

Optimizing Orthodontic Appliances Efficiency with Remote Dental Monitoring

2021 Grants

Dr Mohammed Elnagar

melnagar@uic.edu
O: 312-292-8454

FollowUp Form

Award Information

In an attempt to make things a little easier for the reviewer who will read this report, please consider these two questions before this is sent for review:

- Is this an example of your very best work, in that it provides sufficient explanation and justification, and is something otherwise worthy of publication? (We do publish the Final Report on our website, so this does need to be complete and polished.)*
- Does this Final Report provide the level of detail, etc. that you would expect, if you were the reviewer?*

Title of Project*

Optimizing Orthodontic Appliances Efficiency with Remote Dental Monitoring

Award Type

Orthodontic Faculty Development Fellowship Award (OFDFA)

Period of AAOF Support

July 1, 2021 through June 30, 2023

Institution

University of Illinois, Chicago

Names of principal advisor(s) / mentor(s), co-investigator(s) and consultant(s)

Veerasathpurush Allareddy, Budi Kusnoto and Phimon Atsawasuan

Amount of Funding

\$20,000.00

Abstract

(add specific directions for each type here)

Respond to the following questions:

Detailed results and inferences:*

If the work has been published, please attach a pdf of manuscript below by clicking "Upload a file".

OR

Use the text box below to describe in detail the results of your study. The intent is to share the knowledge you have generated with the AAOF and orthodontic community specifically and other who may benefit from your study. Table, Figures, Statistical Analysis, and interpretation of results should also be attached by clicking "Upload a file".

final report OFDFA 2021.pdf

Results

objective 1 and 2

535 observation points were examined for the general detection of clear aligner and oral hygiene parameters. The estimated accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the notifications generated by DM on the general detection (which implies the AI can identify if a clinical condition is present or absent in the image) of clinical findings related to clear aligner parameters are shown in Table II and Table III.

For generalized detection of clear aligner parameters, the accuracy, sensitivity, specificity, PPV and NPV of the DM artificial intelligence image analysis software was high (>90%) across most evaluated parameters. Slight unseat was moderately accurate (80.7%). Noticeable unseat (80.1%) and slight unseat (76.8%) were moderately sensitive. PPV was moderate for aligner damage (86.4%) and NPV was low for slight unseat (55.4%).

For generalized detection of oral hygiene parameters, the accuracy, specificity and NPV were high (>90%) for most parameters. The sensitivity was moderate (70-90%) to low (<70%) for most parameters. For gingival recession, the sensitivity as null because the AI incorrectly recognized a crown margin as "recession". The PPV was moderate-to-low for most parameters except for noticeable gingivitis (100%), black triangles (100%) and mucosa irregularity (100%).

After all the generalized detection timepoints were examined, the primary investigator further identified the accuracy, sensitivity, specificity, PPV and NPV of the AI's localized detection capabilities (which implies the AI can identify the exact tooth number where a clinical condition is present or absent). There are variable numbers of notification samples for localized detection due to the discrepancies of identified teeth between the gold standard orthodontist and the Dental Monitoring AI algorithm. The results are shown in Table IV and V.

For localized detection of clear aligner parameters, the accuracy, sensitivity, specificity, PPV and NPV was high (>90%) for most parameters. Noticeable unseat demonstrated moderate accuracy (83.2%), sensitivity (70.4%), and NPV (76.8%). Slight unseat demonstrated low accuracy (54.2%), sensitivity (54.7%), specificity (51.5%), and NPV (17.9%).

For localized detection of oral hygiene parameters, the accuracy, specificity, PPV and NPV was high (>90%) for most parameters. Sensitivity was low (<70%) for most parameters, except for noticeable gingivitis (100%). Slight plaque/food and slight gingivitis demonstrated low (<70%) validity values across all parameters. For gingival recession, the sensitivity as null because the AI incorrectly recognized a crown margin as "recession".

objective 3

(13a) (14a) Sixty-three participants were recruited for this study during July 2022-December 2023. (7a) The sample size was determined based on the number of participants who had been in treatment for a minimum of 6 months at the start of study recruitment. (7b) There were no interim analyses provided for the sample size.

A total of thirty-seven patients were included in this study. (13b) (14b) No subjects were excluded after randomization and the larger study is still recruiting more patients. (15) Descriptive statistics of the three groups were calculated for sex and age (Table I). There was no statistically significant mean difference between the three groups ($p=.134$). Twelve subjects were included in the Group 1 (DM group; 3 males and 9

females; mean age 33.34 ± 12.23 years), fourteen subjects for Group 2 (7-day group; 2 males and 12 females; mean age 45.70 ± 18.03 years), and eleven subjects for Group 3 (14-day group; 4 males and 7 females; mean age 38.26 ± 15.87 years). The majority of subjects included in this study were female (75.7%).

ABO Discrepancy Index (DI) Score Results

(16) (17a) (17b) (18) It is important to note that significant differences between some parameters of the DI score may not necessarily translate to clinical significance. For instance, when grading a case with the ABO discrepancy index (DI), overbite is scored in the following categories: >1 to ≤ 3mm = 0 points; >3 to ≤ 5mm = 2 points; >5 to ≤ 7mm = 3 points; Impinging (100%) = 5 points. Therefore, on average, there must be a 1.67 difference in points to change the overbite score. In this study, there was a 1.000 difference in overbite points and therefore is not considered to be clinically significant. A similar logic can be applied for all discrepancy index parameters, as described in the ABO instructions for calculating ABO Discrepancy Index Score 43

A comparison of ABO DI score change between the three groups T0 and T1 after 6-months of clear aligner treatment are shown in Table II. The paired t-test for group 2 (7-day) indicated two parameters that were statistically significant from T0 to T1. Overbite (mean difference = -0.571, p = .026) indicates that this parameter had a statistically significant decrease from T0 to T1, however this was not clinically significant. Lateral open bite (mean difference = 0.857, p = .034) indicates that this parameter had a statistically significant increase from T0 to T1, however this was not clinically significant. The paired t-test for group 3 (14-day) indicated that crowding (mean difference = -0.636, p = .011) demonstrated a statistically significant change from T0 to T1, however this was not clinically significant.

