Original Article

Diagnostic performance of ClinCheck, Dolphin Imaging, and 3D Slicer software for Bolton discrepancy analysis

Thalita Teixeira Santana[a]; Flávio Copello[b]; Guido Artemio Marañón-Vásquez[c]; Lincoln Issamu Nojima[d]; Eduardo Franzotti Sant'Anna[d]

ABSTRACT

Objectives: To evaluate the diagnostic performance of ClinCheck, Dolphin Imaging orthodontic software, and 3D Slicer for the analysis of Bolton discrepancy (BD).

Materials and Methods: Fifty-five pairs of early-stage digital models of patients treated with Invisalign were printed to measure the BD by manual method with a digital caliper (gold standard). The discrepancy values calculated by ClinCheck were obtained. In addition, the sample STL files were measured using Dolphin Imaging and 3D Slicer software to obtain BD values. To assess reliability, precision, and accuracy of the methods, intraclass correlation coefficients (ICCs), Dahlberg's formula, paired *t*-tests, and the Bland-Altman method were used, respectively. Repeated-measures analysis of variance with Bonferroni post hoc test was used to assess the difference between groups.

Results: The three methods showed reliable measurements (ICC $\geq$ 0.7), with the values of anterior Bolton slightly higher than overall Bolton. Measurements for the anterior Bolton showed higher precision (Dahlberg's formula 0.65, 0.70, and 0.55) than those for the overall Bolton. For anterior Bolton, only the measurements obtained by ClinCheck and Dolphin Imaging were accurate ($P > .05$, no proportion bias), while for overall Bolton, all groups had a significant difference. The Bland-Altman plots demonstrated no consistency for anterior Bolton measurements when 3D Slicer was used and for the overall Bolton.

Conclusions: ClinCheck and Dolphin Imaging showed accuracy to quantify anterior BD. For the overall Bolton measurements, ClinCheck showed a statistical difference from the manual assessment but without relevant clinical significance. (*Angle Orthod*. 2025;95:51–56.)

KEY WORDS: Digital orthodontics; Clear aligners

INTRODUCTION

To produce Invisalign clear aligners, there is a digital workflow in which the interaction between orthodontists and Align Technology takes place through a virtual tool, ClinCheck software (Align Technology, San Jose, Calif).[1] This platform displays virtual patient models with predictions of sequential treatment steps to be performed with the aligners. Therefore, many studies have been carried out to establish validity between predictability and the clinical results obtained after treatment completion, while Invisalign becomes increasingly used widely among professionals.[2–5]

[a] Graduate Student, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

[b] Assistant Professor, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Maryland Baltimore, Maryland, USA.

[c] Assistant Professor, Department of Pediatric Dentistry, School of Dentistry of Ribeirao Preto, University of Sao Paulo.

[d] Associate Professor, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

Corresponding author: Dr Eduardo Franzotti Sant'Anna, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, Universidade Federal do Rio de Janeiro, Av. Professor Rodolpho Paulo Rocco, 325 Ilha do Fundão, Rio de Janeiro, RJ 21941-590, Brazil (e-mail: eduardo.franzotti@gmail.com)

Accepted: August 2024. Submitted: February 2024.
Published Online: September 25, 2024

Among the information provided by ClinCheck, the Bolton discrepancy (BD) is displayed. This analysis identifies the proportions of tooth size between the upper and lower arches[6] and is directly related to the amount of interproximal reduction (IPR) or resin additions that are indicated to achieve ideal occlusion. This analysis can be performed digitally using virtual three-dimensional (3D) models, or it can be assessed by the conventional technique initially described by Bolton, using physical models and calipers.[7,8] For clear aligners, correct measurement of mesiodistal diameter of the teeth is even more important because it provides for the creation of adequate space to achieve desired tooth movements.[9] Also, it is related to the correct indication of IPR to obtain space when necessary,[10,11] minimizing iatrogenic effects in treatment with clear aligners.

Information is lacking on how the BD analysis is obtained through the ClinCheck tool. Therefore, the aim of the present study was to evaluate the diagnostic performance of ClinCheck as well as Dolphin Imaging (Dolphin Imaging and Management Solutions, Chatsworth, Calif) and 3D Slicer software (Kitware, Clifton Park, NY) to assess BD.

