Reliability of torque expression by the Invisalign® appliance: A retrospective study

Raj Gaddam, Elissa Freer, Brett Kerr and Tony Weir
Discipline of Orthodontics, School of Dentistry, The University of Queensland, Brisbane, Australia

Introduction: Incisal torque in orthodontics has always been important due to its relationship with smile aesthetics and the stability of a final treatment outcome. The aim of this study was to quantify the accuracy of torque expression predicted by ClinCheck® planning associated with Invisalign® treatment, compared to clinical outcomes.

Methodology: Forty adult patients consecutively treated using the Invisalign appliance were selected. Torque and the interincisal angle (IIA) were measured at T0 (pretreatment), T1 (predicted post-treatment), and R (end of initial aligner sequence) stages on STL files using metrology software (Geomagic Control X).

Results: The difference between the predicted and achieved torque (torque differential) was statistically significant ($P < 0.001$) in the labial crown torque group with a mean of 6.43° (SD = 7.09) for the upper central incisor, 5.06° (SD = 7.32) for the upper lateral incisor and 2.75° (SD = 5.7) for the lower incisor sub-groups. In contrast, the corresponding lingual/palatal crown torque sub-groups did not show a statistically significant torque differential ($P >0.05$). Similar observations were made for the IIA differential (difference between predicted and achieved IIA) with a mean IIA differential of 9° (SD = 10.73; $P < 0.001$) and -3.4° (SD = 7.9; $P > 0.05$) in the labial and lingual/palatal crown torque IIA groups, respectively.

Conclusions: Torque was under-expressed when the teeth were moving labially, and fully- or over-expressed when moved lingually, indicating Invisalign’s inefficiency in torqueing incisors in the labial direction.

(Aust Orthod J 2021; 37: 3 - 13. DOI: 10.21307/aoj-2021-001)

Received for publication: October 2020
Accepted: February 2021

Raj Gaddam: admin@wellingtonorthodontics.org; Elissa Freer: e.freer@uq.edu.au; Brett Kerr: uqbkerr@uq.edu.au; Tony Weir: Tony@tonyweir.com.au

Introduction
The optimal torque of teeth is considered essential for an ideal occlusal relationship, the stability of an orthodontic result and for smile aesthetics.¹ In the early years of Invisalign treatment, clinicians believed that this aesthetic alternative might be an ideal system for delivering torque, since aligners envelop the entire tooth crown thereby creating longer moment arms.² To date, research has not supported that assumption.³

Invisalign treatment has regularly been found to be inferior to fixed appliances in achieving desired torque.⁴-⁶ Using the Objective Grading Scoring (OGS) developed by the American Board of Orthodontics (ABO), Invisalign-treated cases lost 4.19 points in comparison to 2.81 points for cases treated by fixed appliances.⁵

To improve mechanical requirements, Align Technology introduced a new aligner material (SmartTrack™), and modifications by way of resin tooth attachments and Power Ridges™. SmartTrack, a multilayer thermoplastic polymer made of polyurethane and a co-polyester, was introduced in 2012, claiming better stress relaxation than its predecessor, Exceed30™.⁷ Researchers studied the torque outcomes utilising resin attachments and Power Ridges, and found that, while both were practical, an overall loss of torque of up to 50% was noted in relation to the predicted outcomes.⁸ Despite this, ClinCheck® predictions from Align Technology suggested results that differed from the actual attained outcomes.⁹

Previous research provides limited and unclear information regarding the accuracy of clear aligners
and their ability to generate a desirable expression of torque. The aim of the present study was to quantify the accuracy of ClinCheck torque predictions compared to the actual clinical outcome.

The null hypothesis (H₀) states that there is no difference between the predicted and clinically achieved incisor torque produced by Invisalign treatment.

Methodology

Ethical approval was granted in June 2018 by the University of Queensland Human Research Ethics Committee (Project number: 1836).

A sample comprised of 40 subjects (29 females, 11 males; mean age 25.5yrs, SD = 3.2) treated using the Invisalign appliance (SmartTrack) and strictly meeting pre-determined inclusion criteria was selected in a consecutive treatment order (Table I). The majority of the subjects (37) displayed pretreatment crowding and the remaining subjects (3) showed mild spacing. Compliance was monitored by visual observation of wear and the clinical assessment of aligner fit. Only subjects who displayed full compliance were included in the study. The sample was identified from a database collected by a single practitioner (TW), and was comprised of all cases treated by Invisalign-trained and experienced specialist orthodontists between the years of 2013 and 2018. All participants gave explicit written consent for their complete records to be used for educational and research purposes.