After comparing ABO DI score changes from T0 (pre-treatment) to T1 (6-months of treatment) across the three groups, the overall DI score changes are not statistically significant between any groups and suggests similar effectiveness across all groups. However, two parameters, overbite (p=.055) and lateral open bite (p=.008), showed statistically significant differences from T0 to T1, but they were not clinically significant.

ABO Objective Grading System (OGS) Results

By following similar logic as discussed above and carefully following instructions of the ABO OGS, it is important to note that statistically significant differences between some parameters of the OGS score may not necessarily translate to clinical significance.

A comparison of ABO OGS score change between the three groups T0 and T1 after 6-months of clear aligner treatment are shown in Table III. The paired t-test for group 1 (DM) indicated three parameters that were statistically significant from T0 to T1. Overall OGS score (mean difference = -5.583, p = .012) indicates that this parameter had a statistically significant decrease from T0 to T1. Alignment (mean difference = -7.583, p = .002) indicates that this parameter had a statistically significant decrease from T0 to T1. Occlusal contacts (mean difference = 5.000, p = .014) indicates that this parameter had a statistically significant increase from T0 to T1. All three parameters are considered clinically significant due to the instructions followed for ABO OGS.

The paired t-test for group 2 (7-day) indicated three parameters that were statistically significant from T0 to T1. Overall OGS score (mean difference = -8.643, p = .009) indicates that this parameter had a statistically significant decrease from T0 to T1. Alignment (mean difference = -12.857, p = <.001) indicates that this parameter had a statistically significant decrease from T0 to T1. Occlusal contacts (mean difference = 4.857, p = .008) indicates that this parameter had a statistically significant increase from T0 to T1. All three parameters are considered clinically significant due to the instructions followed for ABO OGS.

The paired t-test for group 3 (14-day) indicated that alignment (mean difference = -7.727, p = .001) had a statistically significant decrease from T0 to T1. This is also considered a clinically significant difference.

After comparing ABO OGS score changes from T0 (pre-treatment) to T1 (6-months of treatment) across the three groups, the overall OGS score changes are not statistically significant between any groups and suggests similar effectiveness across all groups. However, alignment (p=.043) showed statistically significant differences from T0 to T1, but it was not significant after further testing.

Treatment Efficiency Results

A comparison of treatment efficiency parameters between the three groups within 6 months of clear aligner treatment are shown in Table IV. The total visits within 6 months, number of emergency/unscheduled visits in 6 months, and number of initial trays showed no significant difference between the three groups, indicating similar conditions across these parameters. The number of scheduled office visits in 6 months ($p=.036$) and tray number at 6 months ($p=.006$) showed statistically significant differences between the three groups. After further examination, the number of scheduled office visits in 6 months (mean difference = -0.988, $p=.032$) demonstrated statistically significant differences between group 1 (DM) and group 2 (7-day), indicating that the DM group may have a lower number of scheduled office visits in 6 months of treatment time. The tray number at 6 months (mean difference = 8.857, $p=.005$) showed statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have progressed through their aligners at a faster rate compared to the 14-day group; this result may be expected as the 7-day group had less initial time in each tray compared to the 14-day group. (19) No harm or unintended effects were identified in each group during the study.

Were the original, specific aims of the proposal realized?*

yes

This study compared the Dental Monitoring (DM) artificial intelligence (AI) image analysis performance against the “gold standard” orthodontist in detecting clear aligner issues from DM scans. This study revealed the AI’s high accuracy in detecting clinical conditions (aligner unseats, attachment loss, aligner damage, button loss) in general but challenges in localization of more subtle aligner tracking issues. The findings shed light on the DM AI’s current limitations and emphasized the need for improvement in the AI’s diagnostic capability for assessing clinical conditions.

This study also highlighted the capabilities and challenges that the DM AI software demonstrates in detecting oral hygiene issues in clear aligner treatment. Following a similar pattern as aligner tracking issues, the DM AI image analysis software is highly accurate in detecting the absence or presence of clinical conditions, but often has issues localizing the condition to identify the specific teeth. This suggests that the AI remote monitoring can be utilized during treatment, but close management is required for more subtle hygiene issues.

Lastly, a study was implemented to investigate the impact of DM on orthodontic clear aligner treatment outcomes over the first 6 months of treatment in three groups with different aligner change intervals (DM, 7-days, 14-days). Initial findings indicate no significant differences in treatment effectiveness or efficiency among the groups, suggesting DM’s potential to seamlessly integrate with traditional treatment protocols without compromising patient outcomes. Furthermore, Dental Monitoring may offer benefits such as reduced office visits, highlighting its viability for enhancing treatment monitoring and patient compliance while accommodating individual patient needs.