MATERIALS AND METHODS

The study was approved by the Research Ethics Committee of the Clementino Fraga Filho University Hospital (CEP-HUCFF) of the Federal University of Rio de Janeiro (UFRJ) under No. 5.202.503.

Sample Selection

A priori sample size calculations were performed using data from a pilot study (n = 10), following two approaches: (1) calculation for comparison of anterior and overall BDs in millimeters between two groups using the paired *t*-test (conventional evaluation vs ClinCheck), and (2) calculation for comparison of the same outcomes among three groups using within-subjects repeated-measures analysis of variance (ANOVA; ClinCheck vs Dolphin Imaging vs 3D Slicer). The estimation that resulted in a larger sample size was used for the present study (conventional evaluation vs. ClinCheck for overall Bolton; mean difference = −0.520 ± 1.265 mm). The parameters established for the calculation were effect size dz = −0.411, a = 0.05, and power = 0.8. The estimation resulted in a total of 49 models, rounded to 55, considering possible losses during the study. Sample size estimates were performed with G*Power software version 3.1.9.6 (Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany).

The sample comprised 55 patients aged 18–60 years, with no sex predilection, who were treated with Invisalign. Selection was based on the following inclusion criteria: complete permanent dentition; Class I, II, or III molar relationship; and lower crowding of up to 6 mm. Cases that had teeth with anatomic anomalies, occlusal reduction, as well as individuals with other orofacial malformations were excluded. After selection, the dental arches were scanned using an iTero Element 2 scanner.

The STL files in were prepared using the Meshmixer software version 3.5.474 (Autodesk, Inc., San Rafael, Calif). The digital models were printed with Anycubic Photon Mono SE 3D printer (Shenzhen, China), and the BD analysis was performed.

Bolton Discrepancy Assessment

Measurements of the mesiodistal diameter of the teeth were performed using a digital caliper (Mitutoyo©, Suzano, São Paulo, Brazil) on the 3D-printed models by the same previously calibrated operator. Teeth were measured from the first molar on one side to the first molar on the opposite side, ensuring the 12 teeth of each dental arch were measured. Then the BD analysis[6] was performed for the anterior region (considering incisors and canines: 3-3) and overall area (first molar to first molar: 6-6), in millimeters.

For the ClinCheck tool, the values and location referring to the BD were identified in the software and tabulated.

Finally, the BD analysis was performed using the Dolphin Imaging orthodontic software version 11.95 Premium and 3D Slicer version 5.1.0. For each software program, the STL digital models were exported, and the same Bolton measurement methodology was performed.

Method Error

Twenty percent of the sample was randomly selected to assess the method error. Tooth measurements and subsequent calculation of the anterior and overall BDs were performed on the selected models two times with an interval of at least 2 weeks. The intraclass correlation coefficient (ICC) and the Bland-Altman method (proportion bias) were used to assess the reliability and accuracy of measurements reported in millimeters.

Data Analysis

The distribution and normality of the data for the anterior and overall BDs (in millimeters) were evaluated through the Shapiro-Wilk test, histograms, and Q-Q plots.

ICC was used to assess reliability by evaluating the degree of agreement between measurements at two times or using two methods. Random and systematic

errors were assessed to determine validity of the methods. Dahlberg's formula was used to measure the random error (precision), that is, how closely the values from two measurements were. The paired *t*-test was used to detect significant differences between measurements using different methods. The Bland-Altman method was also used since, in addition to determining the random error, it assessed the presence of systematic error by estimating the proportion bias with its corresponding 95% confidence interval. Assessments of the BD using the conventional method were used as the reference (gold standard) for all evaluations.

Additionally, the differences in discrepancy values in millimeters between each digital method and the conventional evaluation were independently calculated. Subsequently, these differences were compared using repeated-measures ANOVA. To detect which groups were different, post hoc pairwise comparisons were performed using the Bonferroni test.

All analyses were performed in Jamovi software version 2.0 (Sydney, Australia), adopting a significance level of 5%.