Metrology software (Geomagic Control X) was used for processing digital STL files, generated through intraoral arch scanning (iTero® Element™, Align Technology, Inc.). The long axis of the teeth was auto-generated utilising the flood-selection tool (Figure 1). A transverse reference plane for the upper arch was custom generated by the software on the pretreatment digital model using the flood-selection tool on the most distal right and left maxillary molars and the upper right central incisor. Similarly, a reference plane in the lower arch was generated by using the most distal right and left mandibular molars and the lower left central incisor (Figure 2). The angle between the reference plane and the vector (simulating the virtual long axis) of the incisors represented the incisor torque (Figure 3). Angular measurements were recorded for pretreatment (T₀), predicted post-treatment (T₁) and end of initial aligner sequence stage (R) digital models. Superimpositions were performed using the best-fit surface registration (global and fine) feature with a 50-iteration count (Figure 4). The generated reference
plane on the T₀ model was used for recording the torque expression on both the superimposition entities (T₀−T₁ and T₀−R) (Figure 5). The interincisal angle (IIA) was calculated by measuring the angle between the long axis of the upper central incisor to its counterpart in the lower arch (Figure 6). The upper and lower arches were occluded by the software with the aid of a pre-registered wax bite taken from the patient. The T₀ and T₁ models were superimposed to analyse predicted changes in torque and interincisal angles. Similarly, the T₀ and R stage models were superimposed to assess the achieved changes. A difference between predicted and achieved change will result in a torque differential or IIA differential (Table II).

The study sample was divided into two major groups, labial and lingual/palatal crown torque groups, based on the direction of prescribed crown torque. The groups were further sub-divided into upper central incisor (UC), upper lateral incisor (UL) and lower incisor (LI) sub-groups. Additionally, the difference between the predicted and achieved IIA was analysed.

Figure 2. Generation of reference plane.

Figure 3. Measurement of torque.
Table I. Criteria for sample selection.

Inclusion criteria:
- Non-growing patients
- Complete permanent dentition
- Orthodontic treatment with Invisalign (SmartTrack)
- All subjects finished the initial aligner wear as determined by ClinCheck
- Two-week aligner wear protocol
- Simple Class I malocclusion comprising spacing (< 4 mm) to crowding (< 6 mm)
- Non-extraction treatment
- No interproximal reduction
- No intermaxillary elastics
- No orthognathic surgery
- No restorative treatment to incisors and distal most molars during orthodontic treatment
- All the subjects included in this study met the criteria of compliance

Exclusion criteria:
- Periodontally compromised patients
- Medications which alter bone metabolism (e.g.: Bisphosphonates, PG inhibitors)
- Patients treated with Power Ridges

Table II. Calculation of torque, interincisal angle (IIA) differential and mean accuracy.

- Predicted change in torque = torque at T1 – torque at T0
- Achieved change in torque = torque at R – torque at T0
- Torque differential = predicted change in torque – achieved change in torque

- Predicted change in IIA = IIA at T1 – IIA at T0
- Achieved change in IIA = IIA at R – IIA at T0
- IIA differential = predicted change in IIA – achieved change in IIA

- Mean accuracy for torque or IIA = 100 - (predicted change-achieved change/predicted change) X 100

Figure 4. Best-fit superimposition (global and fine).
for the two major torque groups, as IIA was predicted to reduce in the labial and increase in the lingual/palatal crown torque groups.

**Statistics**

The upper right central incisor was randomly selected (Excel random number generator) to determine the torque differential. It was calculated that a minimum of 15 patients would be required to achieve a power of 0.8 at an alpha level of 0.05 through a power analysis (G*power software). The validity of measurement was assessed by comparing predicted and achieved torque values for 20 participant records, randomly selected, for the central incisor. Measurements were repeated after two weeks by the same operator (RG) to determine intra-operator error. The inter-operator error was also measured by repeating the measurements twice by a second clinician (AM) following an interval of two weeks. Complete intra- and inter-examiner reliability was noted (ICC = 1), which may be explained by the generation of values by custom algorithms inherent to the metrology software.

