Orthodontics & Craniofacial Research

REVIEW ARTICLE

G. Perinetti
J. Primožič
A. Castaldo
R. Di Lenarda
L. Contardo

Is gingival crevicular fluid volume sensitive to orthodontic tooth movement? A systematic review of split-mouth longitudinal studies

Authors' affiliations:
G. Perinetti, A. Castaldo, R. Di Lenarda, L. Contardo, Department of Medical, Surgical and Health Sciences, School of Dentistry, University of Trieste, Trieste, Italy
J. Primožič, Department of Dental and Jaw Orthopaedics, Medical Faculty, University of Ljubljana, Ljubljana, Slovenia

Correspondence to:
Giuseppe Perinetti
Struttura Complessa di Clinica Odontoiatrica e Stomatologica
Ospedale Maggiore
Piazza Ospitale 1
34129 Trieste
Italy
E-mail: G.Perinetti@fmc.units.it

Perinetti G., Primožič J., Castaldo A., Di Lenarda R., Contardo L.
Is gingival crevicular fluid volume sensitive to orthodontic tooth movement? A systematic review of split-mouth longitudinal studies
Orthod Craniofac Res 2013; 16:1–19.

Abstract
To assess the scientific evidence for detectable volume changes of gingival crevicular fluid (GCF) incident to orthodontic tooth movement (OTM). A literature survey of longitudinal split-mouth studies was performed searching PubMed, SCOPUS and the Cochrane Library, with the last access in 15 April 2012. After selection, 13 articles qualified for the final analysis. One study was judged to be of medium/high quality, six were of medium quality and the rest of low quality. Across all studies, there was very little or no statistically significant change in GCF volume incident to OTM. The changes seen were generally ascribed to the clinical or subclinical inflammation consequent to placement of the fixed orthodontic appliance. A reappraisal of the data provided in these studies was performed by comparing the GCF volume from the test (moved) and control (non-moved) teeth, in terms of effects-size (ES) coefficients and variations, as percentages. Generally, the ES coefficients and the variations were below 1.0 and 20%, respectively. Therefore, when using current methods to collect and measure GCF volume, there was no meaningful diagnostic potential for the GCF volume as an index of tissue remodelling incident to OTM. The GCF volume is not a reliable index for tissue remodelling incident to OTM.

Key words: gingival crevicular fluid; orthodontic tooth movement; orthodontics; tissue remodelling

Introduction

The use of biomarkers in orthodontics is advocated for non-invasive monitoring of tissue remodelling during orthodontic treatment on an individual and site-specific basis. This is particularly true in the determination of the optimum force magnitude, which is difficult to perform *in vivo* (1).

Date:
Accepted 10 August 2012

DOI: 10.1111/ocr.12005

The gingival crevicular fluid (GCF) is a transudate of interstitial tissues which is produced by an osmotic gradient (2), and it is released into the crevicular sulcus at a flow rate of about 3 μl/h (3). However, during periodontal inflammation, the main mechanism of GCF formation becomes exudative (4), with an increase in its flow rate, and thus volume, up to 44 μl/h (3).

Previous evidence has shown that the GCF volume might be a better indicator of gingival inflammation than standard clinical assessments (5). Considering that tissue remodelling incident to orthodontic tooth movement (OTM) is triggered by an inflammatory process [for review, see (6)], and it has been hypothesized that the volume of GCF production will reflect these tissue changes [for review, see (7)]. In contrast to the quantification of specific biomarkers within the GCF, which requires a dedicated analytical set-up, the GCF volume *per se* can be determined easily and cheaply in any clinical setting using a Periotron (8). In addition, clinical recording of the periodontal conditions may be used to exclude that gingival inflammation, because of plaque accumulation, is responsible for the increase in GCF production rather than tooth movement. However, contrasting results have been reported in the literature, with studies showing both increased or unchanged GCF volumes incident to OTM.

This systematic review of previous longitudinal split-mouth studies was thus aimed at evaluating whether or not an orthodontic appliance *per se* or OTM can induce detectable changes in GCF volume in healthy young or adult orthodontic patients. Furthermore, the potential diagnostic accuracy of the GCF volume as a biomarker for tissue remodelling incident to OTM was also evaluated.

Materials and methods

Search strategy

This systematic review follows the PRISMA statements (9) and used two previous systematic reviews as a template (10, 11). Herein, all of the studies that examined the effects of OTM on GCF volume were identified through a literature survey carried out through PubMed (http://www.ncbi.nlm.nih.gov/pubmed), SCOPUS (http://www.scopus.com) and the Cochrane Library (http://www.thecochranelibrary.com) from inceptions to the last access in 15 April 2012.

The following algorithm (with MESH indication for PubMed database and asterisk indicating truncation) was used: {crevicular fluid [MESH] and [orthodontic* (MESH) or tooth movement (MESH)]}. No language restrictions were used, and through the whole Cochrane Library (set at 'search all text'), no restrictions as to the record status was set. Finally, a manual search was also performed by scoring the references within the studies examined and the titles of the papers published over the last 10 years among the following major journals in which investigations on the GCF response to OTM mainly have appeared: 1) American Journal of Orthodontics and Dentofacial Orthopedics; 2) European Journal of Oral Sciences; 3) European Journal of Orthodontics; 4) Journal of Clinical Periodontology; 5) Journal of Dental Research; 6) Journal of Periodontology; 7) Journal of Periodontal Research; 8) Orthodontics and Craniofacial Research; 9) Progress in Orthodontics; 10) The Angle Orthodontist; and 11) World Journal of Orthodontics.

Study selection and data items

The studies retrieved had to be either randomized clinical trials (RCTs) or controlled clinical trials (CCTs) on healthy young or adult orthodontic patients using a split-mouth design to investigate the effects of OTM on GCF volume, irrespective of whether or not this was a primary or secondary outcome of the study. These studies had to make use of a fixed appliance and had to include at least a control group of teeth with no orthodontic forces (irrespective of the presence of an orthodontic appliance). No restrictions were set regarding the methods used to collect the GCF or to measure its volume or the duration of the treatments. Case reports, case series, reviews and opinion articles were excluded. Of note, studies in which site-specific clinical monitoring of the periodontal conditions of the experimental sites was not included or declared were also excluded. Further details regarding these inclusion and exclusion criteria are listed in Table 1.

Eligibility assessment and data collection process were performed independently by two authors (GP and LC). The data collection was carried without blinding to the authors, and intra-examiner conflicts were resolved by discussion of each article, until a consensus was reached.

Data items

The following data items were collected: study design, sample size (with ages and sex distribution), teeth investigated (test and control), treatment and orthodontic force (including type of stress on the test sites), time points, GCF collection site and collection modality, procedure for GCF volume measurement, degree of OTM, periodontal conditions over time, GCF volumes recorded (as the principal summary measure of the quantitative analysis), main results and author conclusions about the GCF volume incident to OTM (Table 2). If a study reported merged data about the GCF volumes recorded, because of a lack of statistically significant differences among groups or over time, it was also included. Similarly, studies in which data regarding the GCF volume were shown only at a certain time point were also included, with detailed specifications. Moreover, in studies including more than one group of subjects, that is, juvenile and adult, the results were analysed individually for each group of subjects. Therefore, a comparison of the two specific groups *per se*, independent of the OTM effects on the GCF, was not considered in this analysis. Finally, other results of the studies included, that is, biochemical composition of the GCF incident to OTM, were not considered here.

