Martensson, B., and H. Ryden.: The holodent system, a new technique for measurement and storage of dental casts. *Am J Orthod Dentofacial Orthop,* 102:113-9, 1992.

Mavreas, D., and A. E. Athanasiou.: Factors affecting the duration of orthodontic treatment: a systematic review. *Eur J Orthod,* 30:386-95, 2008.

Mayers, M., A. R. Firestone, R. Rashid, and K. W. Vig.: Comparison of peer assessment rating (PAR) index scores of plaster and computer-based digital models. *Am J Orthod Dentofacial Orthop*, 128:431-4, 2005.

Noble, J., F. J. Hechter, N. E. Karaiskos, N. Lekic, and W. A. Wiltshire.: Future practice plans of orthodontic residents in the United States. *Am J Orthod Dentofacial Orthop,* 135:357-60, 2009.

Okunami, T. R., B. Kusnoto, E. BeGole, C. A. Evans, C. Sadowsky, and S. Fadavi.: Assessing the American Board of Orthodontics objective grading system: digital vs plaster dental casts. *Am J Orthod Dentofacial Orthop,* 131:51-6, 2007.

Otuyemi, O. D., and S. P. Jones.: Methods of assessing and grading malocclusion: a review. *Aust Orthod J,* 14:21-7, 1995.

Parrish, L. D., W. E. Roberts, G. Maupome, K. T. Stewart, R. W. Bandy, and K. S. Kula.: The relationship between the ABO discrepancy index and treatment duration in a graduate orthodontic clinic. *Angle Orthod,* 81:192-7, 2011.

Pulfer, R. M., C. T. Drake, G. Maupome, G. J. Eckert, and W. E. Roberts.: The association of malocclusion complexity and orthodontic treatment outcomes. *Angle Orthod,* 79:468-72, 2009.

Quimby, M. L., K. W. Vig, R. G. Rashid, and A. R. Firestone.: The accuracy and reliability of measurements made on computer-based digital models. *Angle Orthod*, 74:298-303, 2004.

Redmond, W. R.: Digital models: a new diagnostic tool. *J Clin Orthod*, 35:386-7, 2001.

Richmond, S., W. C. Shaw, K. D. O'Brien, I. B. Buchanan, R. Jones, C. D. Stephens, C. T. Roberts, and M. Andrews.: The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod*, 14:125-39, 1992.

Santoro, M., S. Galkin, M. Teredesai, O. F. Nicolay, and T. J. Cangialosi.: Comparison of measurements made on digital and plaster models. *Am J Orthod Dentofacial Orthop*, 124:101-5, 2003.

Savastano, N. J., Jr., A. R. Firestone, F. M. Beck, and K. W. Vig.: Validation of the complexity and treatment outcome components of the index of complexity, outcome, and need (ICON). *Am J Orthod Dentofacial Orthop*, 124:244-8, 2003.

Schafer, S. M., G. Maupome, G. J. Eckert, and W. E. Roberts.: Discrepancy index relative to age, sex, and the probability of completing treatment by one resident in a 2-year graduate orthodontics program. *Am J Orthod Dentofacial Orthop*, 139:70-3, 2011.

Shaw, W. C., S. Richmond, and K. D. O'Brien.: The use of occlusal indices: a European perspective. *Am J Orthod Dentofacial Orthop*, 107:1-10, 1995.

Stevens, D. R., C. Flores-Mir, B. Nebbe, D. W. Raboud, G. Heo, and P. W. Major.: Validity, reliability, and reproducibility of plaster vs digital study models: comparison of peer assessment rating and Bolton analysis and their constituent measurements. *Am J Orthod Dentofacial Orthop*, 129:794-803, 2006.

Summers, C. J.: The occlusal index: a system for identifying and scoring occlusal disorders. *Am J Orthod*, 59:552-67, 1971.

Tomassetti, J. J., L. J. Taloumis, J. M. Denny, and J. R. Fischer, Jr.: A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. *Angle Orthod*, 71:351-7, 2001.

Veenema, A. C., C. Katsaros, S. C. Boxum, E. M. Bronkhorst, and A. M. Kuijpers-Jagtman.: Index of Complexity, Outcome and Need scored on plaster and digital models. *Eur J Orthod*, 31:281-6, 2009.

Vu, C. Q., W. E. Roberts, J. K. Hartsfield, Jr., and S. Ofner.: Treatment complexity index for assessing the relationship of treatment duration and outcomes in a graduate orthodontics clinic. *Am J Orthod Dentofacial Orthop*, 133:9 e1-13, 2008.

Watanabe-Kanno, G. A., J. Abrao, H. Miasiro Junior, A. Sanchez-Ayala, and M. O. Lagravere.: Reproducibility, reliability and validity of measurements obtained from Cecile3 digital models. *Braz Oral Res,* 23:288-95, 2009.

Zilberman, O., J. A. Huggare, and K. A. Parikakis.: Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. *Angle Orthod*, 73:301-6, 2003.

VITA

NAME:	Kristina Dragstrem
EDUCATION:	University of Illinois, Chicago, IL, M.S., Oral Sciences, and May 2014
	University of Illinois, Chicago, IL, Certificate, Orthodontics, May 2014
	University of Illinois, Chicago, IL, D.D.S., May 2011
	University of Illinois, Urbana-Champaign, IL, B.S., Molecular and Cellular Biology, December 2006
HONORS:	American Institute of Orthodontic Research Award, May 2011
	Student Achievement Award, May 2011
	UIC Clinic and Research Clinical Science Award, February 2011
	Dr. John W. Davis Award, May 2010
PUBLICATION:	Dragstrem, K.G., Yuan, J.C., Lee, D.J., Sukotjo, C., Galang, M.T. 2012. “Gender equality in orthodontic literature and leadership in the United States.” Orthodontics (Chic.) no 13(1): 176-83.
PROFESSIONAL MEMBERSHIPS:	American Association of Orthodontists, 2011-present
	Illinois Society of Orthodontists, 2011-present
	American Dental Association, 2007-present
	Chicago Dental Society, 2007-present

APPENDIX

UNIVERSITY OF ILLINOIS
AT CHICAGO

Office for the Protection of Research Subjects (OPRS)
Office of the Vice Chancellor for Research (MC 672)
203 Administrative Office Building
1737 West Polk Street
Chicago, Illinois 60612-7227

Notice of Determination of Human Subject Research

January 28, 2013

20130090-72610-1

20130090-72610-1

Kristina Dragstrem, BS, DDS
Orthodontics
801 S. Paulina
M/C 841
Chicago, IL 60612
Phone: (630) 881-3340 / Fax: (312) 996-0873

RE: **Protocol # 2013-0090**
Accuracy and Reliability of Digital Discrepancy Index Calculations

Sponsor: None

Dear Dr. Dragstrem:

The UIC Office for the Protection of Research Subjects received your "Determination of Whether an Activity Represents Human Subjects Research" application, and has determined that this activity **DOES NOT meet the definition of human subject research** as defined by 45 CFR 46.102(f).

You may conduct your activity without further submission to the IRB.

If this activity is used in conjunction with any other research involving human subjects or if it is modified in any way, it must be re-reviewed by OPRS staff.