

The validity of little's irregularity index in the maxillary and mandibular arches. -In Vitro study-

Dina Sabry Balboul¹, Khalid Moustafa Fawzy², Dina Hussien El-Ghoul³.

Abstract:

Objective: to assess the validity of Little Irregularity Index in both the maxillary and mandibular arches. **Materials and methods:** The study was an invitro retrospective trial performed in the Digital Center of the Orthodontic Department, Faculty of Dentistry, Ain-Shams University. The study was conducted on 100 maxillary and 100 mandibular study casts representing a wide range of crowding. Prior to the initiation of the study, all study casts were converted into digital models by the primary investigator. The study procedure was divided into two phases. Phase 1, where five orthodontic professors assessed the 200 study casts selected for the study to assign a subjective score describing the anterior crowding and phase 2, where Four examiners of different orthodontic background levels used the digitized version of the study models to determine quantitatively the amount of anterior crowding as recommended by Dr. Little. **Results:** ICC for the maxillary and mandibular scores by the five orthodontic professors was notably high (>0.97). Also, for the measurements conducted by the four examiners on the digital models of both arches were (>0.99). The LOA were excessively broad in the maxillary scores, $(2.42 + 9.84) = 12.26$. Likewise, for the mandibular scores the LOA were again too extensive, $(2.67 + 8.17) = 10.84$. CCC values between the scores of the

five orthodontic professors and the scores of the four examiners for the maxillary was 0.418 and for the mandibular was 0.534. **Conclusion:** The validity of LII in the maxillary and mandibular arches could not be confirmed.

Introduction:

Anterior dental crowding is one of the most frequently encountered and widely spread malocclusions in orthodontics. Not only does it cause biological problems, but also affects the facial esthetics as a whole, making it one of the most frequent patients' chief complaint.

Dental crowding was stated the most prevalent malocclusion trait according to a study done in Egypt and Saudi Arabia in 2020 [1]. It represents 63.3% of malocclusions in Saudis, and 57.7% in Egyptians. Another survey stated that 78% of the American population has varying degrees of incisor irregularity, 15% of which were classified as severe to extreme [2].

The term dental crowding is too vague to be used for description of the severity of the condition on its own, as it allows a wide range of interpretation, also adjectives like mild, moderate, severe and significant are subjective and lack quantification. To assess incisor crowding, several methods were proposed, but

¹ BDS 2014, Misr International University.

² Professor of Orthodontics, Department of Orthodontics, Faculty of Dentistry, Ain-Shams University.

³ Professor of Orthodontics, Department of Orthodontics, Faculty of Dentistry, Ain-Shams University.

none of which offered simplicity and true quantification.

Little's irregularity index (LII) was originally formulated to objectively score mandibular incisor alignment for epidemiological studies and insurance programs, also to assess the degree of initial malalignment to compare it with post treatment and post retention results[3]. The traditional application of Little's irregularity index involves horizontal linear measurement of the displacement of the anatomical contact points of each mandibular incisor from the mesial contact point of the left canine to the mesial contact point of the right canine. The sum of the five displacements represent the relative degree of anterior irregularity [3]. All vertical discrepancies, as well as mesiodistal spacing of contact points must be ignored, only the labiolingual shift from proper arch form is recorded [3].

The orthodontic community in recent years has extended the application of Little's Irregularity Index to include the maxillary arches as well as the mandibular arches to allow the evaluation of various orthodontic brackets [4,5], retainer effectiveness[6–8]and treatment modalities [8,9]. The use of Little Irregularity Index on the maxillary anterior teeth has not been studied nor modified to be used on such different teeth anatomy and size as it was originally invented for describing the mandibular anterior crowding. For instance, the identification of anatomical contact points on maxillary incisors may require more guess work compared with the corresponding mandibular teeth due to the increased labio-

lingual width of the maxillary anterior teeth[4],and thicker proximal aspect so that it's no longer a contact point, but rather a contact area.

The use of 3D digitizing technologies has been the subject of considerable innovation in recent years. Digital models of dental arches require less storage space than traditional study casts[10] and are not susceptible to damage or degradation over time [11]. Moreover, digital models of dental arches can be viewed instantly in multiple locations [12]. Digital models made the whole storage, diagnosis and treatment planning much easier and time saving. Also, the reliability and validity of the digital measurement of LII was proven[13]. For that reason, we decided to use 3D digital models to carry out our study to evaluate the validity and reliability of Little's Irregularity Index in the maxillary and mandibular arches.