RESULTS

Method Error

Measurements of BDs in millimeters using the conventional method, Dolphin Imaging software, and 3D Slicer were reliable and accurate (Table 1). The ICC for repeated measures of the anterior and overall Bolton ranged from 0.86 to 0.96 and from 0.80 to 0.91, respectively.

Main Findings

Results of the reliability, precision, and accuracy assessments of the measurements of BD in millimeters are described in Table 2. The three methods showed reliable measurements (ICC $\geq$ 0.7). Those obtained for the anterior Bolton analysis were slightly more reliable than those obtained for the overall Bolton. Similarly, measurements for the anterior Bolton showed higher precision compared with the overall Bolton. Although the methods showed good ICC values (high correlation between assessments), only the measurements obtained for the anterior Bolton using the ClinCheck tool and Dolphin Imaging were accurate (paired *t*-test $P > .05$; no proportion bias). These methods had a precision of 0.65 mm and 0.70 mm, respectively. All the other measurements were inaccurate, evidencing proportion bias.

The Bland-Altman plots showed that the measures for the anterior Bolton in 3D Slicer and for the overall Bolton using the three software programs were scattered widely, above zero, demonstrating no consistency with measurements of the conventional method and tending either to (1) decrease the excess in the mandibular arch; (2) increase the excess in the maxillary arch; or (3) in case of minor discrepancies, indicate that the excess was in the upper arch when it was in the lower arch (Figure 1). When the mean differences between each of the tested methods and the conventional method were compared, no significant differences were observed for the anterior Bolton between the ClinCheck and Dolphin Imaging assessments ($P > .999$), while measurements using 3D Slicer software did show differences with both ($P < .01$; Figure 2). When the overall Bolton was analyzed, no significant difference was observed among the groups (Figure 2).

DISCUSSION

Success with Invisalign occurs when the occlusal results predicted by ClinCheck are consistent with those obtained during actual treatment.[2] Thus, some parameters of this tool have been studied, such as the predictability of overbite correction,[12,13] transverse changes in the arch,[14,15] accuracy of measurements performed in the anterior region such as arch length, intercanine distance, overjet, overbite, dental midline shift,[1] and accuracy of tooth movements predicted by this system.[5,16,17] However, information is lacking on the validity of BD data indicated by ClinCheck, and considering its importance to diagnosis and orthodontic planning procedures, the present study was undertaken.

The gold standard reference for comparison was measurements performed on printed models using stereolithographic technology since the prototype models have demonstrated similar precision and clinical application to plaster models.[18,19] The prototypes

Table 1. Method Error for the Measurement of Bolton Discrepancy in Millimeters[a]

Method	Estimates	Anterior Bolton	Overall Bolton
Conventional evaluation	ICC	0.88	0.82
	BA bias—Est (95% CI)	0.25 (−0.23, 0.72)	0.07 (−0.54, 0.68)
Dolphin Imaging®	ICC	0.96	0.91
	BA bias—Est (95% CI)	0.20 (−0.09, 0.49)	−0.15 (−0.50, 0.20)
3D Slicer	ICC	0.86	0.80
	BA bias—Est (95% CI)	−0.29 (−0.71, 0.13)	0.36 (−0.38, 1.10)

[a] ICC indicates intraclass correlation coefficient; BA, Bland-Altman; Est, estimate; and CI, confidence interval.

Table 2. Reliability, Precision, and Accuracy of the Methods to Measure the Bolton Discrepancy in Millimeters[a]

Method	Estimates	Anterior Bolton	Overall Bolton
ClinCheck	ICC	0.75	0.71
	Dahlberg's formula	0.65	1.12
	Mean difference ± SD	−0.11 ± 0.93	0.84 ± 1.35
	Paired *t*-test	0.389	<0.001
	BA bias—Est (95% CI)	−0.11 (−0.36, 0.14)	0.84 (0.48, 1.21)
Dolphin Imaging®	ICC	0.75	0.70
	Dahlberg's formula	0.70	1.13
	Mean difference ± SD	−0.17 ± 0.99	0.80 ± 1.40
	Paired *t*-test	0.216	<0.001
	BA bias—Est (95% CI)	−0.17 (−0.44, 0.10)	0.80 (0.43, 1.18)
3D Slicer	ICC	0.87	0.77
	Dahlberg's formula	0.55	0.91
	Mean difference ± SD	0.35 ± 0.70	0.45 ± 1.22
	Paired *t*-test	<0.001	0.008
	BA bias—Est (95% CI)	0.35 (0.16, 0.54)	0.45 (0.13, 0.78)