Statistical analysis was performed using Excel (Microsoft, WA, USA) and SPSS software (version 24; SPSS Inc., IL, USA). The level of statistical significance was set at $P < 0.05$. The results were subjected to paired sample $t$-test to evaluate the difference between predicted and achieved change. Pearson correlation coefficients and regression equations were calculated to determine the correlation between the predicted and achieved change in incisor inclination.

**Results**

Descriptive statistics for the torque and IIA differential are presented in Table III and Figure 7. A positive sign for the resultant differential indicates under-expression, while a negative sign indicates over-expression. The results showed significant under-expression of predicted torque change in the labial crown torque group. This contrasts with the lingual/palatal crown torque group, which showed over-expression in UC and LI sub-groups, and marginal under-expression in the UL sub-group. The mean accuracy for the torque differential is displayed in Figure 8.

The paired student $t$-test (Table III) showed the torque differential in all labial crown torque subgroups to be statistically significant ($P < 0.001$), resulting
in rejection of the null hypothesis of no difference between the predicted and achieved change in torque from a prescribed labial crown torque. Conversely, in the lingual/palatal crown torque group, all sub-groups demonstrated the opposite effect as the torque differential was statistically insignificant ($P > 0.05$) leading to acceptance of the null hypothesis.

A similar trend of torque differential was observed in the IIA differential, as under-expression of predicted IIA change (Table III) in the labial crown torque -IIA group was observed in contrast to the lingual/palatal crown torque -IIA group. The paired student $t$-test suggested that the change in the IIA was statistically significant ($P < 0.001$) in the labial crown torque group, but was insignificant in the lingual torque group ($P > 0.05$). The mean accuracy of the IIA differential is displayed in Figure 8.

A linear regression analysis (Table IV) indicated a strong correlation ($R^2$ ranged from 0.25–0.74) between the predicted and achieved change. In the labial torque group, the coefficients ($B$) showed that, for every additional degree in predicted change of torque, the achieved change increased by 0.49° for UC and 0.36° for the UL sub-groups, and an improved 0.91° for the LI sub-group. In the lingual/palatal crown torque group, for every additional degree in predicted change of torque, the achieved change increased by 0.9°, 0.66° and 0.77° for the UC, UL and LI sub-groups, respectively. The achieved change in the interincisal angle for the labial and lingual/palatal crown torque-IIA groups increased by 0.57° and 0.7°, respectively, for every additional degree of predicted change of IIA, demonstrating a similar trend of greater expression of torque change in the lingual/palatal crown torque groups. The negative sign to the constant (Y intercept of the regression line) values in the labial crown torque groups further promoted a tendency towards under-expression of desired torque change, which was the opposite in the lingual/palatal crown torque groups.

Discussion

Non-growing patients were chosen for the present study as they represent the majority of patients currently receiving aligner treatment worldwide. In addition, they may provide better compliance in appliance wear, further reducing the possibly of bias resulting from a lack of adherence of wear instructions. The response to ‘torqueing’ forces may
differ in growing patients who have less bone density and higher metabolic turnover. Non-growing patients also eliminated the issues of change in arch width and tooth eruption resulting from normal development, which could compromise the superimpositions.

The grouping of the sample was not performed based on gender or a distinction between corresponding teeth on the contralateral side, due to the lack of substantial evidence that these variables could affect the outcome. Despite weak evidence for a correlation between age and the rate of tooth movement using the Invisalign appliance by Chisari et al., the current study did not include age as a variable as all sample subjects were adults.

The superimpositions could not be conducted on stable reference planes/areas (e.g., palatal rugae) as the pretreatment stage digital models were to be superimposed with the predicted post-treatment digital models generated by the ClinCheck planning, which lacks stable reference areas. The chosen superimposition method was the best-fit (global and fine) method introduced by Grünheid et al., which allowed the models to be superimposed on the teeth that moved the least. It is noteworthy that the interincisal angles measured without superimpositions and reference planes demonstrated a similar result trend, which independently validated the superimposition method to a reasonable extent.