Were the results published?*

Yes

Have the results of this proposal been presented?*

Yes

To what extent have you used, or how do you intend to use, AAOF funding to further your career?*

Funding from AAOF was crucial for my research project. The award will help me recruit patients to participate in my study and obtain the needed equipment and software for my project. Furthermore, it assisted in the development of my career as an Educator, Clinician, and Scientist

Accounting: Were there any leftover funds?

\$0.00

Published

Citations*

You indicated results have been published. Please list the cited reference/s for publication/s including titles, dates, author or co-authors, journal, issue and page numbers

Snider V, Homsy K, Kusnoto B, Atsawasuwan P, Viana G, Allareddy V, Gajendrareddy P, Elnagar MH. Clinical evaluation of Artificial Intelligence Driven Remote Monitoring technology for assessment of patient oral hygiene during orthodontic treatment. Am J Orthod Dentofacial Orthop. 2024 May;165(5):586-592. doi: 10.1016/j.ajodo.2023.12.008. Epub 2024 Feb 15. PMID: 38363256.

The result was presented in the AAO 2024 annual session

in addition 2, manuscripts ready for submission

Comment: *The AAOF commends you for completing this project, which will contribute to advancing the knowledge base in the rapidly evolving AI space. We encourage you to continue your career development and engagement with the AAOF through grant applications and publishing your results.*

Was AAOF support acknowledged?

If so, please describe:

Yes, AAOF was acknowledged as the funding source during the presentation and publication

Presented

Please list titles, author or co-authors of these presentation/s, year and locations:*

Mohammed H. Elnagat "Embracing the Future: Orthodontics Revolutionized by Artificial Intelligence"

The 2024 (159th) Midwinter Meeting, Chicago Dental Society, Chicago, IL

Mohammed H Elnagar, Veerasathpurush Allareddy, Phimon Atsawasuwana, Budi Kusnoto, Karen Homsy, Vivian Shnyder, Lia Taher "Evaluation of AI-driven remote monitoring technology for tracking tooth movement and reconstruction of 3D models during orthodontic treatment"

The 51st Annual Moyers Symposium (Artificial Intelligence: What Should Clinicians Know?) and 48th Annual International Conference on Craniofacial Research

The University of Michigan, Ann Arbor, MI.

Mohammed H Elnagar "Artificial Intelligence-driven Orthodontic Practice: Welcome to the Future"

Doctor's Session, the 124th Annual Session - American Association of Orthodontists, May 2024, New Orleans, LA.

Was AAOF support acknowledged?

If so, please describe:

yes

Internal Review

Reviewer Comments

Reviewer Status*

Approved

File Attachment Summary

Applicant File Uploads

- final report OFDFA 2021.pdf

Final report for 2021 Orthodontic Faculty Development Fellowship Award

Mohammed H Elnagar

Title of Project:

Award Type: Orthodontic Faculty Development Fellowship Award

Period of AAOF Support: July 1, 2021 through June 30, 2024

Institution: University of Illinois, Chicago

Amount of Funding: 20,000

Aim and Objectives

The primary aim of this study is to evaluate the efficacy of implementing remote Dental Monitoring™ in orthodontic clear aligner treatment.

Our objectives are 1) to evaluate the validity and accuracy of Dental Monitoring™ AI image analysis notification software in identifying generalized and localized detection of clinical conditions regarding clear aligner issues and 2) other oral hygiene concerns compared to the “gold standard” orthodontist visual exam of uploaded scans; and 3) to conduct a pilot study assessing the first six months of treatment outcomes based on ABO DI and ABO OGS indices when implementing remote dental monitoring in clear aligner treatment between three groups: Dental Monitoring, 7-day, 14-day aligner change interval.

clinically evaluate the accuracy of Dental Monitoring’s (DM) artificial intelligence (AI) image analysis and oral hygiene notification algorithm in identifying oral hygiene and mucogingival conditions.

Detailed results and inferences:

Generalized and Localized Detection for Clear Aligner Parameters Results

In this study, 535 observation points for 23 patients were examined for the general detection of clear aligner parameters. Based on the crosstabulations in Appendix Table 1, the estimated accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the notifications generated by DM on the general detection (which implies the AI can identify if a clinical condition is present or absent in the image) of clinical findings related to clear aligner parameters are shown in Table III. The accuracy estimation the DM artificial intelligence image analysis software was high (>90%) across all evaluated parameters, except for moderate accuracy of slight unseat (80.7%). The sensitivity was high (>90%) for all parameters, except for moderate sensitivity for noticeable unseat (80.1%) and slight unseat (76.8%). The

specificity was high (>90%) for all parameters. The PPV estimation was high (>90%) for all parameters, except for moderate PPV for aligner damage (86.4%). The NPV was high (>90%) for all parameters, except for low NPV for slight unseat (55.4%).

Table I. Validity Values of Dental Monitoring Notifications for General Detection of Clinical Findings Regarding Clear Aligner Parameters During Clear Aligner Treatment

Validity Values of DM Notifications for General Detection in Clear Aligner Parameters						
General Clear Aligner DM Notification	Accuracy	Sensitivity	Specificity	PPV	NPV	Total Notification Sample
Slight Unseat	80.7%	76.8%	93.7%	97.5%	55.4%	535
Noticeable Unseat	93.1%	80.1%	99.2%	97.9%	91.4%	535
Attachment Loss	98.5%	94.4%	99.8%	99.2%	98.3%	535
Aligner Damage	99.3%	95.0%	99.4%	86.4%	99.8%	535
Button Loss	100.0%	100.0%	100.0%	100.0%	100.0%	535

Abbreviations: PPV = positive predictive value; NPV = negative predictive value.

