Stability of occlusal outcome during long-term retention: the time-dependent variation of the American Board of Orthodontics index

Angst, Christine ; Eliades, Theodore ; Papageorgiou, Spyridon N

Abstract: BACKGROUND The aim of this retrospective cohort study was to assess long-term occlusal changes at debond and a median of 8 years afterwards (in retention) with the American Board of Orthodontics (ABO) objective grading system and identify risk factors. MATERIALS AND METHOD Fifty patients (median age: 14.3 years at debond; 60% female) treated with fixed appliances (25 with and 25 without premolar extractions) were included. The occlusal outcome was assessed with the ABO tool and analyzed statistically at 5%. RESULTS Extraction treatment was associated with better occlusal outcome than non-extraction treatment (34.2 versus 40.9 points; P = 0.009). In retention, ABO scores improved by 7.4 points, while patients with worse debond finishing improved more afterwards (P = 0.001). Alignment/rotations deteriorated in 58% of the cases and occlusal relationships in 38% of the cases. Marginal ridges improved more for extraction than non-extraction patients (28% versus 0%; P = 0.001). Occlusal relationships improved more for cases that 'passed' the ABO requirements at debond than failed cases (64% versus 28%; P = 0.02). Furthermore, patients with worse debond ABO scores were more likely to deteriorate at alignment/rotations in retention. Finally, the proportion of cases passing the ABO requirements improved considerably between debond (28%) and in retention (54%) as half (47%) of the cases that had failed the ABO requirements at debond passed them in retention. CONCLUSIONS Considerable long-term occlusal changes are seen post-debond, which mostly favour improved settling. Extraction treatment and higher finishing quality at debond significantly influenced the chance for improvement. However, setting a cut-off score to denote treatment excellence showed considerable instability through time.

DOI: https://doi.org/10.1093/ejo/cjaa004

TITLE PAGE

Stability of occlusal outcome during long-term retention: the time-dependent variation of the American Board of Orthodontics index

Christine Angst,[1] Theodore Eliades[1], Spyridon N. Papageorgiou[1]

[1] Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Zurich, Switzerland

Correspondence to: Spyridon N. Papageorgiou, Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Plattenstr. 11, 8032 Zurich, Switzerland. E-mail snpapage@gmail.com.

Short title: Long-term occlusal changes post-debonding

ORCIDs

S.N. Papageorgiou: 0000-0003-1968-3326 / T. Eliades: 0000-0003-2313-4979

Words in abstract: 250/250

Words in text: 2812

Funding

No funding was available.

Conflict of interest

All authors of this study declare no conflicts of interest or other relationship/condition/circumstances that present a potential relationship of conflict of interest.

Ethical approval

Ethical approval for the use of the anonymized patients' data from the Ethical Committee of Canton Zurich (BASEC-Nr: 2018-01688).

Stability of occlusal outcome during long-term retention: the time-dependent variation of the American Board of Orthodontics index

SUMMARY

Background: The aim of this retrospective cohort study was assess long-term occlusal changes at debond and a median of 8 years afterwards (in-retention) with the American Board of Orthodontics (ABO) objective grading system and identify risk factors. **Materials and method:** Fifty patients (median age: 14.3 years at debond; 60% female) treated with fixed appliances (25 with and 25 without premolar extractions) were included. The occlusal outcome was assessed with the ABO tool and analyzed statistically at 5%. **Results:** Extraction treatment was associated with better occlusal outcome than non-extraction cases (34.2 versus 40.9 points; *P*=0.009). In-retention, ABO scores improved by 7.4 points, while patients with worse debond finishing improved more afterwards (*P*=0.001). Alignment/rotations deteriorated in 58% of the cases and occlusal relationships in 38% of the cases. Marginal ridges improved more for extraction than non-extraction patients (28% versus 0%; P=0.001). Occlusal relationships improved more for cases that 'passed' the ABO requirements at debond than failed cases (64% versus 28%; *P*=0.02). Furthermore, patients with worse debond ABO score were more likely to deteriorate at alignment/rotations in-retention. Finally, the proportion of cases passing the ABO requirements improved considerably between debond (28%) and in-retention (54%), as half (47%) of the cases that had failed the ABO requirements at debond, passed them in-retention. **CONCLUSIONS:** Considerable long-term occlusal changes are seen post-debond, which mostly favor improved settling. Extraction treatment and higher finishing quality at debond significantly influenced the chance for improvement. However, setting a cut-off score to denote treatment excellence showed considerably instability through time.

Introduction

It is widely accepted that relapse after orthodontic treatment can occur even in cases with good functional occlusion (1-4). The aetiology of relapse is neither fully understood nor can be fully predicted by any single factor (5), but includes factors like the response of displaced periodontal fibers (6), physiological maturation of the human dentition affecting its width, length, or perimeter (1,7,8), adult alterations of the craniofacial complex (9), and parafunction (10). The retention of the treatment results is therefore considered one of the most difficult problems in orthodontics (11), and relapse, especially of the mandibular incisors might be observed even with the use of retention appliances after debonding (12-13).

The majority of existing studies on post-treatment stability assess short-term relapse of the anterior region by measuring mostly incisor irregularity after extraction or non-extraction treatment and comparing different retention schemes. Such studies largely use the Peer Assessment Rating (PAR) index (3,14) that is not sensitive to the fine details of a well-balanced occlusion (like contacts, inclinations, and alignment of each tooth) or assess in-retention changes only in the short-term (15). To our knowledge, only one study (16) used the detailed objective grading system for models and radiographs of the American Board of Orthodontics (ABO) (17) that measures the fine details of a well-finished and well-balanced occlusion. The ABO tool has been used to assess the treatment outcome with or without tooth extractions (18) and with different appliances or protocols (19-22). The single study that used the ABO tool to assess post-treatment changes found that well-treated cases tended to deteriorate, and poorly finished cases tended to improve, while tooth alignment was the only criterion associated with a mean long-term worsening, as well as a less predictable pattern of change. It did not however assess the effect of extractions or the actual impact of finishing quality on in-retention occlusal changes nor did it identify factors affecting in-retention prognosis.

Therefore, the aim of the present study was to assess the long-term changes in the dentition after extraction or non-extraction treatment with fixed appliances using the ABO tool and to identify risk factors related to these beneficial or detrimental occlusal changes.