Materials and methods:

The study was an invitro retrospective trial performed in the Digital Center of the Orthodontic Department, Faculty of Dentistry, Ain-Shams University.

The study was conducted on 100 maxillary and 100 mandibular study casts and digital study models representing a wide range of crowding. Casts were obtained from the orthodontic diagnostic records of patients that were treated or currently being treated at the Outpatient Clinic of the Orthodontic Department, Faculty of Dentistry, Ain-Shams University.

Study casts were then assessed for eligibility according to the following inclusion and exclusion criteria.

Inclusion criteria:

The study casts were selected from the diagnostic records of patients with;

1. Ages ranging from 16 to 30 years.
2. All permanent maxillary and mandibular teeth were present and fully intact, regardless of the presence or absence of third molars.
3. Crowding of anterior maxillary and/or mandibular teeth with varying degrees of severity.

Exclusion criteria:

The casts from the diagnostic records of patients with the following criteria were excluded;

1. Patients with craniofacial syndromes or complex medical history.
2. Patients with previous orthodontic treatment.

3. Patients with anterior fixed prosthesis.
4. Proximal attrition in anterior maxillary and mandibular teeth.
5. Proximal caries in anterior maxillary and mandibular teeth.
6. Anterior maxillary or mandibular tooth anomaly.

The sample size calculation was based on a study by Adam H. Dowling at al. "The reliability of Little's Irregularity Index for the upper dental arch using three dimensional (3D) digital models"[14]. The estimated Intraclass Correlation Coefficient (ICC) was 0.94, considering that the minimum acceptable ICC was 0.91 to ensure the power of 90% with a significance level of 0.05 (Fig. 1). A sample size of 175 casts was initially required, but it was subsequently raised to 200 casts to ensure a 90% power level for other statistical tests employed other than ICC to enhance the accuracy of the study.

Sample Size Calculator (web)

Intraclass Correlation Coefficient (ICC) - Hypothesis Testing¹

Minimum acceptable reliability (ICC) (q_0):	0.91
Expected reliability (ICC) (q_1):	0.94
Significance level (α):	0.05 Two-tailed
Power ($1 - \beta$):	90 %
Number of raters/repetitions per subject (k):	3
Expected dropout rate:	0 %

Calculate Reset

Sample size, $n =$ 175

Sample size (with 0% dropout), $n_{drop} =$ 175

Figure 1: Sample size calculation using Sample Size Calculator (Web)

Collection of the study casts:

A survey was conducted on the diagnostic records of patients that were treated or currently being treated at the Outpatient Orthodontic Clinic, Ain-Shams University involving 1000 orthodontic diagnostic study casts, encompassing those with various degrees of anterior crowding. From this pool, 100 maxillary and 100 mandibular study casts were selected, ensuring adherence to the standards of orthodontic study models. These casts were thoroughly inspected to fulfill our inclusion criteria. If in doubt of any tooth condition, the patient's sheet and photographs were used to further assess the form and condition of the suspected teeth. The selected maxillary and mandibular casts were not necessarily of the same patient.

Converting Study Casts to 3D digital models

Prior to the initiation of the study, all study casts were converted into digital models by the primary investigator. Each coded study cast was scanned using 3-shape R-750 scanner¹ then rendered into a 3D stereolithographic digital model (.stl) through a specific software (ScanIt™ Orthodontics)². Sticky tack was used to fix and stabilize the base of study casts to the 3-Shape R-750 scanner's platform during the scanning process. Each STL file was named after the code that identified each model and saved. The digital models were viewed on a 17 inch monitor (Dell 1708 Flat Panel LCD) with a resolution of 1280 x 1024 pixels using the Digital center's computer at Ain-Shams University to run the software program. Ortho-Analyzer software³ was used for

conducting Little's Irregularity Index measurements.

Study procedure:

The study was designed to assess the validity of Little's Irregularity Index in the maxillary and the mandibular arches. The study procedure was divided into two phases.

Phase 1:

Five orthodontic professors assessed the 100 maxillary and 100 mandibular study casts selected for the study in order to assign a subjective score describing the anterior crowding (from the mesial aspect of the right canine to the mesial aspect of the left canine).

- 0 Perfect alignment
- 1-3 Minimal irregularity
- 4-6 Moderate irregularity
- 7-9 Severe irregularity
- >10 very severe irregularity

The scores that were given by the professors were considered the gold standard. The term "gold standard" refers to a benchmark that is available under reasonable conditions. It does not have to be the perfect test, but merely the best available one that has a standard with known results.[15]

Phase 2:

Four examiners of different orthodontic background levels and clinical experience used the digitized study models of the same 100 maxillary and 100 mandibular study casts to determine quantitatively the amount of anterior crowding (from the mesial aspect of the right canine to the mesial aspect of the left canine) as recommended by Dr. Little[3]. The four examiners included the primary investigator, a

¹ Orthodontic Digital Center, Ain-Shams University.