Assessment of study quality and risk of bias in individual or across studies

Evaluation of methodological quality of published studies gives an indication of the strength of evidence provided. However, no single approach in assessing methodological soundness may be appropriate to all systematic reviews (12). Therefore, contextual, pragmatic, and methodological

Table 1. **Inclusion and exclusion criteria used in this review**

Inclusion criteria
Randomized or controlled clinical trials with split-mouth designs on healthy human subjects
Use of a fixed orthodontic appliance
Inclusion of a control group not subjected to orthodontic forces (independent of a previous initial aligning and levelling phase)
Use of any methodology for the recording of the GCF volume and showing these data
Exclusion criteria
Case series with no statistical analysis, case reports, studies enroling <10 subjects, comments, letters to the editor, reviews
Studies in which predetermined GCF volumes were sampled, that is, 1 μl, by methods based on micropipetting or capillary tubing
Studies with a lack of periodontal clinical monitoring and optimal periodontal conditions at the beginning of the study
Studies in which the GCF volume/flow was not measured during active treatment, but only in follow-up terms subsequent to the removal of the forces
Studies in which in spite of the split-mouth design, the data from different groups receiving different treatments (or different control groups for different treatments) were pooled without prior statistical analysis
Studies limited to the investigation of the effects of separator elastics
Studies that included pathological conditions of the periodontium, that is, gingival overgrowth or periodontitis, or clinically relevant root resorption
Studies evaluating the GCF volume in combination with orthopaedic treatment, that is, palatal expansion
Studies primarily evaluating the effects of anti-inflammatory drugs, mouthwashes or fluoridated elastomeric ligatures or laser therapy
Studies including surgical aids in the tooth movement, that is, on periodontal distraction
Studies evaluating the crevicular fluid volume around miniscrews

Table 2. **Summarized data of the 13 studies included in the present review**

Study	Study design	Sample size and mean age in years (range or SD)	Teeth	Treatment and orthodontic force	Time points	GCF collection site (stress on test site)/modality	GCF volume measure	OTM in mm (mean ± SD)/ Periodontal conditions over time	GCF volume in μl (mean ± SD)	Main results	Conclusions about GCF volume and OTM
Uematsu et al. (15)	CCT	9 M; 3 F 14.4 (0.9a)	Maxillary canine*	Distalizing force of 250 gf (elastic chain)	Baseline 1 h 24 h 168 h	Distal (compression)/ two consecutive paper strips (PerioPapers) each kept 30 s in the crevice (1 min apart) then pooled	Periotron (version not specified)	1.1 ± 0.1a/ optimal	0.41 ± 0.03 (24 h)	No statistically significant difference in GCF volume between the experimental teeth	Not discussed
			Contralateral canine	Fixed appliance (activation not specified)				ND/optimal	0.35 ± 0.05 (24 h)		
			Antagonist canine	None				Optimal			
Perinetti et al. (16)	CCT	6 M; 10 F 15.5 (11–21)	Maxillary first molar*	Distalizing force of 250 gf (Ni-Ti coil)	Baseline 1 h 7 days 14 days 21 days 28 days	Mesial (tension) and distal (compression)/ a single endodontic paper point per site and kept 30 s in the crevice	Weight determination through analytical balance (assuming 1.0 g/ml as specific weight)	1.7 ± 0.3/slightly worsened	0.15 ± 0.08 (overall)	GCF volume significantly greater in the sites wearing orthodontic appliance irrespective of the OTM	GCF volume increase was mainly due to periodontal inflammation consequent to the fixed appliance placement, rather than to OTM *per se*
			Contralateral first molar	Aligning and levelling force from archwire				ND/slightly worsened	0.15 ± 0.07 (overall)		
			Antagonist first molar	None				–/optimal	0.13 ± 0.07 (overall)		
Ren et al. (17)	CCT	41 M 24 (1.6) (adult group)	Maxillary lateral incisor*	Labial tipping of 70 cN (0.012-in nitinol archwire)	Baseline 24 h	Mesiobuccal and distobuccal (compression)/a single paper strip (PerioPaper) per site, kept 30 s in the crevice and then pooled	Periotron 6000	NA/optimald	0.10a (baseline) 0.13a (24 h); tests and controls pooled	No significant differences between moved and non-moved teeth within either adult or juvenile group. Only in the adult group a modest but significant increase was seen over time in both experimental teeth	Not discussed
			Contralateral lateral incisor	Passive fixed appliance				ND/optimal			
		43 M 11 (0.7) (juvenile group)	Maxillary lateral incisor*	Labial tipping of 70 cN (0.012-in nitinol archwire)	Baseline 24 h			NA/optimal	0.17a (baseline) 0.20a (24 h); tests and controls pooled		
			Contralateral lateral incisor	Passive fixed appliance				ND/optimal			

Table 2. **Continued**

Study	Study design	Sample size and mean age in years (range or SD)	Teeth	Treatment and orthodontic force	Time points	GCF collection site (stress on test site)/modality	GCF volume measure	OTM in mm (mean ± SD)/ Periodontal conditions over time	GCF volume in μl (mean ± SD)	Main results	Conclusions about GCF volume and OTM
Perinetti et al. (18)	CCT	8 M; 10 F 16.1 (11–22)	Maxillary first molar*	Distalizing force of 250 gf (Ni-Ti coil)	Baseline 1 h 7 days 14 days 21 days 28 days	Mesial (tension) and distal (compression)/ a single endodontic paper point per site and kept 30 s in the crevice	Weight determination through analytical balance (assuming 1.0 g/ml as specific weight)	1.9 ± 0.5/slightly worsened	0.16 ± 0.04 (overall)	GCF volume significantly greater in the sites wearing orthodontic appliance irrespective of the OTM	GCF volume increase was mainly because of periodontal inflammation consequent to the fixed appliance placement rather than OTM *per se*
			Contralateral first molar	Passive fixed appliance				ND/slightly worsened	0.17 ± 0.03 (overall)		
			Antagonist first molar	None				–/optimal	0.14 ± 0.02 (overall)		
Sugiyama et al. (19)	CCT	5 M 22.5 ± 2.8 5 F 23.4 ± 3.9	Maxillary canine*	Distalizing force of 250 gf (elastic chain) (previous aligning and levelling phase not specified)	Baseline 1 h 24 h 168 h	Distal (compression)/ two consecutive paper strips (PerioPapers), each kept 30 s in the crevice (1 min apart) then pooled	Periotron 8000	1.2 ± 0.2/optimal	0.38 ± 0.04 (24 h)	The GCF volume was similar between the experimental teeth and over the study term	Not discussed
			Contralateral canine	Same appliance with no chain (previous aligning and levelling phase not specified)				ND/optimal	0.37 ± 0.05 (24 h)		

Table 2. **Continued**

Study	Study design	Sample size and mean age in years (range or SD)	Teeth	Treatment and orthodontic force	Time points	GCF collection site (stress on test site)/modality	GCF volume measure	OTM in mm (mean ± SD)/ Periodontal conditions over time	GCF volume in μl (mean ± SD)	Main results	Conclusions about GCF volume and OTM
Perinetti et al. (20)	CCT	8M; 9 F 16.1 (11–22)	Maxillary first molar*	Distalizing force of 250 gf (Ni-Ti coil)	Baseline 7 days 14 days 21 days	Mesial (tension) and distal (compression)/ a single endodontic paper point per site and kept 30 s in the crevice	Weight determination through analytical balance (assuming 1.0 g/ml as specific weight)	1.1 ± 0.3/ slightly worsened	0.17 ± 0.04 (overall)	GCF volume significantly greater in the sites wearing orthodontic appliance irrespective of the OTM	GCF volume increase was mainly because of periodontal inflammation consequent to the fixed appliance placement rather than OTM *per se*
			Contralateral first molar	Passive fixed appliance				ND/slightly worsened	0.16 ± 0.03 (overall)		
			Antagonist first molar	None				–/optimal	0.14 ± 0.02 (overall)		
Kawasaki et al. (21)	CCT	6 M; 9 F 31 (3.6) (adult group)	Maxillary canine*	Distalizing force of 250 gf (elastic chain) after an aligning and levelling phase	Baseline 1 h 24 h 168 h	Distal (compression)/ two consecutive paper strips (PerioPapers) each kept 1 min in the crevice (1 min apart) then pooled	Periotron 8000	0.82 ± 0.22/ optimal	0.39 ± 0.04 (overall)	No significant differences between moved and non-moved teeth at each time point or over time within test and control teeth, for both the adult and juvenile groups.	Not discussed
			Contralateral/ antagonist canine	Passive fixed appliance after an aligning and levelling phase				ND/optimal	0.39 ± 0.05 (overall)		
		7 M; 8 F 15.1 (2.8) (juvenile group)	Maxillary canine*	Distalizing force of 250 gf (elastic chain) after an aligning and levelling phase	Baseline 1 h 24 h 168 h			1.23 ± 0.15/ optimal	0.46 ± 0.08 (overall)		
			Contralateral/ antagonist canine	Passive fixed appliance after an aligning and levelling phase				ND/optimal	0.45 ± 0.06 (overall)		