After each observation timepoint was examined, the examiner identified individual teeth where the clinical finding was detected. Therefore, there are variable numbers of notification samples for localized detections (which implies the AI can identify the exact tooth number where the clinical condition is present or absent) of clear aligner parameters. Based on the crosstabulations in Appendix Table 2, the estimated accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the notifications generated by DM on the local detection of clinical findings related to clear aligner parameters are shown in Table IV. The accuracy estimation the DM artificial intelligence image analysis software was high (>90%) across all evaluated parameters, except for moderate accuracy of noticeable unseat (83.2%) and low accuracy of slight unseat (54.2%). The sensitivity was high (>90%) for attachment loss and button loss, moderate for noticeable unseat (70.4%) and low for slight unseat

(54.7%). The specificity was high (>90%) for all parameters except for low specificity for slight unseat (51.5%). The PPV estimation was high (>90%) for all parameters, except for moderate PPV for slight unseat (85.5%) and aligner damage (89.0%). The NPV was high (>90%) for all parameters, except for moderate NPV for noticeable unseat (76.8%) and low NPV for slight unseat (17.9%).

Table II. Validity Values of Dental Monitoring Notifications for Localized Detection of Clinical Findings Regarding Clear Aligner Parameters During Clear Aligner Treatment.

Validity Values of DM Notifications for Localized Detection in Clear Aligner Parameters						
Localized Clear Aligner DM Notification	Accuracy	Sensitivity	Specificity	PPV	NPV	Total Notifications Sample
Slight Unseat	54.2%	54.7%	51.5%	85.5%	17.9%	1425
Noticeable Unseat	83.2%	70.4%	95.7%	94.1%	76.8%	739
Attachment Loss	97.8%	94.0%	99.8%	99.5%	96.9%	626
Aligner Damage	97.0%	89.0%	98.3%	89.0%	98.3%	1111
Button Loss	100.0%	100.0%	100.0%	100.0%	100.0%	545

Abbreviations: PPV = positive predictive value; NPV = negative predictive value.

Generalized and Localized Detection for Oral Hygiene Parameters Results

In this study, 535 observation points for 23 patients were examined for the general detection of oral hygiene parameters. Based on the crosstabulations in Appendix Table 3, the estimated accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the notifications generated by DM on the general detection (presence or absence) of clinical findings related to oral hygiene parameters are shown in table V. The accuracy estimation the DM artificial intelligence image analysis software was high (>90%) across all evaluated parameters, except for moderate accuracy of slight plaque/food (83.6%) and slight gingivitis (87.3%). The sensitivity was moderate for slight plaque/food (68.6%), slight gingivitis (68.9%) and buccal calculus (75%) and low for noticeable plaque/food (28.6%), noticeable

gingivitis (8.3%), tooth spot (32.1%), black triangles (5.6%), mucosa irregularity (25.0%). The sensitivity was null for gingival recession because the AI incorrectly recognized a crown margin as “recession”. The specificity was high (>90%) for all parameters except for moderate specificity for slight plaque/food (87.8%). The PPV was high (>90%) for noticeable gingivitis (100%), black triangles (100%), and mucosa irregularity (100%). The PPV was moderate for slight gingivitis (78.2%) and buccal calculus (87.5%). The PPV was low for slight plaque/food (61.4%), noticeable plaque/food (66.7%), tooth spot (58.6%), and gingival recession (0.0%). The NPV was high (>90%) for all parameters.

Table III. Validity Values of Dental Monitoring Notifications for General Detection of Clinical Findings Regarding Oral Hygiene Parameters During Clear Aligner Treatment.

Validity Values of DM Notifications for General Detection in Oral Hygiene Parameters						
General Oral Hygiene DM Notification	Accuracy	Sensitivity	Specificity	PPV	NPV	Total Notifications Sample
Slight Plaque/ Food	83.6%	68.6%	87.8%	61.4%	90.8%	535
Noticeable Plaque/ Food	98.9%	28.6%	99.8%	66.7%	99.1%	535
Slight Gingivitis	87.3%	68.9%	93.5%	78.2%	89.9%	535
Noticeable Gingivitis	97.9%	8.3%	100.0%	100.0%	97.9%	535
Tooth Spot	91.0%	32.1%	97.5%	58.6%	92.9%	535
Black Triangles	96.8%	5.6%	100.0%	100.0%	96.8%	535
Mucosa Irregularity	99.4%	25.0%	100.0%	100.0%	99.4%	535
Gingival Recession	99.4%	#NULL!	99.4%	0.0%	100.0%	535
Buccal Calculus	98.1%	75.0%	99.4%	87.5%	98.6%	535

Abbreviations: PPV = positive predictive value; NPV = negative predictive value.

After each observation timepoint was examined, the examiner identified individual teeth where the clinical finding was detected. Therefore, there are variable numbers of notification samples for localized detections (identify which individual teeth have the finding) of oral hygiene parameters. Based on the crosstabulations in Appendix Table 4, the estimated accuracy, sensitivity,

specificity, positive predictive value (PPV) and negative predictive value (NPV) for the notifications generated by DM on the local detection of clinical findings related to oral hygiene parameters are shown in Table VI. The accuracy estimation the DM artificial intelligence image analysis software was high (>90%) for all parameters except for moderate accuracy of tooth spot (84.6%) and low accuracy of slight plaque/food (49.7%) and slight gingivitis (39.9%). The sensitivity was low (<70%) for all parameters except for high sensitivity for noticeable gingivitis (100%) and null for gingival recession (for reasons stated above). The specificity was high (>90%) for all parameters except for low specificity for slight plaque/food (69.4%) and slight gingivitis (68.3%). The PPV was high (>90%) for black triangles, mucosa irregularity and buccal calculus, moderate for noticeable plaque/food (71.4%) and low for slight plaque/food (27.4%), slight gingivitis (33.5%), noticeable gingivitis (11.8%), tooth spot (33.3%), gingival recession (0.0%). The NPV was high (>90%) for all parameters except for low NPV for slight plaque/food (57.4%) and slight gingivitis (41.8%).