Materials and Methods

This retrospective cohort study included patients treated orthodontically in the post-graduate clinic of the []. Ethical approval for the use of the anonymized patient data was acquired from []. A priori sample

size calculation to identify a difference in total ABO-OGS score between extraction and non-extraction treatment was performed based on a previous similar study (23), using: (i) basis mean of 25.7 points, (ii) Standard Deviation (SD) of 8.7 points - assumed common between groups, (iii) a clinically meaningful difference in ABO OGS of 30% of the basis mean, (iv) use of an unpaired Student's t-test, (v) alpha of 5%, and (vi) beta of 20%. A needed sample of 22 patients/group was calculated, which was rounded up to 25 patients per group (to a total of 50 patients).

The following selection criteria had to be met: (i) patients of any age or sex, (ii) no history of syndromes, clefts, or dentofacial deformities, (iii) no history of periodontal disease, (iv) completion of comprehensive orthodontic treatment with fixed appliances on all permanent teeth up to at least the first molar, and (v) dental casts available immediately after debond and after 7-10 years. In our clinic, patients are invited 3, 5, 7, and 10 years after appliance removal to check stability and long-term changes. The cast of either the 7- or 10-year old appointment (hereafter termed in-retention) were selected from this study according to availability.

All patients were treated with a 0.018"-slot edgewise fixed appliance according to the usual protocols of our clinic without taking into account if additional appliances were used (transpalatal arches, headgears, etc). Additionally, half of the included patients were treated with extraction of 2 or four premolars in conjunction to the fixed appliance treatment. At debond all patients were fitted with fixed retainers for both the upper jaw (0.016x0.016" stainless steel wire; bonded on central and lateral incisors) and the lower jaw (0.028" stainless steel wire; bonded only on canines), which were still in place at the time of this investigation. Plaster dental casts of each patient at debond (T1) and in-retention (T2) were anonymized and scored manually using the ABO OGS and the special ABO gauge. Seven out of the eight ABO OGS criteria were scored with a special gauge, including alignment/rotations, marginal ridges, buccolingual inclination, overjet, occlusal contacts, occlusal relationships, and interproximal contacts. An ideal occlusion would get receive zero points, while for each deviation of a criterion from ideal occlusion 1 or 2 penalty points are given per tooth. Root angulation was not scored, as orthopantomograms are not made routinely at the in-retention visits due to ethical reasons. The sum of all criteria was the main outcome, while changes for each separate criterion were the secondary outcomes. Finally, according to the original publication (17), cases would 'pass' the ABO examination if they had a score of up to 30 points.

Initially, the principal investigator (CA) completed the necessary calibration process as instructed by the ABO. After scoring 30% of all models the principal investigator performed another calibration with the second author (SNP) who has experience with the ABO OGS. Finally, all models were scored independently by both investigators (CA and SNP) to assess inter-examiner agreement, followed by discussion to reach a consensus score for each model. Three weeks after the initial measurement the main investigator (CA) rescored 50% of models to assess intra-examiner agreement. The Concordance Correlation Coefficient (CCC) (24) and the Bland-Altman limits of agreement (25) were used to assess repeatability and agreement, respectively.

Normality was checked by visual inspection and with the Shapiro-Wilk test. Descriptive statistics included means with SDs, medians with Interquartile Ranges (IQRs), or absolute/relative frequencies, as appropriate. Differences between groups were assessed with independent Student's t-tests or Mann-Whitney tests, while follow-up changes (T2 minus T1 value) were assessed with paired t-tests or Wilcoxon signed-rank tests, as appropriate. After checking assumptions, factors associated with T1 ABO score or T2-T1 ABO changes were assessed with linear regression and expressed as regression coefficients and 95% Confidence Intervals (CI). Factors associated with worsening of ABO scores were assessed with generalized linear model models for the binomial family with Relative Risks (RRs) and their 95% CIs. Alpha was set at a two-sided 0.05 and all analyses were run in Stata SE14.2 (StataCorp LP, College Station, TX) with an open dataset (26).

Results

Fifty patients were included with a median age of 14.3 years (IQR: 13.4-15.2 years; range: 11.9-34.1 years) and 30 patients (60%) being female (Table 1). Half of those patients were treated with tooth extractions (8 with two upper and 17 with four premolars) and half of them without extractions and followed after T1 for a mean period of 8.0 years (IQR: 7.2-9.2 years; range: 6.4-13.5 years, as some of the patients came earlier/later than their 7- or 10-year check-ups). No significant differences between extraction/non-extraction patients were seen for sex or follow-up period, but extraction patients were significantly older at T1 than non-extraction patients (15.2 versus 13.4 years; P=0.001). The average ABO score at T1 was 37.6 ± 9.3 points, with extraction patients having significantly better finishing than non-extraction patients (ABO scores of 34.2 versus 40.9 points; P=0.009). On the other hand, patient age at T1 and sex had no influence on the ABO score at T1 (P=0.19 and P=0.82, respectively).

After a median period of 8.0 years the ABO score at T2 was 30.2 ± 10.3 points, with no significant differences between extraction / non-extraction patients (*P*=0.29). On average, there was a statistically significant improvement in ABO scores from T1 to T2 by -7.4 points (Table 2). On average, 1 criterion worsened (alignment/rotations), 2 criteria remained relatively stable (occlusal relationships and interproximal contacts), and 4 criteria improved (marginal ridges; buccolingual inclination; overjet; occlusal contacts). Patient age at T1, sex, and T1-T2 duration had no significant effect on the T1-T2 total ABO change (Table 3). However, total ABO score at T1 was significantly associated with T1-T2 change in ABO (Table 3; Figure 1), with patients with worse ABO scores at T1 experiencing a larger ABO improvement during follow-up (*P*=0.001).

When looking at individual patients, the total ABO score improved in the 8 years of follow-up for the majority of cases (82%), worsened for a few patients (16%), and remained the same for one patient (2%) (Table 4). For each specific ABO criterion, the majority of cases: (a) improved for marginal ridges, buccolingual inclination, overjet, and occlusal contacts; or (b) worsened for alignment/rotation. For the remaining criterion 'occlusal relationships' almost equal proportions improved or worsened (about 40%) and the remaining (20%) remained stable. Significant influences in the progress after T1 for two criteria were seen for the incorporation of extractions in the treatment plan and for cases passing the ABO standards at T1. As seen in Table 5, the marginal ridges of extraction cases were more likely to improve (28% versus 0%) and less likely to worsen (52% versus 96%) compared to non-extraction cases. Additionally, the occlusal relationships of 'ABO pass' cases at T1 were more likely to improve (64% versus 28%) and less likely to worsen (14% versus 53%) compared to cases getting a 'fail' score at T1.