Table 2. **Continued**

Study	Study design	Sample size and mean age in years (range or SD)	Teeth	Treatment and orthodontic force	Time points	GCF collection site (stress on test site)/modality	GCF volume measure	OTM in mm (mean ± SD)/ Periodontal conditions over time	GCF volume in μl (mean ± SD)	Main results	Conclusions about GCF volume and OTM
Nishijima et al. (22)	CCT	4 M 14.5 (2.4) 6 F 15.4 (3.1)	Maxillary canine*	Distalizing force of 250 gf (elastic chain) (previous aligning and levelling phase not specified)	Baseline 1 h 24 h 168 h	Distal (compression)/ two consecutive paper strips (PerioPapers) each kept 1 min in the crevice (1 min apart) then pooled	Periotron 8000	1.6 ± 0.6/ optimal	0.44 ± 0.05 (24 h)	No significant difference in the GCF volume between the test and control teeth at 24 h or in the overall values	Not discussed
			Contralateral/ antagonist canine	Passive fixed appliance (previous aligning and levelling phase not specified)				–/optimal	039 ± 0.05 (24 h)		
Yamaguchi et al. (23)	CCT	3 M 21.3 ± 2.8	Maxillary canine*	Distalizing force of 250 gf (elastic chain) after an aligning and levelling phase	Baseline 1 h 4 h 8 h 24 h 72 h 120 h 168 h	Distal (compression)/ two consecutive paper strips (PerioPapers) each kept 1 min in the crevice (time interval apart not specified) then pooled	Periotron 8000	1.5 ± 0.4/ optimal	0.40 ± 0.06 (overall)	No significant difference in the mean volume of GCF at any time point between the experimental teeth	Not discussed
		6 F 23.1 ± 2.4	Contralateral canine	Passive fixed appliance after an aligning and levelling phase				ND/optimal	0.39 ± 0.07 (overall)		
Yamaguchi et al. (24)	CCT	4 M; 6 F 25.1 ± 4.4	Maxillary canine*	Aligning and levelling with low-friction appliance (0.014-in copper Ni-Ti archwire, force not specified)	Baseline 1 h 24 h 168 h	Mesiobuccal area (not specified)/two consecutive paper strips (PerioPapers) each kept 1 min in the crevice (time interval apart not specified) then pooled	Periotron 8000	NA/optimal	0.39 ± 0.08 (overall)	The GCF volume was similar between the experimental teeth and over the study term	Not discussed
			Contralateral canine*	Aligning and levelling with high-friction appliance (0.014-in copper Ni-Ti archwire, force not specified)				NA/optimal	0.40 ± 0.06 (overall)		
			Antagonist canine	No appliance				–/optimal	0.40 ± 0.08 (overall)		

Table 2. **Continued**

Study	Study design	Sample size and mean age in years (range or SD)	Teeth	Treatment and orthodontic force	Time points	GCF collection site (stress on test site)/modality	GCF volume measure	OTM in mm (mean ± SD)/ Periodontal conditions over time	GCF volume in μl (mean ± SD)	Main results	Conclusions about GCF volume and OTM
Dilsiz et al. (25)	CCT	11 M; 11 F 14.4 (1.1)	Maxillary canine*	Distalizing force of 250 gf (elastic chain) after an aligning and levelling phase	Baseline 1 h 24 h 168 h	Distal (compression)/ four consecutive paper strips (PerioPapers) each kept 30 s in the crevice (1 min apart) then pooled	Periotron 8000	1.56 ± 0.64/ optimal	0.88 ± 0.29 (overall)	The GCF volume was similar between the experimental teeth and over the study term	GCF volume is not influenced by OTM but by inflammation of the periodontal tissue
			Contralateral canine	Same appliance with no chain after an aligning and levelling phase				ND/optimal	0.88 ± 0.28 (overall)		
Luppanapornlarp et al. (26)	RCT	2 M 20.8 ± 1.2	Maxillary canine*	Distalizing force of 50 gf (Ni-Ti coil)	Baseline 1 h 24 h 1 week 1 month 2 months	Distal (compression)/ two consecutive paper strips (PerioPapers) each kept 30 s in the crevice (90 s apart) then pooled	Periotron 8000	NA/optimal	0.38 ± 0.26 (overall)	GCF volumes showed no significant difference among or within experimental teeth at any time point	The absence of change in GCF volume demonstrated good gingival health throughout experimental period
		14 F 20.2 ± 1.6	Contralateral canine*	Distalizing force of 150 gf (Ni-Ti coil)				NA/optimal	0.38 ± 0.29 (overall)		
			Antagonist canine	None				–/optimal	0.37 ± 0.25 (overall)		
Drummond et al. (27)	RCT	9 M; 7 F 17.7 (13–27)	Maxillary canine*	Distalizing force of 150 gf (Ni-Ti coil)	Baseline 1 h 24 h 7 days 14 days 21 days	Mesial (tension) and distal (compression)/ a single paper strip (PerioPaper) per site and kept 30 s into the crevice	Periotron 8000	Not shown/ optimal	0.50 ± 0.20 (mesial, overall) 0.58 ± 0.18 (distal, overall)	A modest but significant increase in the GCF volume over time was sometimes seen in both the experimental teeth. No differences between the experimental teeth or sites within teeth	GCF volume increase was mainly because of subclinical periodontal inflammation consequent to the fixed appliance placement rather than OTM *per se*
			Contralateral canine	Passive fixed appliance				–/optimal	0.51 ± 0.17 (mesial, overall) 0.59 ± 0.19 (distal, overall)		

RCT, randomized clinical trial; CCT, controlled clinical trial; OTM, orthodontic tooth movement; M, male; F, female; ND, not detectable; NA, not available; '–', not applicable.

*Test tooth; standard error of the mean instead of SD; a, median.

considerations are followed when assessing study quality (12). Herein, a modified quality evaluation method from Antczak et al. (13) was used that followed pre-established characteristics, along with the systematic scores that were assigned to the individual retrieved articles:

1. Study design (RCT: 2 points; CCT: 1 point).
2. Adequacy of sample selection description based on four criteria: 1) age and gender; 2) systemic health conditions; 3) periodontal status throughout the study; and 4) any further condition, that is, use of drugs that might have altered tooth movement or GCF volume (full description: 2 points; partial description: 1 point).
3. Adequacy of treatment description based on three criteria: 1) intensity of the force; 2) the orthodontic appliance; 3) degree of movement of the test teeth (full description: 2 points; partial description: 1 point).
4. Inclusion of a control group of teeth that were wearing a fixed appliance but were not subjected to orthodontic forces, either before or during the study period (1 point).
5. Description of the error analysis method for GCF volume determination (2 points).
6. Adequacy of statistics (parametric or nonparametric tests used where appropriate: 2 points; parametric tests used when nonparametric tests would be more appropriate, multiple comparisons with uncorrected *p* values, statistical analysis only partially described: 1 point).
7. Prior estimation of sample size (1 point).
8. Point measures and measures of variability for the key outcome (full description: 2 points; partial description: 1 point).