The transverse reference plane used in the present study was selected from past research conducted by Tepedino. This plane is unique in its ability to precisely calculate the torque of the lateral incisors, despite those teeth, in contrast to the central incisors, being on a curved area of the arch. A conventional coronal reference plane would not be able to accurately record the torque of the lateral incisors. The vectors reflecting the virtual long axis of the tooth may not represent the true long axis. Magkavali-Trikka et al. found the difference between the predicted virtual long axis and the true long axis could be highly variable, ranging between 2° and 37.6°. This disparity would not affect the current study as the difference was calculated between the two virtual long axes as predicted or achieved change, which, in theory, will not be different to an angle calculated between two true long axes.

The values are considered as clinically significant if they range from –2° to +2°, as calculated by Grünheid et al. and Tai et al. following equivalence testing using two one-sided t-tests. The descriptive statistics showed that all of the sub-groups in the labial crown torque group fell significantly short ($P < 0.001$) in achieving the prescribed level of torque. In the labial torque group, the mean torque differential in the UC and UL sub-groups was 6.43° and 5.06°, denoting a clear clinically significant difference. The mean torque
The under-expression of predicted change in torque in the labial crown torque group may be explained from a biomechanical perspective. Incisors prescribed for palatal root/labial crown torque will experience a resultant extrusive component of force. This, coupled with the flexible nature of the aligners, results in a gap between the tooth and the edge of aligner on the palatal aspect, which then compromises the necessary correction in crown torque. The LI sub-group showed a better expression of crown torque than other sub-groups, which could be attributed to lower incisor proclination during the treatment levelling of the curve of Spee. Very few teeth in the labial crown torque group (5 of 56 in the UC, 5 of 51 in the UL, 21 of 112 in the LI sub-groups) showed an over-expression of desired change in torque.

In contrast, the mean torque differential in the lingual/palatal crown torque group was as low as -0.73° for the UC, 0.36° for the UL, and -0.67° for the LI sub-groups, showing no statistical ($P > 0.05$) or clinical significance. However, this is a misleading conclusion of negligible torque differential from the descriptive statistics of this group. On closer inspection, the over-expression of desired change in torque in a significant number of subjects (14 of 24 in UC, 14 of 29 in UL, 25 of 48 in the LI sub-groups) negated the tendency of under-expression in the remaining subjects in the relevant subgroups. Hence, it may be surmised that
the nature of torque expression is quite unpredictable in magnitude and direction when lingual/palatal crown torque is prescribed.

The described results are in partial agreement with a prospective study conducted by Kravitz et al., who found better accuracy in the expression of lingual/palatal crown torque (53.1%) when compared to the labial crown torque (37.6%) between maxillary incisors. The lower incisors in the present study expressed the desired change in torque more reliably than the upper incisors in both the labial and lingual/palatal crown torque groups. This concurs with research performed using F22 aligners by Lombardo et al. in which it was found that an accuracy of 86.1% of torque expression of the lower incisors was achieved in comparison to 64.5% in the upper incisors. However, Lombardo et al. did not differentiate between labial and lingual torque expression. The over-expression of lingual/palatal crown torque in the upper central incisors was also reported in a study of upper first premolar extraction conducted by Dai et al. It was found that the achieved change was greater than the predicted change by 5.16 ± 5.92°.

Several factors could affect the torque expression of incisors other than the mechanical force system delivering prescribed torque by manufactured appliances. The correction of rotations can affect the measurement of incisor torque in predicted final or end of initial aligner sequence stage models, because rotational correction will not always occur purely along the long axis of the teeth. The amount of pre-existing spacing or crowding requiring a transverse contraction or expansion can lead to a corresponding lingually- or labially-directed force (bow-string effect) affecting torque expression. The disparity resulting from an underestimation of the mesio-distal width of the teeth by the ClinCheck plan could result in tighter final aligners, possibly contributing to a lingually-directed force on the incisors. The thickness of attachments on the labial surface of the incisors can cause a lingually-directed force from the lips. As most of the described factors result in a lingually-directed force compared to a labially-directed force, the under-expression of labial torque and over-expression of lingual torque noted in the present study could be justified. Additionally, Elkholy et al., in a study of three different types of aligner thickness and material, found aligner material thickness could contribute to a varying amount of labial and lingual force.

The overall results achieved in the current study for the upper incisors showed improved mean accuracy of 61.4% (range: 15.5–116.3) in torque expression, when compared to the study by Simon et al. A mean accuracy of 49.1% was found (range: 29.9–71.6) between the upper incisors with attachments that were predicted to express torque by more than 10°. However, Simon’s study did not distinguish between labial and lingual torque expression. The improved overall accuracy of the present study could reflect the advancements made by Invisalign over recent years, possibly in the replacement of the Exceed30 material by the SmartTrack material.