As far as factors influencing worsening of ABO scores (totals or within each criterion) are concerned, some significant associations were seen for patient age at T1, T1-T2 duration and total ABO score at T1 (Supplementary Tables 1-2). Cases debonded at a later age were more likely to worsen for overjet and interproximal contacts (RRs of 1.09 and 1.12, respectively), while they were less likely to worsen for alignment/rotations (RR of 0.85). Additionally, the older the patient at T1 was, the less likely she/he was to worsen for overjet. Furthermore, the worse a patient was finished at T1 (the larger the total ABO score), the more likely she/he was to worsen for alignment/rotations (RR of 1.03).

At debond, 28% (n=14) of the cases fulfilled the ABO requirement of a case to pass the examination (i.e. had a score less than 30 points), while 54% (n=27) of the cases passed the

requirements in-retention. Stratifying changes from T1-T2 indicated high variability (Figure 5), as half (47%; 17/36) of the cases that initially did not pass the ABO requirements at debond, passed these requirements in-retention. Likewise, one third (29%; 4/14) of cases that initially passed the requirements, failed these requirements in-retention.

Finally, the observed agreement pre-calibration was moderate (concordance=0.69; Bland-Altman limits=-8.4 to 18.3) (Supplementary Table 3), but turned to almost perfect for both intra- (concordance=0.98; Bland-Altman limits=-3.1 to 5.2) and inter-examiner agreement (concordance=0.90; Bland-Altman limits=-5.2 to 10.5).

Discussion

The current study assessed the long-term changes in the dentition of patients treated with fixed appliances and with or without extractions with the ABO tool after a median period of 8 years (range: 6.4 to 13.5 years). The main finding of the study was that considerable changes occur in the dentition after orthodontic treatment, even with the use of upper and lower fixed retainers. This was seen on average as an improvement of the total ABO score, which reduced by 7.4 points across all patients (Table 2). Interestingly, patients with initially worse occlusal outcome (higher ABO score at debond) showed significantly greater improvements in the retention period (Figure 1). This might be due to the fact that patients with worse ABO scores had more occlusal flaws and therefore greater leeway for improvement that better-finished cases. It must be noted however that this must be regarded with some skepticism, since all included cases were finished to clinically acceptable result with solid occlusal contacts across the dental arch, minimum overjet and well-established occlusal relationships. It might well be that cases that are prematurely or incompletely finished might experience the same spontaneous settling improvement.

The sole exceptions to this tendency for improvement after debond were the criteria of alignment/rotation and occlusal relationships, where 58% and 38% of the patients deteriorated, respectively. As far as alignment/rotations are concerned, this is an easily explainable finding, since mesial migration and tertiary crowding have been well observed in early adulthood (1, 7, 8). It might be assumed that if a Class I relationship with solid cusp-to-fossa intercuspation is not established therapeutically, then the relationship might not be able to withstand any changes that might occur during adolescence and early treatment.

A crucial finding of this study was that the higher finishing quality at debond significantly influenced the prognosis of the dentition, as far as improvement of the various ABO criteria is concerned. First, ABO scores at debond were proportional to the deterioration risk of the alignment/rotations criterion (RR=1.03; $P<0.001$). This might indicate that a balanced occlusion of a well-finished case might act as an obstacle to the mesial migratory forces responsible for tertiary crowding. Second, the occlusal relationships of cases that passed the ABO requirement at T1, were more likely to improve (64%) in-retention, while 21% stayed stable and 14% deteriorated. This was significantly different ($P=0.02$) from ABO failures at T1, which were more likely to deteriorate (53%) in-retention, while 28% improved and 19% remained stable. Again, this might indicate that well-established occlusal relationships with antagonists, might remain more stable or even improve after debond. On another note, it has been also been reported that post-orthodontic dental occlusion meeting ABO standards contributes to more balanced activation of the anterior temporalis muscle during function and better subjective chewing function (27).

The use of a 30-point cut-off for the ABO score however, to denote excellence in treatment outcomes was more problematic, since this categorization into ‘passing’ or ‘failing’ cases proved to be rather fluid. Half of the cases that initially failed the ABO requirements at debond, subsequently passed these in-retention. Even worse, one-third of cases deemed a success with the ABO requirements at debond deteriorated into ‘failures’ in-retention. Finally, from the 8 patients with positive T2-T1 changes in ABO score (i.e. overall worsening of the occlusion), 4 were ‘passing’ and 4 were ‘failing’ the ABO requirements at debond. This indicates that the use of this cut-off might not be based on sound clinical or biological grounds.

Premolar extractions were performed in half of the patients in this study and this had a profound effect on the occlusal outcome. For one, extraction patients were finished significantly better than non-extraction patients at debond ($P=0.009$). This agrees with a previous study (28) that indicated a difference of 5.9 points between extraction / non-extraction patients. Furthermore, the marginal ridges of extraction cases were significantly better aligned and deteriorated less than those of non-extraction patients (Table 5). This indicates that except from a better occlusal outcome, incorporation of extractions might directly influence the stability or even the improvement of the dentitions settling after debonding, which is a novel finding. This contradicts previous findings indicating similar PAR scores and re-crowding of the anterior dentition for extraction and non-extraction patients (29). However, the

patients in that study were followed only for 2.4 years and occlusal outcome was assessed with the PAR index, which is not well suited as the ABO tool to capture the fine occlusal details (30).

Alignment/rotations deteriorated less among older patients (RR=0.85), which is logical since fewer craniofacial changes are expected with increasing age. On the other hand, overjet and interproximal contacts deteriorated more among older patients (RRs of 1.09 and 1.12). Regarding overjet this observation might be explained by reduced orthopedic treatment effects with increasing age (31) and an increased initial overjet that are corrected mostly dentally and tend to relapse more easily. Regarding interproximal spaces, the findings of this study support reports of a tendency for re-opening of spaces among older patients (32).

The present study has also several limitations. For one, it was designed retrospectively, since its long follow-up period might present feasibility difficulties, and therefore might be more biased than a prospective study. Furthermore, both assessors were stricter when grading the included cases with the ABO tool than other known studies in the field using this tool, since our aim was to identify as many fine details and deviations from the ideal occlusion. Therefore, absolute comparisons with other studies might not be easy. Additionally, only seven of the eight ABO criteria were scored, since it was not ethically justifiably to routinely take orthopantomograms in-retention without clinical indications. This means that the present results of this study might be conservative or not directly comparable to other studies and root parallelism might influence post-debond stability, also others refute this (33). Finally, even though the ABO tool is well described in the accompanying documentation, it does not contain an objectively measured value or a clear-cut yes/no response and we found that the examiner's judgement often needs to be employed, which introduced a subjective component.