The quality of the studies, with maximum possible score of 14, was considered as follows:

1. Low: total score ≤8 points.
2. Medium: total score >8 and ≤10 points.
3. Medium/high: total score >10 and ≤12 points.
4. High: total score >12 points.

Moreover, the PRISMA statements (9) for the assessment of the risk of bias of individual studies have been considered herein. According to these statements, the following items should be used:

1. Concealment of randomisation.
2. RCT/CCT stopped early.
3. Patients blinded.
4. Healthcare providers blinded.
5. Data collectors blinded.
6. Outcome assessors blinded.

However, adaptation of these items was deemed necessary because of the particular designs used in the studies included in this systematic review. More in detail, these investigations do not have as a primary goal the assessment of the benefit of a given therapy; therefore, an early interruption for participants does not necessarily denote beneficial/negative effects of those interventions, that is, the use of orthodontic forces. As all these investigations were based on split-mouth designs, the patients blindness would be meaningless, and the health care provider/data collector blindness difficult to pursue. Therefore, only the concealment of randomization and outcome assessor blindness were considered in the risk of bias of individual studies.

For the assessment of the risk of bias across studies, incomplete reporting of the data on the GCF volume has been considered, that is, when the outcome was not reported for all the time points.

Summary measures and synthesis of results

The principal summary measure considered in this quantitative analysis was the GCF volume, reported as microlitres/sample, and expressed as mean/median and SD/percentiles of the compared groups. Other measures analysed but not included in the quantitative analysis were all the clinical parameters including plaque and bleeding on probing indexes, and probing depth.

When examining the studies on the effects of OTM on GCF volume, a problem that arose was that none of these studies included a sample size analysis. Therefore, when non-significant results are seen, a clear distinction has to be made whether or not a lack of statistical power was among the cause of these results. In this quantitative analysis, the effects-size (ES) coefficient (14) was calculated to perform a synthesis, that is, reappraisal, of previously reported results in terms of differences that were clinically significant, independent of the statistical significance. In this case, the ES coefficient is the ratio of the difference between the GCF volume recorded for the different test and control teeth (within each time point, or for the overall study term), divided by the within-group standard deviation (SD), calculated as follows:

$$\mathrm{ES} = \frac{m_t - m_c}{\sqrt{\dfrac{(\mathrm{SD}_t^2 \cdot n_t) + (\mathrm{SD}_c^2 \cdot n_c)}{\mathrm{SD}_t + \mathrm{SD}_c}}}$$

where m_t and m_c are the mean GCF volumes for the test and control teeth, SD_t and SD_c are the corresponding standard deviations, n_t and n_c are the corresponding sample sizes (which are equal in the split-mouth designs).

For the effects of OTM on GCF volume, an ES coefficient reported by a given study had to be >0.6 to be associated with a medium effect, and above 0.80 to be considered to be associated with a 'large' effect (14). However, when dealing with the diagnostic potential of a given parameter, such as here for GCF volume as an index of tissue remodelling incident to OTM, an ES coefficient equal to or above 1.0 has been considered as the minimum value (11). This would mean that the difference between test and control teeth in the corresponding GCF volumes must be at least equal to the variability observed within the same groups. Whenever possible, the ES coefficients were calculated using the comparisons reported in the articles identified, and the publications were then sorted according to the highest ES coefficients obtained (11). Moreover, the lowest ES coefficient retrieved in each study was also reported.

Results

Study search

The results of the automatic and manual searches are summarized in Fig. 1. According to the

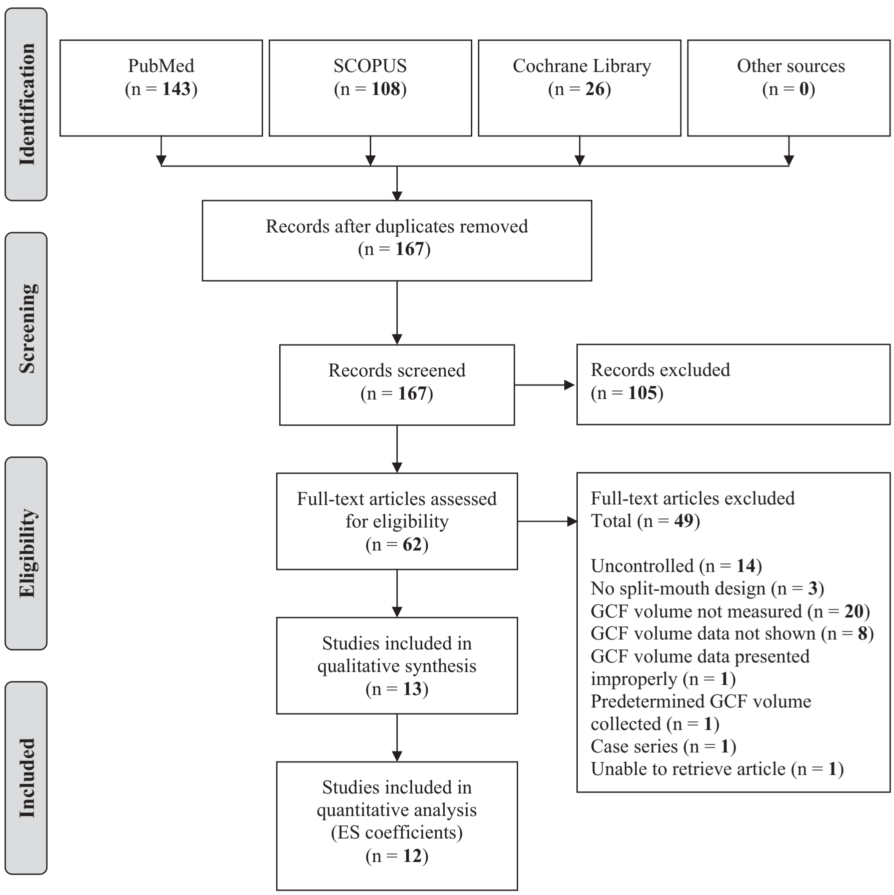

Fig. 1. Flow diagram of the search strategy.

automatic search, a total of 167 articles were retrieved. Among these, 13 studies (15–27) were judged to be relevant to this study according to the inclusion/exclusion criteria. Finally, in the manual search, no more relevant studies were included, while only one article (28) could not be retrieved on internet search, through the local library facility, and after having contacted the editorial office of the journal. The full details of the studies included are summarized in Table 2.

Study designs and population

The 13 studies included comprised 11 CCTs (15–25) and 2 RCTs (26, 27). No systematic reviews or meta-analyses were retrieved.

Only one study (17) included specifically male subjects, while all the rest monitored both sexes. The ages of the subjects were variable among the studies, but relatively narrow within the studies. These investigations included mainly adolescents (25) and adults (19, 23, 24, 26), or with both age ranges in a single group (15, 16, 18, 20, 22, 27). Two studies (17, 21) included two groups of adolescent and adult subjects, defined as juvenile and adult, respectively. Overall, the youngest and oldest groups had mean ages of 11 (17) and 31 (21) years, respectively. In all of the studies, the subjects were systemically and periodontally healthy and had never undergone any previous orthodontic treatment.

Teeth investigated, and treatments and experimental terms

Most of the studies (15, 19, 21–27) used a maxillary canine as the test tooth, and the contralateral and/or antagonist as the control teeth. In all of these studies, the test canine was moved distally (after premolar extraction) with the exception of one investigation (24), in which the test canine was monitored during the aligning and levelling phases. Other studies (16, 18, 20) used a maxillary first molar as the test tooth (again, moved distally) and the contralateral and antagonist teeth as controls. In only one study (17), a maxillary lateral incisor (tipped labially) was chosen as the test tooth, with the contralateral used as the control.