² 3Shape A/S, Copenhagen, Denmark

³ 3Shape A/S, Copenhagen, Denmark

colleague in the same academic year, and two examiners who were one academic year ahead and one academic year behind.

Starting with the primary investigator, the anterior crowding of the maxillary and mandibular digital models were measured using Little's irregularity index as proposed by Dr. Little.[3] Only 20 models were measured during each session to avoid error in measurement due to eye fatigue and to decrease the chance of subjective error.

Working in isolation, the examiner opened a 3D digital model file at random. Each model was rotated to the occlusal view and the linear measurement tool in Ortho-Analyzer software⁴ was used to measure Little's Irregularity Index by placing point markers on each of the 10 anatomic contact points from the mesial aspect of the upper right canine to the mesial aspect of the upper left canine (UR3-UR2, UR2-UR1, UR1-UL1, UL1-UL2, UL2-UL3) for the maxillary casts and from the mesial aspect of the lower right canine to the mesial aspect of the lower left canine (LR3-LR2, LR2-LR, LR1-LL1, LL1-LL2, LL2-LL3) for the mandibular casts then saved the file (Figs. 2, 3). It was possible for the examiner to enlarge the model by "zooming in" (to a maximum of x10) to accurately select each contact point. The localization of each contact point was examined in all three dimensions by manipulating the digital model. In the event where contact point identification was obscured by crowding, the examiner selected the contact point by estimating its position based on the visible portion of the tooth anatomy, choosing the most probable location.

Measurements were taken only when the points were properly aligned antero-posteriorly and in the horizontal plane. All the anatomic contact points were recorded parallel to the occlusal plane to make sure that all vertical discrepancies were disregarded. The sum of contact point displacement of each cast was then calculated to give a final index score.

After completing the whole 100 maxillary and 100 mandibular digital casts; the same process was then repeated two more times so that all casts were measured three times in total, three weeks apart. The other three examiners then followed the same protocol after being provided with a tutorial on the capabilities of the Ortho-Analyzer software⁵, specially how to locate the relevant anatomic contact points, followed by a 20-minutes practice session.

A total of 12,000 measurements were collected resulting from 5 contact point displacement on each digitized cast multiplied by 200 digitized casts across 3 occasions by 4 examiners and then used for statistical analysis.

⁴ 3Shape A/S, Copenhagen, Denmark

⁵ 3Shape A/S, Copenhagen, Denmark

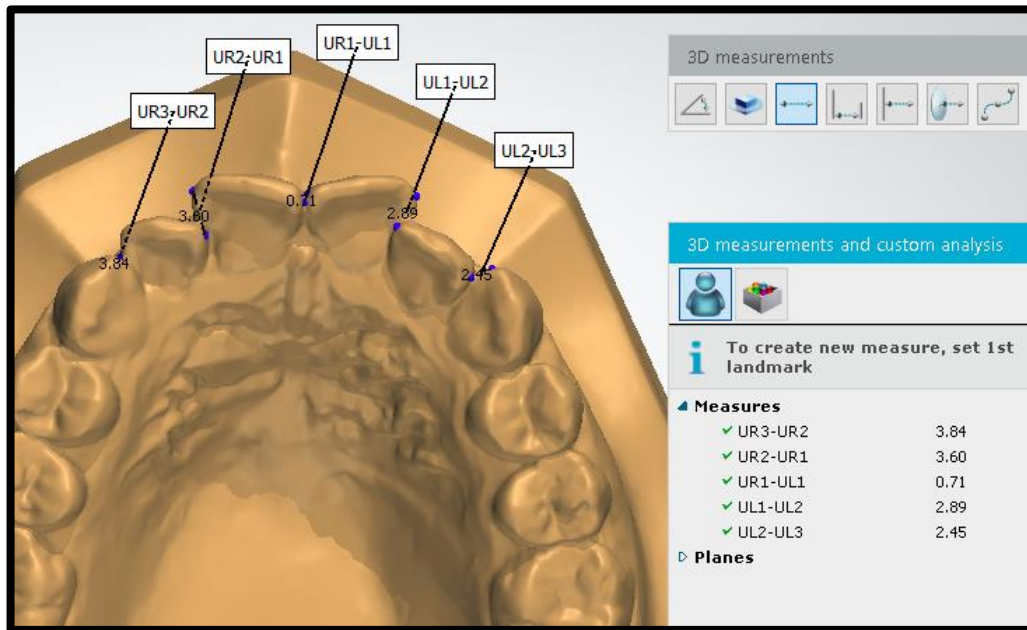


Figure 2: Little's Irregularity Index measured on scanned maxillary cast using the linear measurement tool in Ortho-Analyzer Software.