Conclusions

The results of this present retrospective cohort study indicated that considerable long-term changes in the dentition occur after removal of the orthodontic appliances, which for most ABO criteria favored an improved occlusal settling. Extraction treatment and higher quality of finishing at debond significantly influenced the chance for improvement in-retention. However, setting a cut-off score to denote treatment excellence showed considerably instability through time.

Funding

No funding was available.

Conflict of interest

All authors of this study declare no conflicts of interest or other relationship/condition/circumstances that present a potential relationship of conflict of interest.

References

1. Little, R.M., Riedel, R.A. and Artun J. (1988) An evaluation of changes in mandibular anterior alignment from 10 to 20 years postretention. *American Journal of Orthodontics and Dentofacial Orthopedics*, 93, 423-8.
2. Al Yami, E.A., Kuijpers-Jagtman, A.M. and van 't Hof, M.A. (1999) Stability of orthodontic treatment outcome: follow-up until 10 years postretention. *American Journal of Orthodontics and Dentofacial Orthopedics*, 115, 300-4.
3. de Freitas, K.M., Janson, G., de Freitas, M.R., et al. (2007) Influence of the quality of the finished occlusion on postretention occlusal relapse. *American Journal of Orthodontics and Dentofacial Orthopedics*, 132, 428.e9-14.
4. Lyotard, N., Hans, M., Nelson, S., Valiathan, M. (2010) Short-term postorthodontic changes in the absence of retention. *The Angle Orthodontist*, 80, 1045-50.
5. Little, R.M. (1990) Stability and relapse of dental arch alignment. *British Journal of Orthodontics,* 17, 235-41.
6. Reitan, K., Kvam, E. (1971) Comparative behavior of human and animal tissue during experimental tooth movement. *The Angle Orthodontist*, 41, 1-14.
7. Sinclair, P.M., Little, R.M. (1983) Maturation of untreated normal occlusions. *American Journal of Orthodontics*, 83, 114-23.
8. Thilander, B. (2009) Dentoalveolar development in subjects with normal occlusion. A longitudinal study between the ages of 5 and 31 years. *European Journal of Orthodontics*, 31, 109-20.
9. Behrents, R.G., Harris, E.F., Vaden, J.L., Williams, R.A., Kemp, D.H. (1989) Relapse of orthodontic treatment results: growth as an etiologic factor. *Journal of the Charles H. Tweed International Foundation*, 17, 65-80.
10. Lang, G., Alfter, G., Göz, G., Lang, G.H. (2002) Retention and stability--taking various treatment parameters into account. *Journal of Orofacial Orthopedics*, 63, 26-41.
11. Oppenheim, A. (1934) The crisis in orthodontia Part I 2. Tissue changes during retention. Skogborg's septotomy. *American Journal of Orthodontics*, 6, 542-554.
12. Renkema, A.M., Renkema, A., Bronkhorst, E., Katsaros, C. (2011) Long-term effectiveness of canine-to-canine bonded flexible spiral wire lingual retainers. *American Journal of Orthodontics and Dentofacial Orthopedics*, 139, 614-21.

13. Steinnes, J., Johnsen, G., Kerosuo, H. (2017) Stability of orthodontic treatment outcome in relation to retention status: An 8-year follow-up. *American Journal of Orthodontics and Dentofacial Orthopedics*, 151, 1027-33.
14. Ormiston, J.P., Huang, G.J., Little, R.M., Decker, J.D., Seuk, G.D. (2005) Retrospective analysis of long-term stable and unstable orthodontic treatment outcomes. *American Journal of Orthodontics and Dentofacial Orthopedics*, 128, 568-74.
15. Hoybjerg, A.J., Currier, G.F., Kadioglu, O. (2013) Evaluation of 3 retention protocols using the American Board of Orthodontics cast and radiograph evaluation. *American Journal of Orthodontics and Dentofacial Orthopedics,* 144, 16-22.
16. Nett, B.C., Huang, G.J. (2005) Long-term posttreatment changes measured by the American Board of Orthodontics objective grading system*. American Journal of Orthodontics and Dentofacial Orthopedics,* 127, 444-50.
17. Casko, J.S., Vaden, J.L., Kokich, V.G., Damone, J., James, R.D., Cangialosi, T.J., Riolo, M.L., Owens, S.E. Jr and Bills, E.D. (1998) Objective grading system for dental casts and panoramic radiographs. American Board of Orthodontics. *American Journal of Orthodontics and Dentofacial Orthopedics*, 114, 589–599.
18. Anthopoulou, C., Konstantonis, D. and Makou, M. (2014) Treatment outcomes after extraction and nonextraction treatment evaluated with the American Board of Orthodontics objective grading system. *American Journal of Orthodontics and Dentofacial Orthopedics*, 146, 717–723.
19. Johansson, K. and Lundström, F. (2012) Orthodontic treatment efficiency with self-ligating and conventional edgewise twin brackets: a prospective randomized clinical trial. *The Angle Orthodontist*, 82, 929–934.
20. Penning, E.W., Peerlings, R.H.J., Govers, J.D.M., Rischen, R.J., Zinad, K., Bronkhorst, E.M., Breuning, K.H. and Kuijpers-Jagtman, A.M. (2017) Orthodontics with Customized versus Noncustomized Appliances: A Randomized Controlled Clinical Trial. *Journal of Dent Research*, 96, 1498–1504.
21. Papageorgiou, S.N., Koletsi, D., Iliadi, A., Peltomaki, T. and Eliades T (2019) Treatment outcome with orthodontic aligners and fixed appliances: a systematic review with meta-analyses. *European Journal of Orthodontics*, doi: 10.1093/ejo/cjz094. [Epub ahead of print].