When the OTM was obtained by means of elastic chains (15, 19, 21–23, 25), these all exerted 250 gf. When nickel–titanium (Ni-Ti) coil springs were used, these exerted 250 gf (16, 18, 20), 150 gf (27) or both 50 gf and 150 gf on two contralateral test teeth (26). Finally, in the last two studies, the orthodontic forces were obtained by the use of nitinol archwires that exerted controlled forces of 70 cN (17), and copper Ni-Ti archwires (24), with this last not reporting the forces exerted. However, in only three studies (17, 26, 27), the force exerted by the orthodontic appliance was actually measured using a force gauge.

The inclusion of at least a control group wearing a passive orthodontic appliance was used in seven studies (17, 18, 20, 21, 23, 25, 27). None of the studies measured the degree of residual forces at the end of the experimental term.

These studies included those in which the experimental term followed an initial phase of aligning and levelling of both the test and control teeth (21, 23, 25). As no clinically detectable movement was seen in these control teeth at the end of the experimental term, their appliances were considered as passive, even when not clearly specified by the authors. In one study (16), a treated control group was subjected to active treatment (aligning and levelling phase), while two further studies (24, 26) only had a control group with no appliance. In the last three studies (15, 19, 22), incomplete information did not allow clear classification of the activation of the orthodontic appliance during the experimental term.

The experimental terms were generally short, with most being up to 1 week, that is, 168 h (15, 19, 21–25), with longer experimental terms of up to 21 days (20, 27), 28 days (16, 18) and 2 months (26) also used. One study (17) was limited to changes that occurred within 24 h of the application of the forces. All of the studies included an early time point after the application of the forces, either at 1 h or 1 day (or both).

Collection sites of GCF and methods of volume determination

For the test teeth, in eight studies (15, 17, 19, 21–23, 25, 26), the GCF was collected specifically from compression sites, while in four other studies (16, 18, 20, 27), it was collected from both tension and compression sites. In the last study (24), no indication was given about the type of stress exerted on the periodontium from which the GCF was collected.

All of the studies made use of paper strips to collect the GCF, with no other methods used. With the exception of three studies (16, 18, 20) in which endodontic paper points (#30) were used, the rest used PerioPapers. In five studies (16–18, 20, 27), a single sampling per site was performed, with one study having two sampling sites per experimental tooth (17); in seven studies (15, 19, 21–24, 26), two consecutive samplings were performed per site (and then pooled). In one study (25), a repetition of four samplings per site was followed. The sampling time was 30 s (15–20, 25–27) or 1 min (21–24). The time interval between repeated samplings was 1 min (15, 19, 21–25) or 90 s (26). The GCF volume was determined using a Periotron in all of the studies (versions 6000 or 8000), with the exception of three (16, 18, 20), which used an analytical balance for weight determination, with the volume derived assuming 1.0 g/ml as the specific weight of GCF.

Clinical monitoring

In nine of the studies (15, 16, 18–23, 25), the mean degree of OTM of the test teeth was reported, which ranged from 0.82 ± 0.22 mm (168 h; 21) to 1.9 ± 0.5 mm (28 days; 18). In the rest of the studies (17, 24, 26, 27), the degree of OTM of the test teeth was not reported. In all of the studies using a control tooth wearing an active orthodontic appliance (16), or with a passive appliance where the experimental term began after a previous aligning and levelling phase (15, 18–21, 23, 25), the degree of OTM of these teeth was not detectable. In all other cases, the control teeth underwent no OTM, as they did not wear an orthodontic appliance or the appliance was passive and no previous aligning and levelling phase was performed.

The periodontal conditions of all of the experimental teeth (recorded as plaque accumulation, bleeding on probing and probing depth) were optimal throughout the experimental terms in 10 studies (15, 17, 19, 21–27). On the other hand, in three studies (16, 18, 20), they worsened slightly during the observational term in the test and contralateral control teeth wearing appliances. However, in these studies, the antagonist control teeth that were not wearing any appliance, showed optimal clinical conditions during the study term.

GCF volumes

The GCF volumes recorded were presented as means ± SD in all of the studies, with the exception of two, in which means ± SE (15) and medians (17) were reported. This last study also merged the GCF volumes of the test and control teeth. In some studies (16, 18, 20), overall GCF volumes recorded over time within each experimental group were reported. In other studies (21, 24–27), the full data were shown. The overall means ± SDs were calculated and are reported in Table 2. In the rest of the studies (15, 17, 19, 22), the data on GCF volume were shown only for a 24-h time point, independent of the length of the experimental term (see also Table 3). The GCF collected had means for all of the groups between 0.13 and 0.17 μl when collected by a single endodontic paper point per experimental site (16, 18, 20) and between 0.35 and 0.44 μl when collected twice per experimental site using PerioPapers (15, 19, 21–24, 26). In one study (25), up to 0.88 μl GCF was collected per site by pooling four consecutive samplings. In another study (17), where they used two PerioPapers at two different sites per experimental tooth, that is, not consecutive, the median GCF volumes were in the range of 0.10–0.17 μl. In the last study (27), in which the GCF was collected using a single PerioPaper in a single site per experimental tooth, the mean GCF volume ranged between 0.50 and 0.59 μl.

Main reported results and conclusions

Generally, irrespective of the studies, the GCF volumes were similar between the test and control teeth, with very few statistically significant differences. In eight studies (15, 19, 21–26), no

Table 3. **Quality evaluation of the 13 studies included in the present review**

Study	Design	Sample description	Treatment description	Control group with passive appliance	Method error (GCF volume)	Adequacy of statistical analysis	Previous estimate of sample size	GCF point measures and variability	Quality score	Judged quality standard
Uematsu et al. (15)	CCT	Partial	Full	No	No	No (parametric)	No	Partial	6	Low
Perinetti et al. (16)	CCT	Full	Full	No	No	Yes (parametric)	No	Partial	8	Low
Ren et al. (17)	CCT	Full	Partial	Yes	No	Yes (parametric)	No	Partial	8	Low
Perinetti et al. (18)	CCT	Full	Full	Yes	No	Yes (nonparametric)	No	Partial	9	Medium
Sugiyama et al. (19)	CCT	Full	Partial	No	No	No (nonparametric)	No	Partial	6	Low
Perinetti et al. (20)	CCT	Full	Full	Yes	No	Yes (nonparametric)	No	Partial	9	Medium
Kawasaki et al. (21)	CCT	Full	Full	Yes	No	Yes (parametric)	No	Full	10	Medium
Nishijima et al. (22)	CCT	Full	Partial	No	No	No (nonparametric)	No	Partial	6	Low
Yamaguchi et al. (23)	CCT	Full	Full	Yes	No	No (nonparametric)	No	Partial	8	Medium
Yamaguchi et al. (24)	CCT	Full	Partial	No	No	Yes (parametric)	No	Full	8	Low
Dilsiz et al. (25)	CCT	Full	Full	Yes	No	Yes (nonparametric)	No	Full	10	Medium
Luppanapornlarp et al. (26)	RCT	Full	Full	No	No	No (not specified)	No	Full	9	Medium
Drummond et al. (27)	RCT	Full	Full	Yes	No	Yes (parametric)	No	Full	11	Medium/high

statistically significant differences were seen for any of the comparisons (including comparisons between experimental teeth within each time point, or over time within each experimental tooth). In one study, similar results were seen for the juvenile group of subjects, while for the adult group, there was a modest but significant increase in GCF volume from 0.17 μl to 0.20 μl (as medians) from baseline to 24 h, respectively (pooled data for experimental and control teeth). In three studies (16, 18, 20), the only statistically significant differences in the GCF volumes were seen between the teeth wearing an orthodontic appliance (either passive or active) and the teeth not wearing an appliance. In the last study (27), the test and control teeth were both wearing orthodontic appliances, and they showed some modest, but statistically significant, increases in GCF volumes over the study term. However, no significant differences were seen between the experimental teeth at each time point, nor between the mesial (tension) and distal (compression) sites of the test tooth.