Table IV. Validity Values of Dental Monitoring Notifications for Localized Detection of Clinical Findings Regarding Oral Hygiene Parameters During Clear Aligner Treatment.

Validity Values of DM Notifications for Localized Detection in Oral Hygiene Parameters						
Localized Oral Hygiene DM Notification	Accuracy	Sensitivity	Specificity	PPV	NPV	Total Notifications Sample
Slight Plaque/ Food	49.7%	18.3%	69.4%	27.4%	57.4%	862
Noticeable Plaque/ Food	98.2%	38.5%	99.6%	71.4%	98.5%	577
Slight Gingivitis	39.9%	14.4%	68.3%	33.5%	41.8%	1150

Noticeable Gingivitis	97.2%	100.0%	97.2%	11.8%	100.0%	538
Tooth Spot	84.6%	30.9%	91.8%	33.3%	90.9%	579
Black Triangles	92.0%	4.3%	100.0%	100.0%	91.9%	604
Mucosa Irregularity	99.3%	20.0%	100.0%	100.0%	99.2%	534
Gingival Recession	99.4%	#NULL!	99.4%	0.0%	100.0%	533
Buccal Calculus	91.1%	58.5%	99.0%	93.5%	90.8%	630

Abbreviations: PPV = positive predictive value; NPV = negative predictive value.

ABO Discrepancy Index (DI) Score Results

A comparison of ABO DI scores between the three groups at pre-treatment timepoint (T0) for clear aligner treatment are shown in Table VII. The overall DI, overjet, anterior open bite, crowding, occlusion, buccal posterior crossbite, cephalometric, and “other” scores showed no significant difference between the three groups, indicating similar initial conditions across these parameters at T0. Overbite ($p=.024$), lateral open bite ($p=.023$), and lingual posterior crossbite ($p=.031$) showed statistically significant differences between the three groups at T0. After further examination, overbite (mean difference=1.000, $p=.025$) demonstrated statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have started with slightly more severe overbite at T0. Lateral open bite (mean difference = -1.636, $p=.036$) showed statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have started with slightly less severe lateral open bite at T0. Lingual posterior crossbite (mean difference = -0.818, $p=.041$) showed statistically significant differences between group 1 (DM) and group 3 (14-day), indicating that the DM group may have started with slightly less severe lingual posterior crossbite at T0. Though these values indicate statistically significant differences between some parameters of the DI score, it may not necessarily translate to clinical significance. For instance, when grading a case with the ABO discrepancy index (DI), overbite is scored in the following categories: >1 to ≤ 3 mm = 0 points; >3 to ≤ 5 mm = 2 points; >5 to ≤ 7 mm = 3 points; Impinging (100%) = 5 points. Therefore, on average, there must be a 1.67

difference in points to change the overbite score. In this study, there was a 1.000 difference in overbite points and therefore is not considered to be clinically significant. A similar logic can be applied for all discrepancy index parameters, as described in the ABO instructions for calculating ABO Discrepancy Index Score (“ABO Discrepancy Index,” n.d.)

Table V. Comparison of ABO DI Scores Between the Three Groups at Pre-Treatment Timepoint (T0) for Clear Aligner Treatment.

ABO DI at T0										
Variable	G1 (DM) n=12		G2 (7-day) n=14		G3 (14-day) n=11		ANOVA	Multiple Comparisons Bonferroni		
	Mean	SD	Mean	SD	Mean	SD		G1- G2	G1- G3	G2- G3
Overall DI Score	13.58	9.904	13.86	6.431	15.73	9.275	0.808	NS	NS	NS
Overjet	1.92	0.793	2.21	0.893	3.00	2.366	0.205	NS	NS	NS
Overbite	0.33	0.778	1.00	1.240	0.00a	0.000	0.024*	0.192	1.000	0.025*
Anterior OB	0.25	0.866	0.21	0.579	0.45	0.688	0.682	NS	NS	NS
Lateral OB	0.08	0.289	0.00	0.000	1.64	2.803	0.023*	1.000	0.061	0.036*
Crowding	0.67	0.778	1.00	1.840	0.73	0.786	0.782	NS	NS	NS
Occlusion	1.17	1.586	1.86	1.657	2.73	2.240	0.138	NS	NS	NS
Lingual PC	0.00	0.000	0.14	0.535	0.82	1.250	0.031*	1.000	0.041*	0.099
Buccal PC	0.00 ^a	0.000	0.14	0.535	0.18	0.603	0.610	NS	NS	NS
Cephalometric	7.33	8.026	5.43	5.302	4.09	6.580	0.507	NS	NS	NS
Other	1.83	2.588	1.86	2.568	2.09	2.212	0.963	NS	NS	NS

Abbreviations: NS = Not Significant, OB = Open Bite, PC = Posterior Crossbite; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

A comparison of ABO DI scores between the three groups at 6-month treatment timepoint (T1) for clear aligner treatment is shown in Table VIII. All parameters of ABO DI, except for lingual posterior crossbite (p=.021), showed no significant difference between the three groups, indicating similar initial conditions across these parameters at T1. After further examination, lingual posterior crossbite (mean difference=-0.636, p=.020) demonstrated statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have a slightly less severe lingual posterior crossbite at T1. Though these values indicate a statistically significant difference, it is not clinically significant, based on the grading system of ABO Discrepancy Index (“ABO Discrepancy Index,” n.d.).