22. Yıldırım, K. and Saglam-Aydinatay, B. (2018) Comparative assessment of treatment efficacy and adverse effects during nonextraction orthodontic treatment of Class I malocclusion patients with direct and indirect bonding: A parallel randomized clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics*, 154, 26-34.e1.
23. Mislik, B., Konstantonis, D., Katsadouris, A., Eliades, T. (2016) University clinic and private practice treatment outcomes in Class I extraction and nonextraction patients: A comparative study with the American Board of Orthodontics Objective Grading System. *American Journal of Orthodontics and Dentofacial Orthopedics*, 149, 253-8.
24. Lin, L.I. (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45, 255-268.
25. Bland, J.M. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307-310.
26. Angst, C., Papageorgiou, S.N., Eliades, T. Alterations between debonding and post-retention: a challenge of the appropriateness of the American Board of Orthodontics index as an indicator of excellence [Data set]. Zenodo. http://doi.org/10.5281/zenodo.3556751.
27. Shim, J., Ho, K.C.J., Shim, B.C., Metaxas, A., Somogyi-Ganss, E., Di Sipio, R. and Cioffi, I. (2019) Impact of post-orthodontic dental occlusion on masticatory performance and chewing efficiency. *European Journal of Orthodontics*, doi:10.1093/ejo/cjz095. [Epub ahead of print]
28. Papageorgiou, S.N., Höchli, D., Eliades, T. (2017) Outcomes of comprehensive fixed appliance orthodontic treatment: A systematic review with meta-analysis and methodological overview. *Korean Journal of Orthodontics*, 47, 401-413.
29. Janson, G., Camardella, L.T., Araki, J.D., de Freitas, M.R., Pinzan, A. (2010) Treatment stability in patients with Class II malocclusion treated with 2 maxillary premolar extractions or without extractions. *American Journal of Orthodontics and Dentofacial Orthopedics*, 138, 16-22.
30. Deguchi, T., Honjo, T., Fukunaga, T., et al. (2005) Clinical assessment of orthodontic outcomes with the peer assessment rating, discrepancy index, objective grading system, and comprehensive clinical assessment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 127, 434-43.
31. Barton, S., Cook, P.A. (1997) Predicting functional appliance treatment outcome in Class II malocclusions--a review. *American Journal of Orthodontics and Dentofacial Orthopedics,* 112, 282-6.

32. Oftedal, B., Wisth, P.J. (1982) Residual extraction sites after orthodontic treatment. Part III. Changes in the post-retention period. *European Journal of Orthodontics,* 4, 203-10.
33. Chiqueto, K., Janson, G., de Almeida, C.T., Storniolo, J.M., Barros, S.E., Henriques, J.F. (2011) Influence of root parallelism on the stability of extraction-site closures. *American Journal of Orthodontics and Dentofacial Orthopedics,* 139, e505-10.

FIGURES

Figure 1. Scatter plot of observed T2-T1 changes in total ABO score against ABO score at T1, including fitted linear regression line.

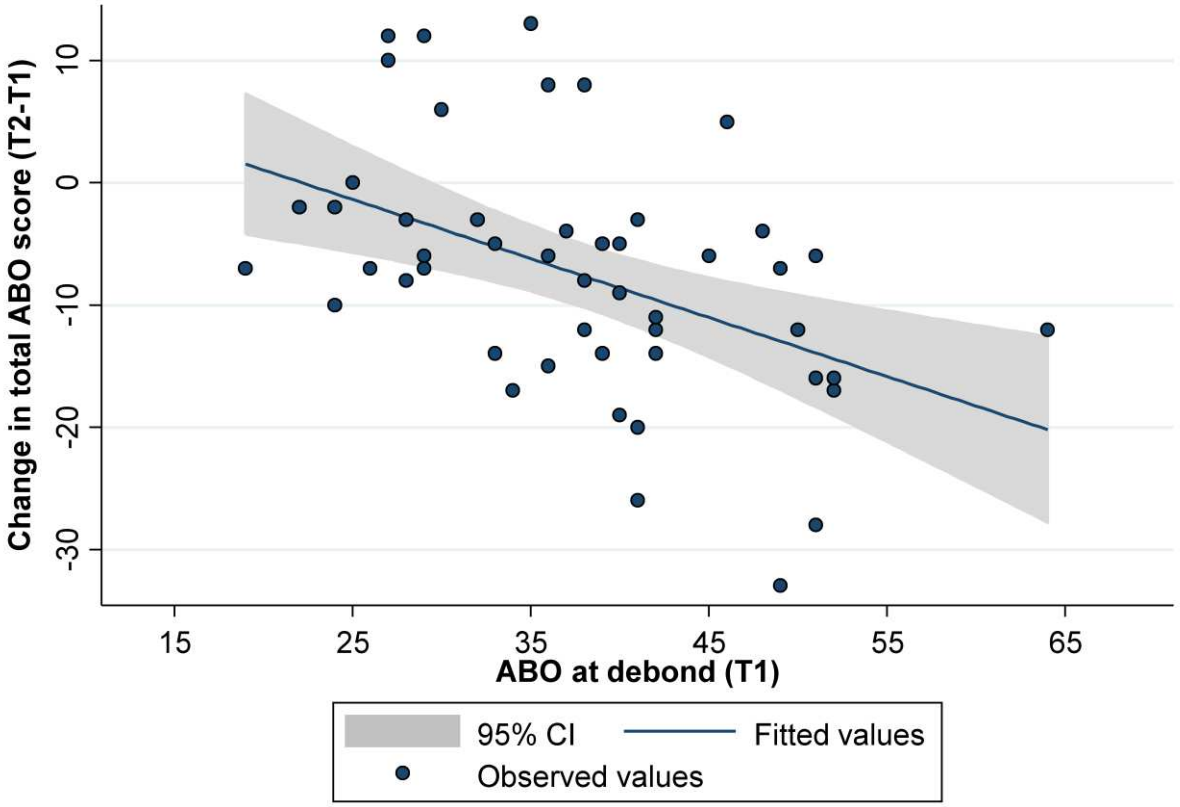


Figure 2. Scatter plot of observed T2-T1 changes in total ABO score. Green dots: patients stable or improved scores; red dots: patients with worse scores.