The current study contradicts previous studies that found minimal or no significant difference in torque differential between predicted change and achieved change. Grünheid et al. did not mention the magnitude of predicted torque, which could have affected the achieved torque as both are strongly correlated in a linear fashion, as shown by the coefficients of determination ($R^2$) in the linear regression analysis of the present study. Grünheid et al. reported statistically significant 1.75 ± 2.86° of torque differential for the upper central incisors and a curious torque differential of only 0.08 ± 2.93° for upper lateral incisors. It was shown, however, that the lower central incisors over-expressed torque in a lingual direction by -0.66 ± 2.61° for the central incisors and -29 ± 2.34° for the lateral incisors.

The above research was reviewed by Tai et al., who reported no significance between predicted and achieved change in the linear dimension of incisor movement, except minor differences for UC of -0.45 ± 0.64 mm in the labiobuccal direction. Both studies grouped the teeth with labial and lingual torque prescriptions together, with similar numerical signage to describe torque values, and drew conclusions with a compensation in statistics. However, the need for compensatory statistics was avoided in the current study by separately grouping the data in two opposing directions. Tepedino et al. found no significant difference between the predicted and achieved change; however, the grouping of data according to the direction of prescribed torque was not specified.

A similar trend supporting the above observations was evident in interincisal angle change. In the labial crown torque–IIA group, the intended change in interincisal angle fell short by approximately one-third (34.8%), which was statistically significant.
The mean difference in IIA expression was $9^\circ$, demonstrating a clinically significant under-expression of the intended angular change. Only a small percentage of the sample (5 of 61) showed an over-expression of the intended change in the interincisal angle.

In contradistinction, in the lingual/palatal crown torque -IIA group, there was a mean over-expression in the change of interincisal angle by $-3.4^\circ$, which was not statistically significant ($P > 0.05$). The insignificance may be a result of the relatively smaller number of patients in this group and a large associated standard deviation ($7.9^\circ$). Despite most of the angular measurements showing over-expression (14 of 19) in this group, there is still a near 25% chance of under-expression of intended change in interincisal angle, contributing to the unpredictability.

The limitations of the study involve a lack of blinding of the operator given the nature of the research, which was a part of postgraduate coursework tailored to meet course curriculum requirements. Due to the limited availability of subjects meeting pre-determined inclusion criteria, the sample was selected from a consecutively-treated patient pool rather than by randomisation. However, the selection bias associated with a retrospective study design was kept minimal by comparing the pretreatment digital models to the end of initial aligner sequence (to calculate achieved change in torque and IIA), which will aid in including patients not only satisfied but also unsatisfied after initial aligner sequence. The present study focused on mild to moderate Class I malocclusions without the need for orthodontic extractions, and so future studies could choose samples involving more complex cases. The lack of a stable structure or area for the superimpositions, and a modification of the existing methodology in superimposition studies to determine the centre of rotation of achieved tooth movement, may also be addressed in forthcoming studies.

Conclusions

1. Incisor torque is under-expressed when incisors are programmed to move labially and over-expressed to a minor extent when incisors are programmed to move lingually.
2. The nature of torque expression is relatively unpredictable in magnitude and direction when lingual/palatal crown torque is prescribed.
3. Lower incisors demonstrated a more reliable expression of torque than the upper incisors.
4. A change in interincisal angle is under-expressed when incisors are programmed to move labially and over-expressed when incisors are programmed to move lingually.

Recently, Invisalign has undergone considerable improvements, yet continues to lag in reliably achieving a desired change in torque. The under-expression of labial crown torque may result in anterior interferences that could significantly contribute towards the development of a lateral open bite, seen inconsistently with clear aligner treatment. Based on the current research findings, overcorrection is recommended when prescribing labial crown torque, which could be staged in the final 5–10 additional aligners, thereby reducing the burden of refinement.

Conflict of interest

The authors declare no financial or personal conflict of interest.

Corresponding author

Raj Gaddam
125 Davey Street
Hobart
Australia
Email: r.gaddam@uqconnect.edu.au

Acknowledgements

This study was supported by student research grants from Australian Society of Orthodontists Foundation for Research and Education (ASOFRE) and the University of Queensland.

References