Table VI. Comparison of ABO DI Scores Between the Three Groups at 6-month Treatment Timepoint (T1) for Clear Aligner Treatment.

ABO DI at T1										
Variable	G1 (DM) n=12		G2 (7-day) n=14		G3 (14-day) n=11		ANOVA	Multiple Comparisons Bonferroni		
	Mean	SD	Mean	SD	Mean	SD		G1-G2	G1-G3	G2-G3
Overall DI Score	13.25	7.175	11.36	5.048	13.09	10.163	0.777	NS	NS	NS
Overjet	1.75	0.866	2.07	1.141	1.73	0.905	0.616	NS	NS	NS
Overbite	0.50	0.905	0.43	0.852	0.00a	0.000	0.227	NS	NS	NS
Anterior OB	0.00	0.000	0.50	1.871	0.91	1.221	0.274	NS	NS	NS
Lateral OB	0.50	1.000	0.86	1.351	0.55	1.036	0.693	NS	NS	NS
Crowding	0.25	0.452	0.29	0.469	0.09	0.302	0.495	NS	NS	NS
Occlusion	1.17	1.586	1.71	1.899	2.36	2.501	0.371	NS	NS	NS
Lingual PC	0.17	0.389	0.00	0.000	0.64	0.924	0.021*	1.000	0.143	0.020*
Buccal PC	0.00a	0.000	0.14	0.535	0.18	0.603	0.610	NS	NS	NS
Cephalometric	7.42	7.573	4.21	3.215	4.36	7.061	0.353	NS	NS	NS
Other	1.50	1.883	1.14	2.282	2.27	3.036	0.511	NS	NS	NS

Abbreviations: NS = Not Significant, OB = Open Bite, PC = Posterior Crossbite; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

A comparison of ABO DI score change between the three groups T0 and T1 after 6-months of clear aligner treatment are shown in Table IX. The paired t-test for group 2 (7-day) indicated two parameters that were statistically significant from T0 to T1. Overbite (mean difference = -0.571, $p = .026$) indicates that this parameter had a statistically significant decrease from T0 to T1, however this was not clinically significant. Lateral open bite (mean difference = 0.857, $p = .034$) indicates that this parameter had a statistically significant increase from T0 to T1, however this was not clinically significant.

The paired t-test for group 3 (14-day) indicated that crowding (mean difference = -0.636, $p = .011$) demonstrated a statistically significant change from T0 to T1, however this was not clinically significant.

After comparing ABO DI score changes from T0 (pre-treatment) to T1 (6-months of treatment) across the three groups, the overall DI score changes are not statistically significant between any groups and suggests similar effectiveness across all groups. However, two parameters, overbite ($p=.055$) and lateral open bite ($p=.008$), showed statistically significant differences from T0 to T1, but they were not clinically significant.

Table VII. Comparison of ABO DI Score Change Between the Three Groups T0 and T1 after 6-months of Clear Aligner Treatment.

ABO DI at T1-T0													
Variable	T-Test/Paired Sample Statistics									ANOVA	Multiple Comparisons Bonferroni		
	G1 (DM) n=12			G2 (7-day) n=14			G3 (14-day) n=11				G1-G2	G1-G3	G2-G3
	Mean Diff.	SD	P Value	Mean Diff.	SD	P Value	Mean Diff.	SD	P Value				
Overall DI Score	-0.333	4.599	0.806	-2.500	6.858	0.196	-2.636	4.081	0.058	0.514	NS	NS	NS
Overjet	-0.167	0.835	0.504	-0.143	0.864	0.547	-1.273	2.494	0.121	0.142	NS	NS	NS
Overbite	0.167	1.030	0.586	-0.571	0.852	0.026*	0.000 ^a	0.000 ^a	0.000 ^a	0.055*	NS	NS	NS
Anterior OB	-0.250	0.866	0.339	0.286	2.016	0.605	0.455	1.128	0.211	0.487	NS	NS	NS
Lateral OB	0.417	1.084	0.210	0.857	1.351	0.034*	-1.091	1.973	0.097	0.008*	1.000	0.063	0.008*
Crowding	-0.417	0.900	0.137	-0.714	1.590	0.117	-0.636	0.674	0.011*	0.804	NS	NS	NS

Occlusion	0.000	1.206	1.00	-0.143	1.460	0.720	-0.364	0.809	0.167	0.773	NS	NS	NS
Buccal PC	0.000 ^a	0.000 ^a	0.000 ^a	-0.143	0.535	0.336	-0.182	0.405	0.167	0.136	NS	NS	NS
Lingual PC	0.167	0.389	0.166	0.000	0.784	1.000	0.000	0.894	1.000	1.000	NS	NS	NS
Cephalometric	0.083	2.314	0.903	-1.214	3.62	0.232	0.273	2.102	0.676	0.359	NS	NS	NS
Other	-0.333	1.155	0.339	-0.714	1.49	0.096	0.182	1.401	0.676	0.277	NS	NS	NS

Abbreviations: NS = Not Significant, OB = Open Bite, PC = Posterior Crossbite; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