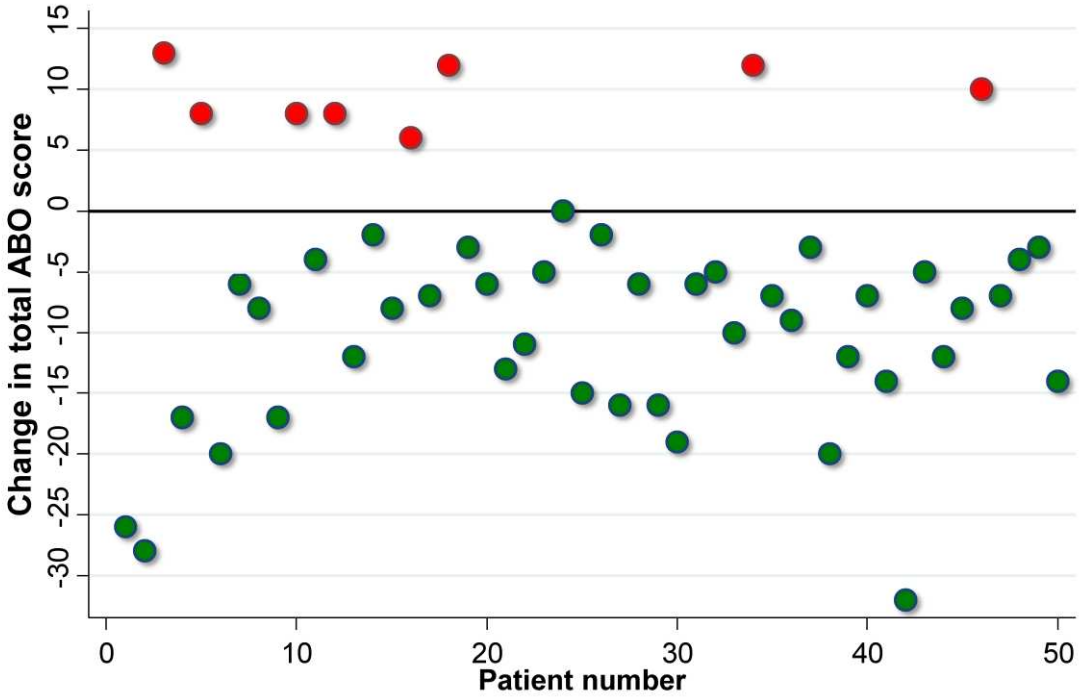


Figure 3. Scatter plot of observed T2-T1 changes in the first four ABO criteria.Green dots: patients stable or improved scores; red dots: patients with worse scores.

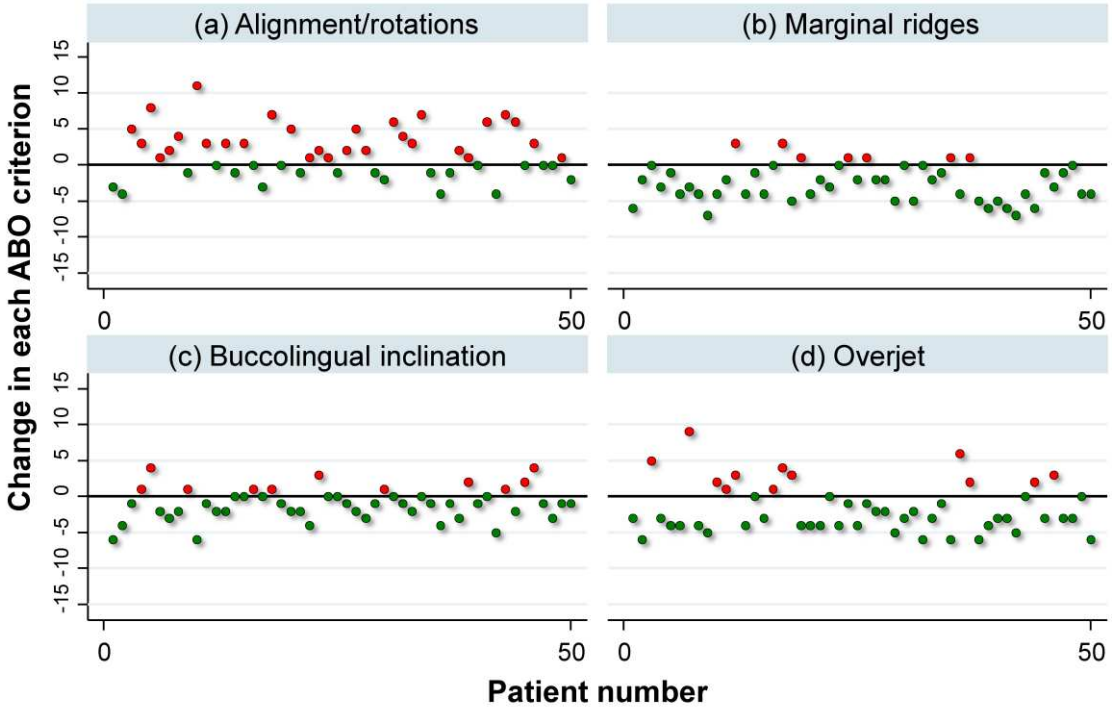


Figure 4. Scatter plot of observed T2-T1 changes in the remaining three ABO criteria for all patients: (a) occlusal contacts, (b) occlusal relationship, (c) interproximal contacts. Green dots: patients stable or improved scores; red dots: patients with worse scores.