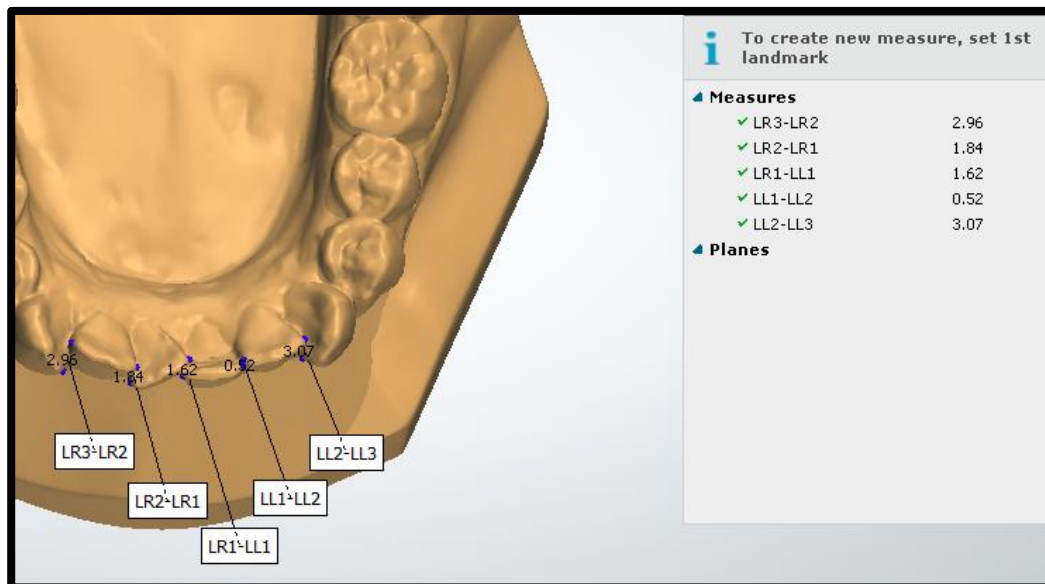


Figure 3: Little's Irregularity Index measured on scanned mandibular cast using the linear measurement tool in Ortho-Analyzer Software.

Methods of assessment and statistical analysis:

All data was collected, tabulated and subjected to statistical analysis by SPSS in general (version 20), while Microsoft office Excel was used for data handling and graphical presentation.

The Multi-rater Intraclass Correlation Coefficient was employed to evaluate the consistency among the scores conducted by the five orthodontic professors. Likewise, the same test was utilized to assess the measurements carried out by the four orthodontic examiners, ensuring uniformity across all examiners' measurements.

Validity:

To assess validity of Little's Irregularity Index, the mean of the final scores obtained in phase 1 by the five orthodontic professors were tested against the means of the final scores obtained in phase 2 by the four examiners for each cast.

The following tests were conducted:

- Dahlberg Error (DE), Relative Dahlberg Error (RDE).**

- Bland and Altman limits of agreement with its 95% confidence limits.**
- Concordance Correlation Coefficient.**

Results:

The Multi-rater Intraclass Correlation Coefficient (ICC) for both the maxillary and mandibular scores provided by the five orthodontic professors was notably high (>0.97), signifying excellent agreement among them. Similarly, for the measurements conducted by the four examiners on the digital models of both maxillary and mandibular arches (>0.99), indicating nearly perfect agreement among the four examiners.

DE and RDE for the maxillary and mandibular Little's Irregularity Index exhibited considerable separation between the means of the scores of the five orthodontic professors (gold standard) and the means of the scores of the four orthodontic examiners. RDE was exceeding 100% in both arches (Tables 1,2).

Table (1): Dahlberg Error (DE) and Relative Dahlberg Error (RDE) between LII maxillary scores of the five professors and the four examiners.

	Mean	SD	Dahlberg error DE	Relative Dahlberg Error RDE
Gold standard (professors)	3.05	2.44	3.42	112.3%
Measured by examiners	6.76	4.57		

Statistically significant RDE.