[a] ICC indicates intraclass correlation coefficient; BA, Bland-Altman; Est, estimate, and CI, confidence interval.

were printed with a layer thickness of 50 μm for accuracy.[20] The models were arranged at 0° to the printing platform, minimizing the expansion effect in the terminal region of the arch.[21]

The results indicated that, although the measurements obtained were reliable, only the ClinCheck tool and Dolphin Imaging software showed accuracy for BD measurements performed in the anterior region. Authors of two previous studies investigated the diagnostic performance of ClinCheck compared with manual methods.[9,22]

Shailendran et al[9] performed segmentation of the teeth to measure their mesiodistal diameter and assess BD. The findings indicated that ClinCheck tended to underestimate tooth width measurements by an average of 0.36 mm, while for molars, the average was 0.9 mm. As ClinCheck accuracy decreased as the measurements became farther posterior in the arch, a greater probability of underestimation was found in these areas, while more accurate values were found in the anteroinferior region,[9] a pattern that was also observed in the current study.

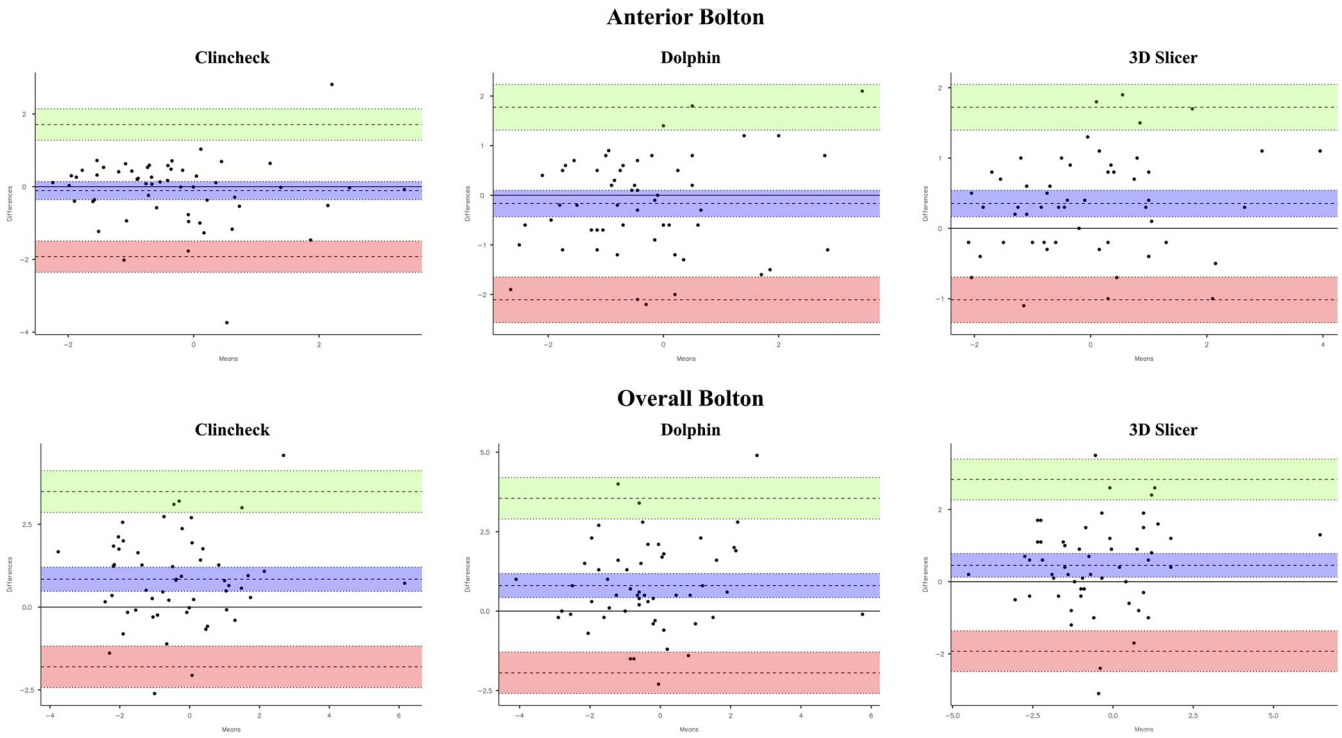

Figure 1. Distribution of measurements after calculating the anterior and overall Bolton discrepancy (BD) by the three methods.