In seven studies (15, 17, 19, 21–24), the results regarding the GCF volume were not discussed. In three studies (16, 18, 20), the increases in GCF volume were ascribed to the periodontal inflammation that was consequent to the fixed appliance placement, rather than to OTM *per se*. In the last three studies where the periodontal conditions were optimal throughout the experimental terms (25–27), the GCF volumes were concluded to have been influenced by periodontal inflammation (even at a subclinical level) and not by OTM.

Quality analysis and risk of bias in individual studies

The results of the quality analysis are shown in Table 3. The quality was medium/high in one study (27), medium in six studies (16–18, 20, 21, 23, 25, 26) and low in all of the other six studies (15–17, 19, 22, 24).

The selection description was classified as adequate (full) in 12 studies (16–27), and as not adequate (partial) in one study (15; Table 4). All of the studies extensively detailed the age, sex and systemic and periodontal health of the subjects

Table 4. **Reappraisal of 12 study reports through calculation of the highest and lowest effect size coefficients and the corresponding percentage variations in the comparisons and other related information**

Study	HighestES	Variation (%)	Compared teeth (time point)	Lowest ES	Variation (%)	Comparisons/time point
Highest ES coefficient ≥ 1.0						
Uematsu et al. (13)	1.46	17.1	Test vs. contralateral* (24 h)	0.40	5.7	Contralateral vs. antagonist† (24 h)
Perinetti et al. (16)	1.18‡	21.4	Contralateral* vs. antagonist† (overall)	0.28	6.3	Test vs. contralateral* (overall)
Perinetti et al. (18)	1.17‡	21.3	Contralateral* vs. antagonist† (overall)	0.28	6.3	Test vs. contralateral* (overall)
Nishijima et al. (20)	1.00	12.8	Test vs. contralateral/antagonist* (24 h)	NA	–	–
Highest ES coefficient ≥0.20 and <1.00						
Drummond et al. (25)	0.57	–22.0	Test vs. contralateral*, distal sites (1 h)	0.04	–2.0	Test vs. contralateral*, mesial sites (24 h)
Yamaguchi et al. (22)	0.53	10.5	Test (high-friction) vs. antagonist† (24 h)	0.0	0.0	Test (low-friction) vs. antagonist† (all time points)
Luppanapornlarp et al. (24)	0.41	–34.4	Test (50 gf) vs. antagonist† (1 h)	0.03	2.3	Test (150 gf) vs. antagonist† (1 h)
Kawasaki et al. (19; juvenile group)	0.40	13.5	Test vs. contralateral/antagonist* (24 h)	0.27	12.5	Test vs. contralateral/antagonist* (168 h)
Perinetti et al. (14)	0.29	15.4	Contralateral§ vs. antagonist† (overall)	0.00	0.0	Test vs. contralateral§ (overall)
Kawasaki et al. (19; adult group)	0.28	28.2	Test vs. contralateral/antagonist* (1 h)	0.22	15.8	Test vs. contralateral/antagonist* (24 h)
Sugiyama et al. (17)	0.22	2.7	Test vs. contralateral* (24 h)	NA	–	–
Yamaguchi et al. (21)	0.15	2.6	Test vs. contralateral* (overall)	NA	–	–
Dilsiz et al. (23)	0.07	2.3	Test vs. contralateral* (1 h and 168 h)	0.0	0.0	Test vs. contralateral† (24 h)

For each of the studies analysed, the highest and lowest effect size (ES) seen in all of the possible comparisons is shown. Variation, corresponding to the ES shown and calculated on the smaller of the two means which were compared. Negative variation values indicate a lower GCF volume for the test teeth.
NA, not applicable.
Control teeth that during the study period: *wore a passive fixed appliance, or †had no appliance, or §were subjected to orthodontic forces.
‡Difference between the corresponding mean values that were statistically significant.

included. Similarly, in all of the studies, the lack of any antibiotics or anti-inflammatory therapies prior to and during the experimental terms was clearly reported for all of the subjects. In the only study with a sample description classified as partial (15), the lack of information regarded the age range, that is, standard deviation, of the subjects.

The treatment description was classified as full in 10 studies (15, 16, 18–21, 23, 25–27). In particular, in one of these studies (27), clinically detectable movement of the test teeth was seen and recorded, although not shown. The treatment description was classified as partial in the other studies. The incomplete treatment descriptions related to potential clinically detectable OTM (17,

24) or the intensity of the force applied to the test teeth (24). In one study (22), whether or not an initial aligning and levelling phase was performed prior to the beginning of the experimental term was neither specified nor derivable. In only seven studies (17, 18, 20, 21, 23, 25, 27), did the control teeth wear a passive orthodontic appliance, while none of the studies included a method error analysis for GCF volume determination.

The statistical methods were judged appropriate in eight studies (16–18, 20, 21, 24, 25, 27). In contrast, in five studies, the statistical analysis was judged inappropriate, because of the use of parametric tests when nonparametric tests would have been more appropriate (derived from the high asymmetrical distributions of the data sets when no more information, that is, data transformation, was provided; 26), or because the multiple nature of the comparisons was not taken into account and corrections of the *p* values in pairwise comparisons were not declared (15). Finally, in three studies (19, 22, 23), tests for comparing independent data sets were used in spite of the paired nature of the compared groups' data.

None of the studies used previous estimates of sample size. Point measures and measures of variability were reported in full in five studies (21, 24, 25, 27) and partially in the rest of the studies.

Regarding the risk of bias assessment in individual studies, no information was provided on the concealment of randomization in the two RCTs (26, 27), and none of the studies reported a blindness of the outcome assessor.

Effects-size coefficients of GCF volume

For four studies (16, 18, 20, 23), the ES coefficients were calculated on the overall GCF volumes of each experimental group. For four further studies (15, 17, 19, 22), the ES coefficients were derived only at a unique time point (according to the data provided in the articles), while for five studies (21, 24–27), the ES coefficients for each time point could be calculated. In one of these studies (27), the ES coefficients were also calculated for mesial and distal sites in the treated group, which corresponded to tension and compression sites, respectively. Finally, for one study, the ES coefficients could not be calculated as the outcomes were reported as medians, and no other data for this calculation were available (17).

A total of 69 ES coefficients were calculated. Among these, 32 (46.4%) were below 0.2; 33 (47.8%) were between 0.2 and 0.8; none was between 0.8 and 1.0; and only 4 (5.8%) were above 1.0. The results of the highest and lowest ES coefficient calculations for each of these 12 studies are shown in Table 4, along with all of the relative information about the OTM treatment and timing of the GCF collection.

Only four studies included in this reappraisal revealed a high ES coefficient of at least 1.0 (range, 1.0–1.46; 15, 18, 20, 22). The coefficients of variation were small for these studies and ranged from 12.8% (22) to 21.4% (18). In two of these studies (18, 20), the highest ES coefficients retrieved (1.18 and 1.17) related to comparisons between a contralateral control tooth wearing a passive appliance and the other antagonist control tooth not wearing any appliance; on the other hand, in the same studies, the lowest ES coefficients retrieved (0.28) related to comparisons between the (moved) test teeth and the (non-moved) contralateral teeth. In all of the other studies, the highest ES coefficient retrieved ranged from 0.07 (25) to 0.57 (27). The corresponding coefficients of variations were also variable, and ranging from −34.4% (meaning a lower GCF volume in the moved teeth; 26) to 28.2% (21). For the same studies, the lowest ES coefficients retrieved were extremely low and not relevant (Table 4).

In general, no clear behaviour for the magnitude of the ES coefficients could be seen with respect to the different time points, nor was there any significant correlation between the coefficients of variations and the corresponding ES coefficients (data not shown).

Discussion

This systematic review was undertaken to address the issue of whether or not the GCF volume is a useful diagnostic aid in monitoring OTM in a site-specific manner.