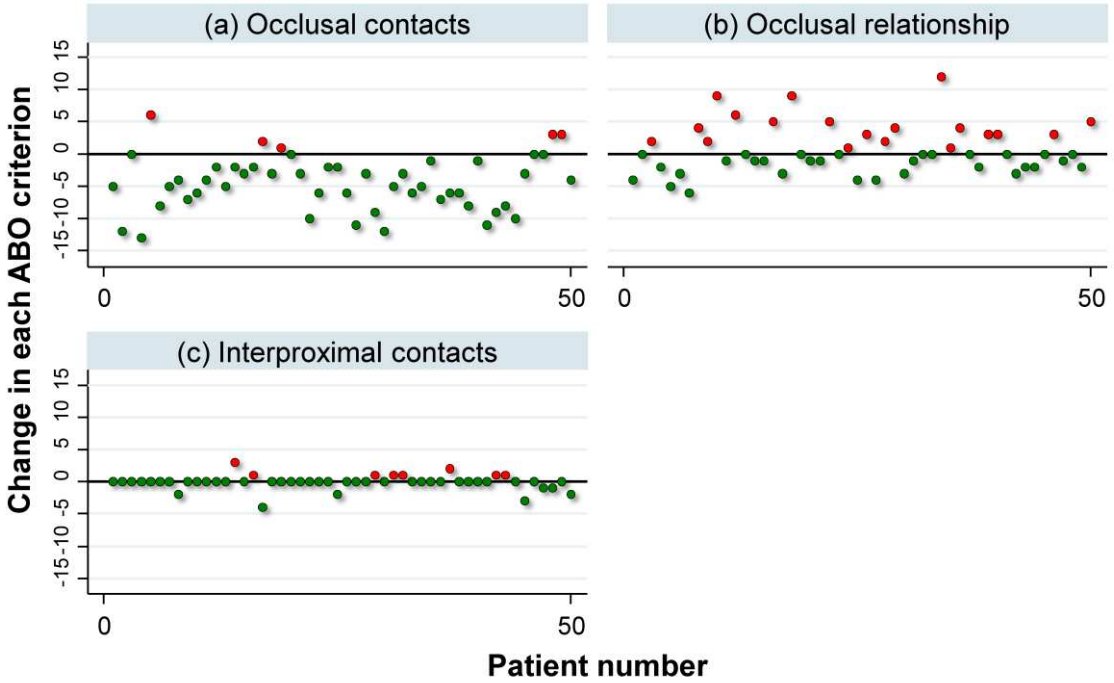
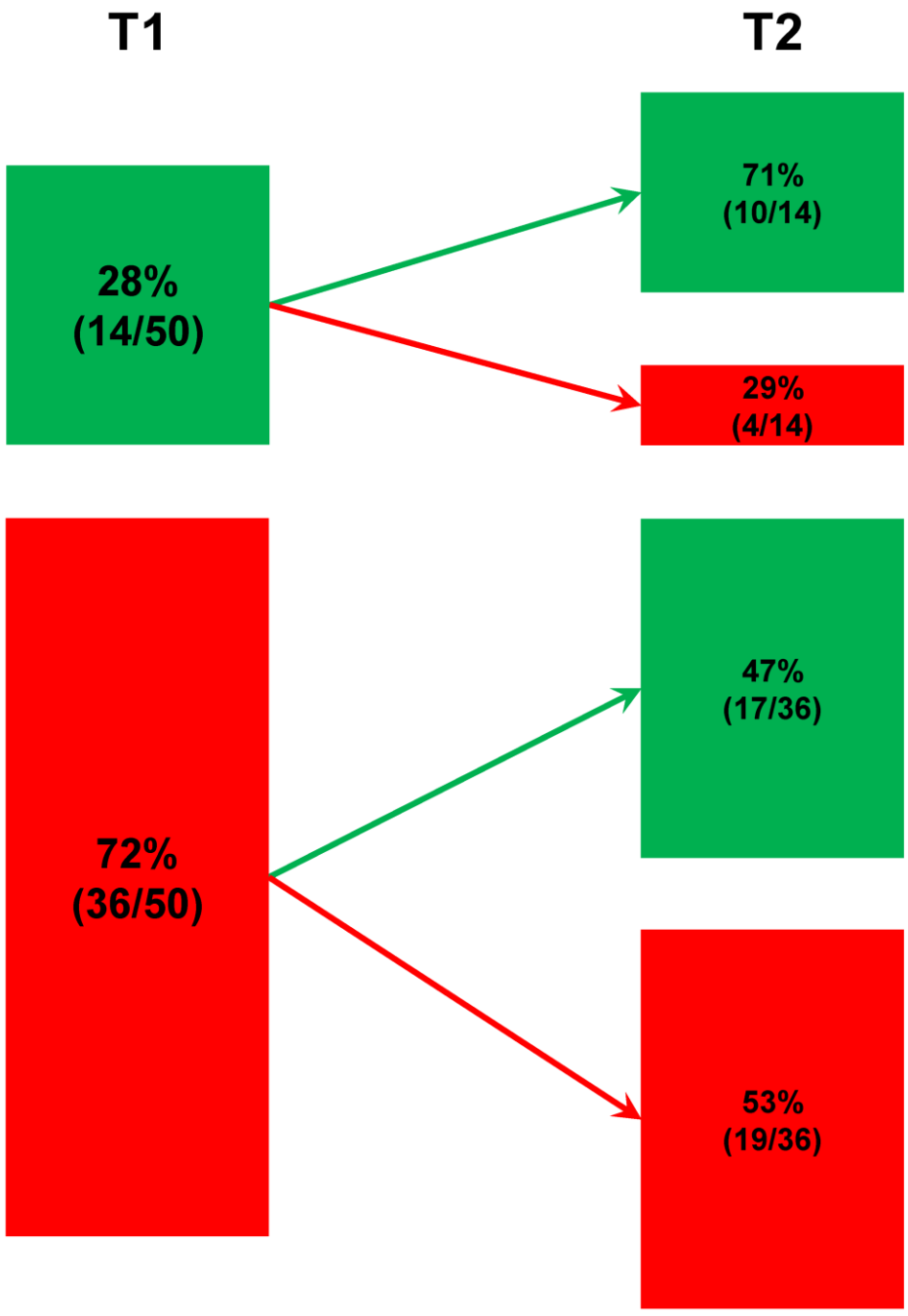


Figure 5. Proportion of patients with a passing ABO score (≤30 points; green) or with a failing score (>30 points; red) at debond (T1) and in-retention (T2).



TABLES

Table 1. Characteristics of the included patients.

		Overall			**Non-Ex**		**Ex**		
Variable		**n**			**n**		**n**		***P***
Age at debond	Median (IQR)	50	14.3 (13.4-15.2)		25	13.4 (13.2-14.1)	25	15.2 (14.3-16.0)	0.001†
	Range		11.9-34.1			11.9-15.8		12.9-34.1	
Female	N (%)	50	30 (60%)		25	14 (47%)	25	16 (53%)	0.56$
Follow-up (years)	Median (IQR)	50	8.0 (7.2-9.2)		25	8.1 (7.5-8.9)	25	8.0 (7.2-9.7)	0.91†
	Range		6.4-13.5			6.4-11.0		6.6-13.5	
ABO at debond	Mean (SD)	50	37.6 (9.3)		25	40.9 (8.7)	25	34.2 (8.7)	0.009*
	Range		19.0-60.0			27.0-60.0		19.0-51.0	
ABO in-retention	Mean (SD)	50	30.2 (10.3)		25	31.8 (11.4)	25	28.7 (9.0)	0.29*
	Range		12.0-52.0			15.0-52.0		12.0-44.0	

ABO, American Board of Orthodontics; Ex, extraction; SD, standard deviation.
† from Mann-Whitney test.
* from t-test for independent samples.
$ from chi-square test.

Table 2. Follow-up changes (in-retention minus debond) in total ABO and each criterion.

Outcome	**Average (dispersion)**	**Range**	***P***
Overall ABO score – mean (SD)	-7.4 (10.1)	-32.0, +13.0	<0.001*
Alignment/rotations – mean (SD)	+1.7 (3.4)	-4.0, +11.0	0.002*
Marginal ridges – mean (SD)	-2.5 (2.5)	-7.0, +3.0	<0.001*
Buccolingual inclination – mean (SD)	-1.0 (2.2)	-6.0, +4.0	0.003*
Overjet – median (IQR)	-3.0 (-4.0, 0)	-6.0, +9.0	<0.001$
Occlusal contacts – mean (SD)	-4.5 (4.2)	-13.0, +6.0	<0.001*
Occlusal relationships – mean (SD)	+0.6 (3.7)	-6.0, +12.0	0.21*
Interproximal contacts – median (IQR)	0 (0, 0)	-4.0, +3.0	0.92$

ABO, American Board of Orthodontics; IQR, interquartile range; SD, standard deviation.
*from paired t-test
$ from Wilcoxon signed-rank test.