RDE < 10%: Perfect validity.
 RDE 10%-20%: Validity under concern.
 RDE>20%: Not valid

Table (2): Dahlberg Error (DE) and Relative Dahlberg Error (RDE) between LII mandibular scores of the five professors and the four examiners.

	Mean	SD	Dahlberg error DE	Relative Dahlberg Error RDE
Gold standard (professors)	2.68	2.41	2.75	102.5%
Measured by examiners	5.43	4.40		

RDE < 10%: Perfect validity.
 RDE 10%-20%: Validity under concern.
 RDE > 20%: Not valid

Statistically significant RDE.

Concerning the maxillary scores, the Limits of Agreement were excessively broad, with a range of $(2.42 + 9.84) = 12.26$. This wide range indicates that the difference between scores conducted by the examiners could deviate from those provided by the professors by a considerable margin of 12.26 (Fig. 4).

Similarly, for the mandibular scores. The Limits of Agreement were again too extensive, with a range of $(2.67 + 8.17) = 10.84$, suggesting that the difference between scores performed by the examiners could vary from those given by the professors by 10.84 (Fig. 5).

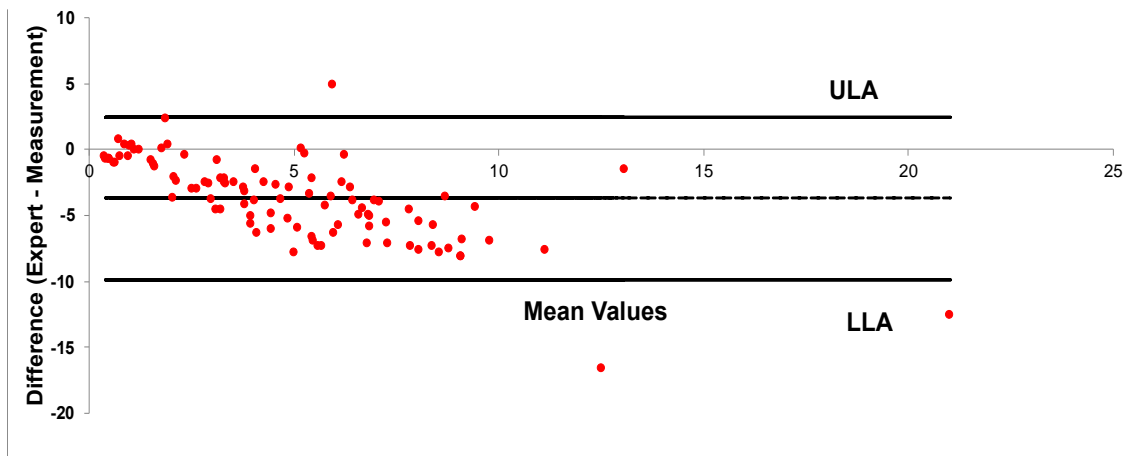


Figure 4: Bland-Altman plot of the maxillary scores.

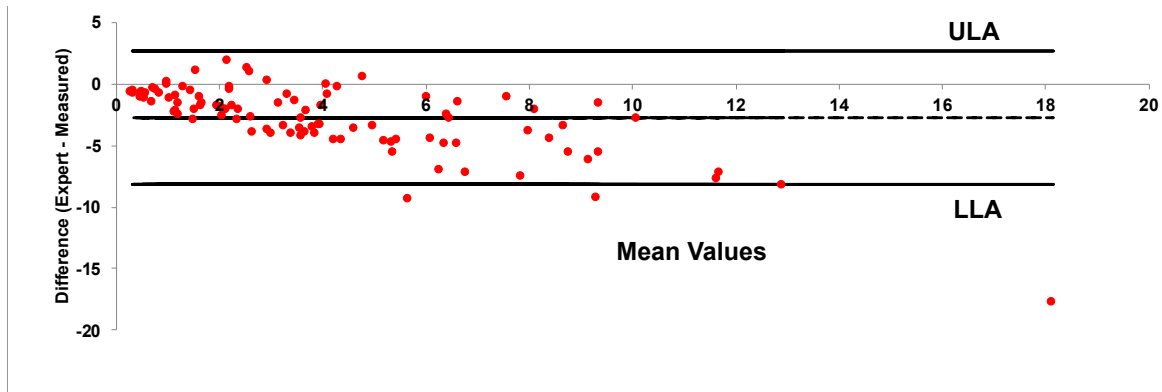


Figure 5: Bland-Altman plot of the mandibular scores.

Concordance Correlation Coefficient (CCC) values between the scores of the five orthodontic professors and the scores of the four examiners for both the maxillary (CCC=0.418) and mandibular (CCC=0.534) arches respectively. The Scatter plots of both arches showed that almost all points were far above the line of equality (Figs. 6,7).