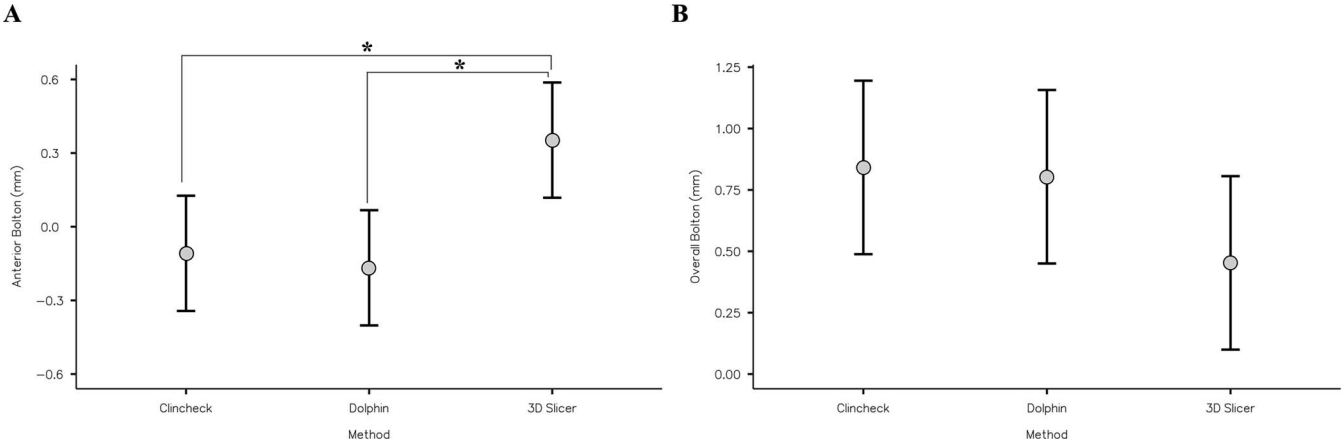

Figure 2. Comparison between the mean differences found between measurements using each of the three tested software methods and the conventional method for anterior and overall Bolton discrepancy (BD).

Martin et al.[22] compared the values indicated by ClinCheck software with measurements performed on plaster models and found higher values provided by the software, with the largest measurement errors also in the molar region. The measurement methodology was like that of the present study, in which the actual position of the teeth in the arch was considered, along with crowding, inclination, and existing dental rotation. Such factors present increased difficulty for accurate measurement of BD[23] but also represent the classic evaluations performed on plaster models, developed by Bolton himself in 1962.[6]

Concerning overall Bolton assessment, ClinCheck, Dolphin Imaging, and 3D Slicer showed moderate reliability, but the values were significantly different than the measurements performed on the printed models, demonstrating that the measurements performed were not accurate. However, the variation was less than 1.5 mm which, according to Proffit et al,[24] would have little clinical implication.

The current findings were like those of previous studies, in which authors observed satisfactory agreement between the Bolton analysis for anterior and overall proportions when comparing measurements in 3D models using software and plaster models. In those studies, the digital models were obtained by digitalizing plaster models and not with intraoral scanning, as performed in the present study.[25–28] Regarding ClinCheck, the correlation found in the study by Shailendran et al[9] was moderate, with an ICC of 0.83 for overall Bolton results, like that found in the present study.

Some explanation may exist as to why the reproducibility of overall Bolton measurements was not better. Acquisition of digital measures may be easier[29] and require less time than manual methods,[25,26] but divergence or imprecision may occur in marking the reference points of the measurements, even when performed by the same operator over a short time interval.[23,27] This limitation may be related to the 3D nature of the models which is viewed two-dimensionally,[29] making it difficult to identify points, axes, and planes by the software.[23,26,30] For Dolphin Imaging, 3D features are lost since the software uses a static image of the dental arch in the occlusal view to mark the points.