ABO Objective Grading System (OGS) Results

A comparison of ABO OGS scores between the three groups at pre-treatment timepoint (T0) for clear aligner treatment is shown in Table X. The alignment, buccolingual inclination, overjet, occlusal relationship, interproximal contacts, and root angulation scores showed no significant difference between the three groups, indicating similar initial conditions across these parameters at T0. Overall OGS score ($p=.029$), marginal ridges ($p=.017$), and occlusal contacts ($p=.038$) showed statistically significant differences between the three groups at T0. After further examination, overall OGS score (mean difference = -16.045, $p=.026$) demonstrated statistically significant differences between group 1 (DM) and group 3 (14-day), indicating that the DM group may have started with less severe overall OGS score at T0. Marginal ridges (mean difference = -2.174, $p=.017$) showed statistically significant differences between group 1 (DM) and group 3 (14-day), indicating that the DM group may have started with slightly less severe marginal ridge score at T0. Occlusal contacts (mean difference = -3.786, $p=.042$) showed statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have started with slightly less severe occlusal contacts score at T0.

Though these values indicate statistically significant differences between some parameters of the OGS score, it may not necessarily translate to clinical significance. By following similar logic as discussed above and carefully following instructions of the ABO OGS, all three

parameters (overall OGS score, marginal ridges, occlusal contacts) show clinically significant differences at T0.

Table VIII. Comparison of ABO OGS Scores Between the Three Groups at Pre-Treatment Timepoint (T0) for Clear Aligner Treatment.

ABO OGS at T0										
Variable	G1 (DM) n=12		G2 (7-day) n=14		G3 (14-day) n=11		ANOVA	Multiple Comparisons Bonferroni		
	Mean	SD	Mean	SD	Mean	SD		G1- G2	G1- G3	G2- G3
Overall OGS Score	43.50	15.716	49.93	13.975	59.55	10.885	0.029*	0.730	0.026*	0.276
Alignment	19.33	8.637	24.21	5.117	25.18	5.016	0.074	NS	NS	NS
Marginal Ridges	2.92	1.505	4.36	1.737	5.09	2.023	0.017*	0.134	0.017*	0.922
Buccolingual Inclination	4.50	4.011	5.07	3.050	4.82	3.281	0.916	NS	NS	NS
Overjet	4.25	3.415	2.57	4.603	2.82	3.920	0.543	NS	NS	NS
Occlusal Contacts	2.00	3.275	1.21	1.888	5.00	5.310	0.038*	1.000	0.166	0.042*
Occlusal Relationship	3.83	4.648	3.79	3.262	6.09	2.879	0.239	NS	NS	NS
Interproximal Contacts	2.83	3.689	4.21	5.820	6.45	5.681	0.255	NS	NS	NS
Root Angulation	3.83	2.406	4.50	3.006	4.09	1.640	0.787	NS	NS	NS

Abbreviations: NS = Not Significant; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

A comparison of ABO OGS score between the three groups at 6-month treatment timepoint (T1) for clear aligner treatment are shown in Table XI. The marginal ridges, buccolingual inclination, overjet, occlusal contacts, occlusal relationship, interproximal contacts, and root angulation showed no significant difference between the three groups, indicating similar initial conditions across these parameters at T1. Overall OGS score (p=.050) and alignment (p=.004) showed statistically significant differences between the three groups at T1. After further examination, overall OGS score (mean difference= -15.538, p=.048) demonstrated statistically significant differences between group 1 (DM) and group 3 (14-day), indicating that the DM group may have a less severe overall OGS score at T1. Alignment (mean difference = -5.705, p=.016) showed statistically significant differences between group 1 (DM) and group 3 (14-day), indicating

that the DM group may have a less severe alignment score at T1. Alignment (mean difference = -6.097, $p=0.007$) also showed statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have a less severe alignment score at T1.

Though these values indicate statistically significant differences between some parameters of the OGS score, it may not necessarily translate to clinical significance. By following similar logic as discussed above and carefully following instructions of the ABO OGS, all three parameters (overall OGS score and alignment) show clinically significant differences at T1.

Table IX. Comparison of ABO OGS Score Between the Three Groups at 6-month Treatment Timepoint (T1) for Clear Aligner Treatment.

ABO OGS at T1										
Variable	G1 (DM) n=12		G2 (7-day) n=14		G3 (14-day) n=11		ANOVA	Multiple Comparisons Bonferroni		
	Mean	SD	Mean	SD	Mean	SD		G1-G2	G1-G3	G2-G3
Overall OGS Score	37.92	15.871	41.29	15.379	53.45	12.25	0.040*	1.000	0.048*	0.143
Alignment	11.75	5.362	11.36	3.177	17.45	5.241	0.004*	1.000	0.016*	0.007*
Marginal Ridges	3.42	2.353	3.64	2.023	5.00	3.225	0.282	NS	NS	NS
Buccolingual Inclination	4.00	3.464	4.36	2.925	4.36	3.585	0.953	NS	NS	NS
Overjet	4.33	4.141	6.00	6.312	4.00	3.633	0.554	NS	NS	NS
Occlusal Contacts	7.00	6.030	6.07	6.281	8.00	3.286	0.685	NS	NS	NS
Occlusal Relationship	2.58	3.988	2.93	3.474	5.64	3.139	0.094	NS	NS	NS
Interproximal Contacts	2.08	3.204	3.36	4.517	5.55	4.480	0.143	NS	NS	NS
Root Angulation	2.75	1.960	3.57	2.652	3.45	1.753	0.607	NS	NS	NS

Abbreviations: NS = Not Significant; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

A comparison of ABO OGS score change between the three groups T0 and T1 after 6-months of clear aligner treatment are shown in Table XII. The paired t-test for group 1 (DM) indicated three parameters that were statistically significant from T0 to T1. Overall OGS score (mean difference = -5.583, $p = .012$) indicates that this parameter had a statistically significant decrease from T0 to T1. Alignment (mean difference = -7.583, $p = .002$) indicates that this parameter had a statistically significant decrease from T0 to T1. Occlusal contacts (mean difference = 5.000, $p = .014$) indicates that this parameter had a statistically significant increase from T0 to T1. All three parameters are considered clinically significant due to the instructions followed for ABO OGS.