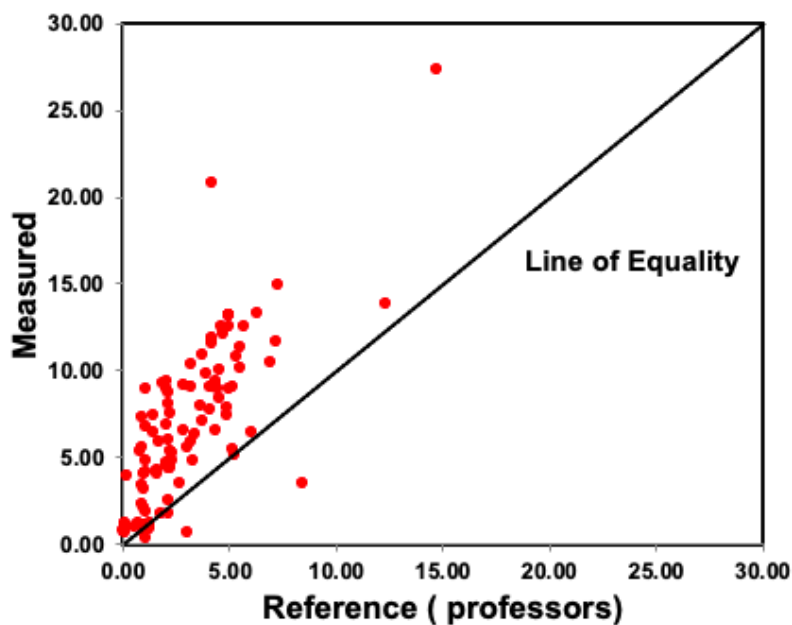


Figure 6: Scatter Plot of the maxillary scores.

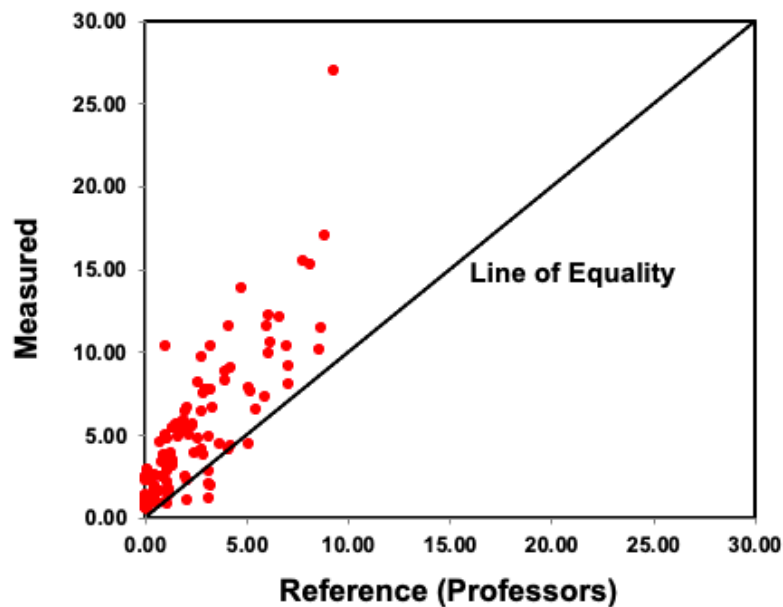


Figure 7: Scatter Plot of the mandibular scores.

Discussion:

In this study, Little's Irregularity Index, being one of the most commonly used methods to quantify anterior crowding in the maxillary and mandibular arches was investigated. While several methods have been proposed to evaluate incisor crowding, none have offered a straightforward and genuinely quantitative approach.

Prior to this study, there was a noticeable absence in the literature of any research evaluating the validity of Little's Irregularity Index specifically in the maxillary arch. Consequently, we adopted a methodology similar to the one used by Dr. Little[3] for assessing the validity of LII in the mandibular arch. This involved a comparison between subjective scores provided by orthodontic professors, considered as the gold standard, during phase 1 of the study and the quantitative

measurements conducted by four examiners using digital study models in phase 2. The objective of the study was to ascertain the extent to which the index aligns with the subjective evaluations of experienced orthodontic professionals "validity".

The findings of our study indicated that the Multi-rater Intraclass Correlation Coefficient (ICC) for both the maxillary and mandibular scores provided by the five orthodontic professors was notably high (>0.97), signifying excellent agreement among them. Similarly, for the measurements conducted by the four examiners on the digital models of both maxillary and mandibular arches (>0.99), indicating nearly perfect agreement among the four examiners.