Another factor would be the difficulty of scanning to faithfully reproduce posterior teeth due to patient-related factors, such as salivation.[31] The software fills in such regions through a mathematical algorithm, leading to changes in tooth shape that generate measurement inaccuracies.[1]

Finally, inherent difficulties may exist in measuring posterior regions using scanning technology. Authors of studies comparing sizes in digital vs conventional models have indicated that, despite having high reproducibility, digital models tended to decrease the intermolar width and increase the palatal depth.[23] This systematic error in the reproduction of the molar surface was also described by Zilberman et al.[30] For the ClinCheck tool, authors of studies indicated that the software overestimated the alignment and posterior occlusal relationships obtained after treatment.[2] Further studies are needed regarding the dental information provided by ClinCheck, especially pertaining to the posterior region of the arch.

Considering common clinical application, evidence of the accuracy of the anterior Bolton values using ClinCheck and Dolphin Imaging is encouraging since intervention in this region is more frequent, either by IPR in cases of excess or by addition of resin or by restoration to modify atypical shape of teeth to achieve proper occlusion. However, in the posterior region, IPR is less often indicated because it may increase the chance of food impaction and caries risk but may still be indicated to alter abnormal shape. It is recommended that some caution may be appropriate when assessing and interpreting BD data obtained through orthodontic software.

CONCLUSIONS

- ClinCheck and Dolphin Imaging software demonstrated precision and accuracy for anterior BD analysis measurements, demonstrating validity to be used during orthodontic diagnosis.
- During analysis of overall Bolton relationships, ClinCheck, Dolphin Imaging, and 3D Slicer showed more variability and tended to decrease excess in the lower arch and increase excess tooth size in the upper arch. However, the variations shown did not have clinically relevant implications.

ACKNOWLEDGMENTS

This study was financed in part by the agency Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES) Financing Code 001.