Studies recording the GCF volume during OTM generally reported very large inter-subject variability (15–27). This aspect would make results from studies, in which the control samples belong to different subjects, less reliable than those in which test and control samples belong to the same person. Moreover, increases in the GCF after the placement of a fixed orthodontic appliance may also be due to subclinical inflammation in the absence of detectable tooth movement (27). Therefore, studies using a split-mouth design, especially those with control sites wearing an appliance but not subjected to forces, and with a close monitoring of the periodontal conditions, have been selected for this systematic review.

Main findings

All of the studies saw little or no changes in the GCF volume incident to OTM, at least for the short experimental terms from 24 h to 2 months (Table 2). When statistically significant differences were seen in the experimental teeth, these were ascribed to the clinical or subclinical periodontal inflammation that was consequent to the placement of the fixed appliances. Of interest, the only study that could not be retrieved here is a very early one (28) that, using a controlled split-mouth design, also reported no significant changes in the GCF volume collected from orthodontically moved teeth and untreated controls. In particular, this information was retrieved from a narrative review article (29).

These results are consistent over the different appliances used to obtain the OTM with forces presumed to be up to 250 gf (Table 2). Indeed, with the exception of only three studies (17, 26, 27), in all of the rest the actual forces exerted by the appliances were presumed to be equal to those provided by the manufacturers being not measured *in situ*.

However, these results are reinforced by the optimal clinical conditions that were present in all of the experimental sites in most of the studies, with the exception of three studies (16, 18, 20), in which a correct control group was nevertheless provided. Therefore, the potential effects on the GCF volume of the gingival inflammation because of plaque accumulation rather than OTM were avoided or controlled in all of these studies. The measure of the GCF volume through current technology is thus not reliable in assessing tissue remodelling incident to OTM. Moreover, these findings argue against reporting constituents of GCF as volumetric concentrations instead of as total levels, while these biomarkers may eventually be expressed relative to total protein content (30, 31) or as ratios of different constituents (32).

Risk of bias at study/outcome and review level

Six of the studies included were of low quality, and only one was of medium/high quality (Table 3). However, a further important aspect was seen in the not completely achievable blinding in such studies, and for this reason, this criterion was not included in this quality analysis. However, the more recent studies were generally of better quality, as the total scores seen here were significantly correlated with the year of publication (Spearman's rho, 0.61, $p < 0.05$). As for the PRISMA risk of bias of individual studies, no proper information on the concealment of randomization was provided in the two RCTs (26, 27), nor a clear indication as to the outcome assessor blindness, referring to the operator in charge of measuring the GCF volume, was reported.

A further important aspect relates to the repeatability of the GCF sampling. In this regard, none of the studies included here performed any analysis for the repeatability of the GCF volume according to the sampling procedures used. A different study (33), however, has shown that when the sampling procedure is standardized, the GCF volume collected should have intra-subject consistency up to 5 weeks. Further studies that evaluate the repeatability of the GCF sampling in greater depth are warranted.

Most of the studies included in this systematic review used the Periotron to measure the GCF volume. The Periotron has been developed for both clinical and research purposes, and both the 6000 and 8000 versions have been reported to be reliable and convenient for measuring GCF volumes. However, for volumes <0.2 μl, the associated errors can be up to 20% (8). This error is of relevance when dealing with healthy periodontal conditions, such as with orthodontic patients in which the total sampled GCF volume is generally equal to or slightly below 0.2 μl/paper strip (see Table 2). Moreover, the coefficients of variation corresponding to the highest ES coefficients retrieved in this analysis were generally below 20% (Table 4). Of interest, none of the studies considered here included a method error analysis of the GCF volume determination. Therefore, whether or not the GCF volume is not sensitive to OTM or is sensitive at a level that remains below the detection capabilities of the measurement methods used to date remains an open issue.

Regarding limitations at the review level, full data on the GCF volume reporting were seen only in few studies (21, 24–27) with the rest reporting selective data for given time points or merging data among time points. Moreover, because of the paired nature of the data, the means and SDs of the GCF volumes reported for each experimental group of teeth were not enough to calculate the confidence intervals (CIs) of the ES coefficients. In spite of these limitations, the results obtained in the quantitative analysis provided herein clearly demonstrate the lack of sensitivity of the GCF volume to OTM.

General interpretation of the results in the context of other evidence

An important issue when dealing with GCF volume is the sampling procedure. Indeed, the use of paper strips allows the collecting of the resting GCF inside the crevice, which is referred as to the GCF volume. In contrast, the use of capillary tubing kept inside the crevice for several minutes is useful for the sampling of the GCF flow, which is a somewhat different entity (7). Interestingly, all of the studies included in this analysis used paper strips to collect GCF, reporting very small changes in terms of GCF volumes incident to OTM (Table 2). In contrast, other studies that used capillary tubing have reported notable increases in GCF flow rate (34–36). However, these studies and others could not be included in the present review as they did not have a split-mouth design (34, 37) or they were uncontrolled (35, 36). From this point of view, further evidence is necessary to establish whether or not GCF flow, rather than GCF volume, might be sensitive to OTM. However, the discomfort of the capillary tubing in sampling the GCF (7) might significantly limit its application in a clinical setting.

A further important issue relates to the features of any diagnostic tool, which has to provide measurement outcomes that can be considered accurate (i.e. with high sensitivity and specificity), in terms of the presence/absence of a given condition. However, a difficulty in reappraising previous data resides in the variability of the number of subjects enroled in the studies examined, rendering the corresponding power of the statistical tests also poorly comparable (38). Moreover, none of the studies included in this analysis carried out a prior estimation for the sample size, and thus, whether or not a lack of statistical power has caused the lack of statistical significance remains an open issue. Therefore, a critical approach to assess the relevance of measurements of the GCF volume as a diagnostic aid in orthodontics has to rely on highly accurate measurements that are recorded at two different teeth, one moved orthodontically and the other not, whereby the GCF volume should show recordable changes when compared to the corresponding variances. Indeed, a low ratio here would be responsible for low sensitivity and specificity, indicating the poor diagnostic performance of a given tool. A statistical approach to quantify this ratio (taking into account the sizes of the study populations) is provided by the calculation of the ES coefficient (14), which has been used as an index of diagnostic performance (39, 40). However, the present ES coefficients were generally low (Table 4), that is, below the threshold of 1.0, and whether or not those above this value would remain as such even considering the corresponding CIs is not proven here.

These results from the retrieved ES coefficients show that in spite of the different methods to achieve OTM and the different time points, the GCF volume *per se* does not have diagnostic potential as a biomarker of tissue remodelling incident to OTM. This result is highly consistent among the studies, with the highest ES coefficient of 1.46 (15; Table 4). Moreover, in two studies (18, 20), the corresponding highest ES coefficients retrieved were seen between two control teeth, the difference between which was the wearing or not of a passive fixed appliance. Therefore, the preferred method to monitor the process of OTM on an individual basis still remains based upon the clinical parameters, that is, degree of tooth displacement, radiographical detection of root/-bone resorption, probing depth or subjective pain. However, these methods remain limited by their failure in taking into account the full biological response of the tissues to forces, as they do not have predictive value in term of tissue damage. In this view, current evidence has shown that the quantification of the several GCF biomarkers, such as those for tissue inflammation (41), necrosis (18, 20) or bone metabolism (42), has potential for the successful monitoring of the different aspects of tissue response incident to OTM.

Conclusions

Using current methods for sampling and measuring, the GCF volume is not a reliable index for tissue remodelling incident to OTM.

Clinical relevance

Biological monitoring of orthodontic tooth movement through analyses of GCF has been advocated for the setting of the optimum force magnitude *in vivo*. However, using current procedures, the measuring of GCF volume does not provide a diagnostic index of tissue remodelling during orthodontic treatment.

Acknowledgements: The authors are grateful to Dr. Christopher Paul Berrie (Telethon Institute for Genetics and Medicine, Naples, Italy) for critical appraisal of the manuscript. This study was funded by the University of Trieste, and no further external source of founding was involved.