In light of these results, the average scores taken by the five professors on the stone casts were confidently regarded as a gold standard against which the average of

measurements acquired by the four examiners on the digital models were compared to evaluate validity.

To evaluate the validity of Little's Irregularity Index in both the maxillary and mandibular arches, three methods were employed. The first method was, Dahlberg Error (DE) and Relative Dahlberg Error (RDE).

In the results of this study, DE and RDE for both the maxillary and mandibular Little's Irregularity Index exhibited considerable separation between the means of the scores of the five orthodontic professors (gold standard) and the means of the scores of the four orthodontic examiners. This high RDE exceeding 100% in both arches indicated that using Little's Irregularity Index to quantify anterior crowding was not a valid method.

The second approach employed to evaluate the validity of Little's Irregularity Index involved the Bland & Altman Limits of Agreement LOA. The Bland-Altman plot analysis serves as a straight forward method to assess bias between mean differences and estimate an agreement interval[16]. Thus, with a single measurement using method 1, it is expected, with 95% confidence, that the variation from a second measurement obtained through an alternative method will fall within the Limits of Agreement (LOA). It is then up to clinical expertise to determine whether such difference is acceptably small for the two methods to be considered in agreement.

Concerning the maxillary scores, the Limits of Agreement were excessively broad, with a range of $(2.42 + 9.84) = 12.26$. This wide range indicates that the difference

between scores conducted by the examiners could deviate from those provided by the professors by a considerable margin of 12.26, making the index insufficient for accurately describing the real situation. Similarly, for the mandibular scores. The Limits of Agreement were again too extensive, with a range of $(2.67 + 8.17) = 10.84$, suggesting that the difference between scores performed by the examiners could vary from those given by the professors by 10.84. In our investigation, with regards to the maxillary and mandibular scores, a substantial negative difference was observed between the scores provided by the professors and those obtained by the four examiners. This points to a notable bias, suggesting that the values measured by the examiners on the digital models significantly overstate the condition compared to the values indicated by the professors (gold standard).

Moreover, this could be clearly illustrated by the Bland-Altman plot illustrations of the maxillary and mandibular measurements where the Limits of Agreement were too wide with most of the points lying below x axis. Such large negative difference indicated large bias that the scores measured by the examiners on the digital models considerably overestimate the condition compared to the values given by the professors.

It is important to highlight that while the limits of agreement for both maxillary and mandibular scores exhibited wide ranges, the maxillary scores demonstrated a broader range. This suggests a higher likelihood of error in the maxillary arch compared to the mandibular arch.

These findings were in contrast to those by Dr. Little[3] who stated that there is no difference between his gold standards and the examiners scores in the mandibular arch. However, he noted that the index has a tendency to exaggerate cases with considerable irregularity but little arch length shortage.

Ultimately, the third and final approach employed to evaluate the validity of Little's Irregularity Index involved the use of the Concordance Correlation Coefficient (CCC). This statistic quantifies the level of agreement between these two measures of the same variable. Similar to a correlation, CCC ranges from -1 to 1, with perfect agreement represented by a value of 1[17].

In our study, we observed low CCC values between the scores of the five orthodontic professors and the scores of the four examiners for both the maxillary (CCC=0.418) and mandibular (CCC=0.534) arches respectively, indicating a lack of agreement. This could also be demonstrated on the scatter plots of the maxillary and mandibular scores, where almost all points were far above the line of equality indicating poor agreement between the scores. Moreover, all the points being above the line of equality indicated that the values measured by the examiners on the digital models considerably overestimate the condition compared to the values given by the professors.

It is noted again, that the outcomes of this test indicated a comparatively poorer agreement in maxillary measurements than in mandibular measurements.

To the best of our knowledge, apart from Dr. Little's original paper[3], there have

been no other studies examining the validity of Little's Irregularity Index. In his research, Dr. Little[3] confirmed the validity of his index, a conclusion that sharply contradicts the outcomes of our study. It is important to mention that we employed a methodology similar in concept. However, our study employed a sample size four times larger.

Finally, validity stands as a fundamental concept denoting the degree to which a given method accurately measures its intended construct[18]. It is commonly acknowledged that high reliability serves as a key indicator of validity. When a measurement exhibits consistency and stability over repeated trials, it tends to suggest that the method is likely to be valid in capturing the essence of what it aims to measure. However, it is important to recognize that reliability alone does not ensure validity. While a reliable test provides a sense of confidence in its results, it does not guarantee that these result accurately represent the underlying reality or phenomenon being studied. Hence, it becomes apparent that validity encompasses a broader spectrum of considerations beyond mere reliability, including factors such as the appropriateness of the measurement tools, the relevance of the constructs being assessed, and the extent to which the results align with theoretical expectations.