REFERENCES

1. Krieger E, Seiferth J, Marinello I, et al. Invisalign® treatment in the anterior region: were the predicted tooth movements achieved? *J Orofac Orthop*. 2012;73:365–376.
2. Morton J, Derakhshan M, Kaza S, Li C. Design of the Invisalign system performance. *Semin Orthod*. 2017;23:3–11.
3. Buschang PH, Ross M, Shaw SG, Crosby D, Campbell PM. Predicted and actual end-of-treatment occlusion produced with aligner therapy. *Angle Orthod*. 2015;85:723–727.
4. Solano-Mendoza B, Sonnemberg B, Solano-Reina E, Iglesias-Linares A. How effective is the Invisalign® system in expansion movement with Ex30′ aligners? *Clin Oral Invest*. 2016;21: 1475–1484.
5. Haouili N, Kravitz ND, Vaid NR, Ferguson DJ, Makki L. Has Invisalign improved? A prospective follow-up study on the efficacy of tooth movement with Invisalign. *Am J Orthod Dentofacial Orthop*. 2020;158:420–425.
6. Bolton WA. The clinical application of a tooth-size analysis. *Am J Orthod*. 1962;48:504–529.
7. Santoro M, Ayoub ME, Pardi VA, Cangialosi TJ. Mesiodistal crown dimensions and tooth size discrepancy of the permanent dentition of Dominican Americans. *Angle Orthod.* 2000;70:303–307.
8. Akyalçin S, Doğan S, Dinçer B, Erdinc AME, Öncağ G. Bolton tooth size discrepancies in skeletal Class I individuals presenting with different dental angle classifications. *Angle Orthod*. 2006;76:637–643.
9. Shailendran A, Weir T, Freer E, Kerr B. Accuracy and reliability of tooth widths and Bolton ratios measured by ClinCheck Pro. *Am J Orthod Dentofacial Orthop*. 2022;161:65–73.
10. Johner AM, Pandis N, Dudic A, Kiliaridis S. Quantitative comparison of 3 enamel-stripping devices in vitro: how precisely can we strip teeth? *Am J Orthod Dentofacial Orthop*. 2013;143:168–172.
11. Kalemaj Z, Levrini L. Quantitative evaluation of implemented interproximal enamel reduction during aligner therapy: a prospective observational study. *Angle Orthod*. 2021;91:61–66.
12. Krieger E, Seiferth J, Saric I, Jung BA, Wehrbein H. Accuracy of Invisalign® treatments in the anterior tooth region. *J Orofac Ortho.* 2011;72:141–149.
13. Blundell HL, Weir T, Kerr B, Freer E. Predictability of overbite control with the Invisalign appliance. *Am J Orthod Dentofacial Orthop*. 2021;160:725–731.
14. Houle JP, Piedade L, Todescan R Jr, Pinheiro FHL. The predictability of transverse changes with Invisalign. *Angle Orthod*. 2017;87:19–24.
15. Riede U, Wai S, Neururer S, et al. Maxillary expansion or contraction and occlusal contact adjustment: effectiveness of current aligner treatment. *Clin Oral Invest*. 2021;25:4671–4679.
16. Charalampakis O, Iliadi A, Ueno H, Oliver DR, Kim KB. Accuracy of clear aligners: a retrospective study of patients who needed refinement. *Am J Orthod Dentofacial Orthop.* 2018;154:47–54.
17. Papadimitriou A, Mousoulea S, Gkantidis N, Kloukos D. Clinical effectiveness of Invisalign® orthodontic treatment: a systematic review. *Prog Orthod*. 2018;19:1–24.
18. Aly P, Mohsen C. Comparison of the accuracy of three-dimensional printed casts, digital, and conventional casts: an in vitro study. *Eur J Dent*. 2020;14:189–193.
19. Ellakany P, Alharbi F, El Tantawi M, Mohsen C. Evaluation of the accuracy of digital and 3D-printed casts compared with conventional stone casts. *J Prosthet Dent*. 2020;127:438–444.
20. Zhang ZC, Li PL, Chu FT, Shen G. Influence of the three-dimensional printing technique and printing layer thickness on model accuracy. *J Orofac Orthop.* 2019;80:194–204.
21. Ko J, Bloomstein RD, Briss D, et al. Effect of build angle and layer height on the accuracy of 3-dimensional printed dental models. *Am J Orthod Dentofacial Orthop*. 2021;160:451–458.
22. Martin MA, Lipani E, Martinez LB, Lorenzo AA, Aiuto R, Garcovich D. Reliability of tooth width measurements delivered by the Clin-Check Pro 6.0 software on digital casts: a cross-sectional study. *Int J Environ Res Public Health*. 2022;19:3581.
23. Sousa MVS, Vasconcelos EC, Janson G, Garib D, Pinzan A. Accuracy and reproducibility of 3-dimensional digital model measurements. *Am J Orthod Dentofacial Orthop*. 2012;142: 269–273.
24. Proffit WR, Fields HW, Sarver DM. *Ortodontia Contemporânea*. 5° Edição ed. St Louis, MO, USA: Elsevier/Mosby; 2012.
25. Tomassetti JJ, Taloumis IJ, Denny JM, et al. A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. *Angle Orthod*. 2001;71:351–357.
26. Mullen, SR, Martin CA, Ngan P, Gladwin M. Accuracy of space analysis with emodels and plaster models. *Am J Orthod Dentofacial Orthop*. 2007;132:346–352.
27. Brandão MM, Sobral MC, Vogel CJ. Reliability of Bolton analysis evaluation in tridimensional virtual models. *Dental Press J Orthod*. 2015;20:72–77.
28. Amuk NG, Karsli E, Kurt G. Comparison of dental measurements between conventional plaster models, digital models obtained by impression scanning and plaster model scanning. *Int Orthod*. 2019;17:151–158.
29. Abizadeh N, Moles DR, O'Neill J, Noar JH. Digital versus plaster study models: how accurate and reproducible are they? *J Orthod*. 2012;39:151–159.
30. Zilberman O, Huggare J, Parikakis KA. Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. *Angle Orthod*. 2003;73:301–306.
31. Flügge TV, Schlager S, Nelson K, Nahles S, Metzger MC. Precision of intraoral digital dental impressions with iTero and extraoral digitization with the iTero and a model scanner. *Am J Orthod Dentofacial Orthop*. 2013;144:471–478.