The paired t-test for group 2 (7-day) indicated three parameters that were statistically significant from T0 to T1. Overall OGS score (mean difference = -8.643, $p = .009$) indicates that this parameter had a statistically significant decrease from T0 to T1. Alignment (mean difference = -12.857, $p = <.001$) indicates that this parameter had a statistically significant decrease from T0 to T1. Occlusal contacts (mean difference = 4.857, $p = .008$) indicates that this parameter had a statistically significant increase from T0 to T1. All three parameters are considered clinically significant due to the instructions followed for ABO OGS.

The paired t-test for group 3 (14-day) indicated that alignment (mean difference = -7.727, $p = .001$) had a statistically significant decrease from T0 to T1. This is also considered a clinically significant difference.

After comparing ABO OGS score changes from T0 (pre-treatment) to T1 (6-months of treatment) across the three groups, the overall OGS score changes are not statistically significant between any groups and suggests similar effectiveness across all groups. However, alignment

($p=.043$) showed statistically significant differences from T0 to T1, but it was not significant after further testing.

Table X. Comparison of ABO OGS Score Change Between the Three Groups T0 and T1 After 6-months of Clear Aligner Treatment

ABO OGS at T1-T0										
Variable	T-Test/Paired Sample Statistics									ANOVA
	G1 (DM) n=12			G2 (7-day) n=14			G3 (14-day) n=11			
	Mean Diff.	SD	P Value	Mean Diff.	SD	P Value	Mean Diff.	SD	P Value	
Overall OGS Score	-5.583	6.445	0.012*	-8.643	10.529	0.009*	-6.091	12.926	0.149	0.717
Alignment	-7.583	6.403	0.002*	-12.857	5.419	<0.001*	-7.727	5.711	0.001*	0.043*
Marginal Ridges	0.500	1.931	0.389	-0.714	2.199	0.246	-0.091	1.814	0.871	0.317
Buccolingual Inclination	-0.500	1.834	0.365	-0.714	2.813	0.359	-0.455	1.572	0.36	0.950
Overjet	0.083	2.575	0.913	3.429	4.363	0.011*	1.182	3.027	0.224	0.056
Occlusal Contacts	5.000	5.954	0.014*	4.857	5.763	0.008*	3.000	6.293	0.145	0.673
Occlusal Relationship	-1.250	2.734	0.142	-0.857	2.598	0.239	-0.455	1.128	0.211	0.716
Interproximal Contacts	-0.750	1.545	0.121	-0.857	3.483	0.374	-0.909	2.548	0.264	0.990
Root Angulation	-1.083	1.782	0.059	-0.929	2.093	0.121	-0.636	1.206	0.111	0.829

Abbreviations: NS = Not Significant; a = correlation and t cannot be computed because the standard error of the difference is 0. Statistical significance set to 0.05.

Treatment efficiency results

A comparison of treatment efficiency parameters between the three groups within 6 months of clear aligner treatment are shown in Table XIII. The total visits within 6 months, number of emergency/unscheduled visits in 6 months, and number of initial trays showed no significant difference between the three groups, indicating similar conditions across these parameters. The number of scheduled office visits in 6 months ($p=.036$) and tray number at 6 months ($p=.006$) showed statistically significant differences between the three groups. After further examination, the number of scheduled office visits in 6 months (mean difference= -0.988, $p=.032$) demonstrated statistically significant differences between group 1 (DM) and group 2 (7-day), indicating that the

DM group may have a lower number of scheduled office visits in 6 months of treatment time. The tray number at 6 months (mean difference = 8.857, p=.005) showed statistically significant differences between group 2 (7-day) and group 3 (14-day), indicating that the 7-day group may have progressed through their aligners at a faster rate compared to the 14-day group; this result may be expected as the 7-day group had less initial time in each tray compared to the 14-day group.

Table XI. Treatment Efficiency Parameters Between the Three Groups After 6-months of Clear Aligner Treatment.

Treatment Efficiency Parameters after 6 months										
Variable	G1 (DM) n=12		G2 (7-day) n=14		G3 (14-day) n=11		ANOVA	Multiple Comparisons Bonferroni		
	Mean	SD	Mean	SD	Mean	SD		G1-G2	G1-G3	G2-G3
Total visits	4.42	2.275	5.21	1.051	4.91	0.944	0.425	NS	NS	NS
Number of scheduled office visits	2.58	1.084	3.57	0.938	3.09	0.701	0.036*	0.032*	0.596	0.622
Number of emergency office visits	0.75	1.422	0.64	0.842	0.82	0.751	0.914	NS	NS	NS
Number of initial trays	24.17	7.661	25.43	8.671	26.27	7.926	0.823	NS	NS	NS
Tray number at 6 months	17.67	7.011	21.86	7.378	13.00	1.764	0.006*	0.288	0.267	0.005*

Abbreviations: NS = Not Significant; a = correlation and t cannot be computed because the standard error of the

difference is 0. Statistical significance set to 0.05.