Conclusion:

- 1. The validity of Little's Irregularity Index in the maxillary and mandibular arches could not be confirmed.**
- 2. The outcomes of all tests to assess the validity of Little's Irregularity Index indicated a comparatively poorer agreement in maxillary scores than in mandibular scores.**

Recommendations:

- 1. Further studies are needed to assess the validity of Little's Irregularity Index.**
- 2. Further studies for introducing a new formula that could overcome the limitations of Little's Irregularity Index.**
- 3. Comparison between all available indices that quantify anterior crowding.**

References:

1. El-Attar H, Salama R, Hussien A. Prevalence of malocclusion and common occlusal traits among adolescent school children in Dakahliya, Egypt and Makkah, Saudi Arabia (A COMPARATIVE STUDY). *Egyptian Orthodontic Journal*. 2020;57:1-9.
2. Proffit WR, FHJMLJ. Prevalence of malocclusion and orthodontic treatment need in the United States: estimates from the NHANES III survey. *Int J Adult Orthodon Orthognath Surg*. 1998;13(2):97-106.
3. Little RM. The Irregularity Index: A quantitative score of mandibular anterior alignment. *Am J Orthod*. 1975;68(5):554-563.
4. Pandis N, Polychronopoulou A, Eliades T. Active or passive self-ligating brackets? A

randomized controlled trial of comparative efficiency in resolving maxillary anterior crowding in adolescents. *American journal of orthodontics and dentofacial orthopedics* . 2010;137(1):12.e1-12.e6.

5. Megat Abdul Wahab R, Idris H, Yacob H, Zainal Ariffin SH. Comparison of self- and conventional-ligating brackets in the alignment stage. *Eur J Orthod*. 2012;34(2):176-181.

6. Thickett E, Power S. A randomized clinical trial of thermoplastic retainer wear. *Eur J Orthod*. 2010;32(1):1-5.

7. Rowland H, Hichens L, Williams A, et al. The effectiveness of Hawley and vacuum-formed retainers: A single-center randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics*. 2007;132(6):730-737.

8. Destang DL, Kerr WJS. Maxillary retention: is longer better? *Eur J Orthod*. 2003;25(1):65-69.

9. Sadowsky C, Schneider BJ, BeGole EA, Tahir E. Long-term stability after orthodontic treatment: Nonextraction with prolonged retention. *American Journal of Orthodontics and Dentofacial Orthopedics*. 1994;106(3):243-249.

10. Stevens D.R. FMC, NB, RDW, HG, MPW. 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 Dentofac Orthop* . 2006;129:794-803.

11. Fleming PS, Marinho V, Johal A. Orthodontic measurements on digital study

models compared with plaster models: a systematic review. *undefined*. 2011;14(1):1-16. doi:10.1111/J.1601-6343.2010.01503.X

12. Sonis A. Comparison of intraoral 3D scanning and conventional impressions for fabrication of orthodontic retainers. *J Clin Orthod*. Published online January 1, 2010. Accessed November 23, 2022. https://www.academia.edu/31741327/Comparison_of_intraoral_3D_scanning_and_conventional_impressions_for_fabrication_of_orthodontic_retainers

13. Roy W Goonewardene; Mithran S Goonewardene; John M Razza; Kevin Murray. Accuracy and validity of space analysis and irregularity index measurements using digital models. *Aust Orthod J*. 2008;24(2):83-90.

14. Burns A, Dowling AH, Garvey TM, Fleming GJP. The reliability of Little's

Irregularity Index for the upper dental arch using three dimensional (3D) digital models. *J Dent*. 2014;42(10):1320-1326. doi:10.1016/j.jdent.2014.07.012

15. Versi E. "Gold standard" is an appropriate term? *BMJ* . 1992;305(6846):187-187.

16. Bland J M & Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. Published online 1986.

17. Lawrence I-Kuei Lin. A Concordance Correlation Coefficient to Evaluate Reproducibility *Biometrics* Vol. 45, No. 1 (Mar., 1989), pp. 255-268. *International Biometric Society*. 1989;45(1):255-268.

18. Fiona F. Middleton. Reliability vs. Validity in Research | Difference, Types and Examples. *scribber*. Published online January 30, 2023.