Table 3. The effect of various factors on ABO OGS change (in-retention minus debond), assessed through linear regression

Factor	**Category**	**b (95% CI)**	***P***
Age at debond	Per year	0.5 (-0.2, 1.1)	0.14
Sex	Male	Reference	
	Female	-1.8 (-7.7, 4.1)	0.55
Treatment	Non-Ex	Reference	
	Ex	3.6 (-2.1, 9.3)	0.21
Follow-up	Per year	-1.3 (-3.2, 0.5)	0.15
ABO at debond	Per point	-0.5 (-0.8, -0.2)	0.001

ABO, American Board of Orthodontics; b, unstandardized regression coefficient; Ex, extraction.

Table 4. Change categories for overall ABO score and for each separate criterion, including the possible influence of extraction treatment or a passing ABO score (≤30 points) at T1.

T1-T2 Change in	Improved	Stable	Worsened		Ex *P**	ABO pass T1 *P* *
Total ABO score	41 (82%)	1 (2%)	8 (16%)		0.70	0.20
Alignment/rotations	14 (28%)	7 (14%)	29 (58%)		0.17	0.20
Marginal ridges	37 (74%)	6 (12%)	7 (14%)		0.001	0.20
Buccolingual inclination	31 (62%)	8 (16%)	11 (22%)		1.00	0.49
Overjet	34 (68%)	4 (8%)	12 (24%)		0.14	0.88
Occlusal contacts	41 (82%)	4 (8%)	5 (10%)		0.42	0.27
Occlusal relationships	21 (42%)	10 (20%)	19 (38%)		0.87	0.02
Interproximal contacts	7 (14%)	35 (70%)	8 (16%)		0.32	0.89

ABO, American Board of Orthodontics; Ex, extraction.

* from Fisher's exact test

Table 5. Stratified analysis of change categories for significant findings of Table 4.

Criterion	Category	Improved	Stable	Worsened	*P**
Marginal ridges	Non-Ex	0 (0%)	1 (4%)	24 (96%)	0.001
	Ex	7 (28%)	5 (20%)	13 (52%)	
Occlusal relationships	ABO fail at T1	10 (28%)	7 (19%)	19 (53%)	0.02
	ABO pass at T1	9 (64%)	3 (21%)	2 (14%)	

ABO, American Board of Orthodontics; Ex, extraction.

* from Fisher's exact test

Supplementary material

Supplementary Table 1. Factors associated with patients having worse score in-retention compared to debond for overall score and the first three criteria.

		ABO score		**Alignment/ rotations**		**Marginal ridges**		**Buccolingual inclination**	
Factor	**Category**	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***
Age at debond	Per year	1.06 (0.98, 1.16)	0.15	0.85 (0.73, 0.99)	0.04	1.08 (0.99, 1.17)	0.08	1.03 (0.94, 1.13)	0.50
Sex	Male	Ref		Ref		Ref		Ref	
	Female	0.67 (0.19, 2.36)	0.53	0.82 (0.52, 1.31)	0.40	0.89 (0.22, 3.55)	0.87	1.17 (0.39, 3.47)	0.78
Ex	No	Ref		Ref		Ref		Ref	
	Yes	0.60 (0.16, 2.25)	0.45	0.81 (0.50, 1.31)	0.39	1.00 (-)	NC	0.83 (0.29, 2.38)	0.73
Follow-up	Per year	0.70 (0.39, 1.24)	0.22	0.90 (0.76, 1.08)	0.27	0.98 (0.62, 1.54)	0.92	1.02 (0.74, 1.42)	0.89
ABO total	Per Point	0.95 (0.88, 1.02)	0.18	1.03 (1.01, 1.04)	<0.001	0.93 (0.85, 1.01)	0.07	0.98 (0.93, 1.04)	0.58

ABO, American Board of Orthodontics; CI, confidence interval; Ex, extraction; NC, non-calculable; Ref, reference; RR, relative risk.

Supplementary Table 2. Factors associated with patients having worse score in-retention compared to debond for overall score and criteria four to seven.

		Overjet		**Occlusal contacts**		**Occlusal relationship**		**Interproximal contacts**	
Factor	**Category**	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***	**RR (95% CI)**	***P***
Age at debond	Per year	1.09 (1.06, 1.12)	<0.001	1.02 (0.87, 1.20)	0.79	0.99 (0.91, 1.08)	0.84	1.12 (1.08, 1.17)	<0.001
Sex	Male	Ref		Ref		Ref		Ref	
	Female	1.33 (0.46, 3.84)	0.59	0.44 (0.08, 2.43)	0.35	1.14 (0.54, 2.40)	0.72	2.00 (0.45, 8.94)	0.36
Ex	No	Ref		Ref		Ref		Ref	
	Yes	1.00 (0.37, 2.68)	1.00	4.00 (0.48, 33.32)	0.20	0.90 (0.44 1.83)	0.77	1.67 (0.45, 6.24)	0.45
Follow-up	Per year	0.58 (0.33, 0.94)	0.03	1.08 (0.66, 1.77)	0.77	0.98 (0.78, 1.24)	0.86	1.26 (0.93, 1.72)	0.13
ABO total	Per Point	1.01 (0.96, 1.06)	0.78	1.03 (0.94, 1.13)	0.50	0.96 (0.93, 1.00)	0.06	0.99 (0.93, 1.06)	0.84

ABO, American Board of Orthodontics; CI, confidence interval; Ex, extraction; NC, non-calculable; Ref, reference; RR, relative risk.

Supplementary Table 3. Comparison between assessors' repeated measurements.

		CCC (95% CI)		Average difference (95% limits of agreement)
Pre-calibration				
	Inter-examiner	0.69 (0.59, 0.78)		4.96 (-8.36, 18.28)
Post-calibration				
	Inter-examiner	0.90 (0.86, 0.94)		2.68 (-5.18, 10.54)
	Intra-examiner	0.98 (0.97, 0.98)		1.05 (-3.11, 5.21)

CCC, concordance correlation coefficient; CI, confidence interval.