References

1. Ren Y, Maltha JC, Kuijpers-Jagtman AM. Optimum force magnitude for orthodontic tooth movement: a systematic literature review. *Angle Orthod* 2003;73:86–92.
2. Pashley DH. A mechanistic analysis of gingival fluid production. *J Periodontal Res* 1976;11:121–34.
3. Goodson JM. Gingival crevice fluid flow. *Periodontol 2000* 2003;31: 43–54.
4. Alfano MC. The origin of gingival fluid. *J Theor Biol* 1974;47:127–36.
5. Griffiths GS, Sterne JA, Wilton JM, Eaton KA, Johnson NW. Associations between volume and flow rate of gingival crevicular fluid and clinical assessments of gingival inflammation in a population of British male adolescents. *J Clin Periodontol* 1992;19:464–70.
6. Krishnan V, Davidovitch Z. Cellular, molecular, and tissue-level reactions to orthodontic force. *Am J Orthod Dentofacial Orthop* 2006;129:469 e461–432.
7. Griffiths GS. Formation, collection and significance of gingival crevice fluid. *Periodontol 2000* 2003;31: 32–42.
8. Chapple IL, Landini G, Griffiths GS, Patel NC, Ward RS. Calibration of the Periotron 8000 and 6000 by polynomial regression. *J Periodontal Res* 1999;34:79–86.
9. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–34.
10. Perinetti G, Turp JC, Primozic J, Di Lenarda R, Contardo L. Associations between the masticatory system and muscle activity of other body districts. A meta-analysis of surface electromyography studies. *J Electromyogr Kinesiol* 2011;21:877–84.
11. Perinetti G, Contardo L. Posturography as a diagnostic aid in dentistry: a systematic review. *J Oral Rehabil* 2009;36:922–36.
12. Centre for Reviews and Dissemination Systematic reviews. *CRD*'s Guidance for Undertaking Reviews in Healthcare. York: University of York NHS Centre for Reviews & Dissemination; 2009.
13. Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized control trials in dental research. I. Methods. *J Periodontal Res* 1986;21:305–14.
14. Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
15. Uematsu S, Mogi M, Deguchi T. Interleukin (IL)-1 beta, IL-6, tumor necrosis factor-alpha, epidermal growth factor, and beta 2-microglobulin levels are elevated in gingival crevicular fluid during human orthodontic tooth movement. *J Dent Res* 1996;75:562–7.
16. Perinetti G, Paolantonio M, D'Attilio M, D'Archivio D, Tripodi D, Femminella B et al. Alkaline phosphatase activity in gingival crevicular fluid during human orthodontic tooth movement. *Am J Orthod Dentofacial Orthop* 2002;122:548–56.
17. Ren Y, Maltha JC, Van't Hof MA, Von Den Hoff JW, Kuijpers-Jagtman AM, Zhang D. Cytokine levels in crevicular fluid are less responsive to orthodontic

force in adults than in juveniles. *J Clin Periodontol* 2002;29:757–62.
18. Perinetti G, Paolantonio M, D'Attilio M, D'Archivio D, Dolci M, Femminella B et al. Aspartate aminotransferase activity in gingival crevicular fluid during orthodontic treatment. A controlled short-term longitudinal study. *J Periodontol* 2003;74:145–52.
19. Sugiyama Y, Yamaguchi M, Kanekawa M, Yoshii M, Nozoe T, Nogimura A et al. The level of cathepsin B in gingival crevicular fluid during human orthodontic tooth movement. *Eur J Orthod* 2003;25:71–6.
20. Perinetti G, Serra E, Paolantonio M, Brue C, Meo SD, Filippi MR et al. Lactate dehydrogenase activity in human gingival crevicular fluid during orthodontic treatment: a controlled, short-term longitudinal study. *J Periodontol* 2005;76:411–7.
21. Kawasaki K, Takahashi T, Yamaguchi M, Kasai K. Effects of aging on RANKL and OPG levels in gingival crevicular fluid during orthodontic tooth movement. *Orthod Craniofac Res* 2006;9:137–42.
22. Nishijima Y, Yamaguchi M, Kojima T, Aihara N, Nakajima R, Kasai K. Levels of RANKL and OPG in gingival crevicular fluid during orthodontic tooth movement and effect of compression force on releases from periodontal ligament cells in vitro. *Orthod Craniofac Res* 2006;9:63–70.
23. Yamaguchi M, Yoshii M, Kasai K. Relationship between substance P and interleukin-1beta in gingival crevicular fluid during orthodontic tooth movement in adults. *Eur J Orthod* 2006;28:241–6.
24. Yamaguchi M, Takizawa T, Nakajima R, Imamura R, Kasai K. The damon system and release of substance p in gingival crevicular fluid during orthodontic tooth movement in adults. *World J Orthod* 2009;10:141–6.
25. Dilsiz A, Kilic N, Aydin T, Ates FN, Zihni M, Bulut C. Leptin levels in gingival crevicular fluid during orthodontic tooth movement. *Angle Orthod* 2010;80:504–8.
26. Luppanapornlarp S, Kajii TS, Surarit R, Iida J. Interleukin-1beta levels, pain intensity, and tooth movement using two different magnitudes of continuous orthodontic force. *Eur J Orthod* 2010;32:596–601.
27. Drummond S, Canavarro C, Perinetti G, Teles R, Capelli J Jr. The monitoring of gingival crevicular fluid volume during orthodontic treatment: a longitudinal randomized split-mouth study. *Eur J Orthod* 2011;34:109–13.
28. Tersin J. Studies of gingival conditions in relation to orthodontic treatment. V. A comparison between the gingival exudation in orthodontically activated teeth and non-activated teeth. *Swed Dent J* 1978;2:137–9.
29. Kavadia-Tsatala S, Kaklamanos EG, Tsalikis L. Effects of orthodontic treatment on gingival crevicular fluid flow rate and composition: clinical implications and applications. *Int J Adult Orthodon Orthognath Surg* 2002;17:191–205.
30. Iwasaki LR, Haack JE, Nickel JC, Reinhardt RA, Petro TM. Human interleukin-1 beta and interleukin-1 receptor antagonist secretion and velocity of tooth movement. *Arch Oral Biol* 2001;46:185–9.
31. Perinetti G, Franchi L, Castaldo A, Contardo L. Gingival crevicular fluid protein content and alkaline phosphatase activity in relation to pubertal growth phase. *Angle Orthod* 26 Mar 2012, Epub ahead of print, doi: 10.2319/123111-123806.123111.
32. Iwasaki LR, Chandler JR, Marx DB, Pandey JP, Nickel JC. IL-1 gene polymorphisms, secretion in gingival crevicular fluid, and speed of human orthodontic tooth movement. *Orthod Craniofac Res* 2009;12:129–40.
33. Chapple IL, Socransky SS, Dibart S, Glenwright HD, Matthews JB. Chemiluminescent assay of alkaline phosphatase in human gingival crevicular fluid: investigations with an experimental gingivitis model and studies on the source of the enzyme within crevicular fluid. *J Clin Periodontol* 1996;23:587–94.
34. Last KS, Donkin C, Embery G. Glycosaminoglycans in human gingival crevicular fluid during orthodontic movement. *Arch Oral Biol* 1988;33:907–12.
35. Pender N, Samuels RH, Last KS. The monitoring of orthodontic tooth movement over a 2-year period by analysis of gingival crevicular fluid. *Eur J Orthod* 1994;16:511–20.
36. Samuels RH, Pender N, Last KS. The effects of orthodontic tooth movement on the glycosaminoglycan components of gingival crevicular fluid. *J Clin Periodontol* 1993;20: 371–7.
37. Tersin J. Studies of gingival conditions in relation to orthodontic treatment. IV. The effect of oral hygiene measures on gingival exudation during the course of orthodontic treatment. *Swed Dent J* 1978;2:131–6.
38. Phillips C. Sample sizes and power: what is enough? *Semin Orthod* 2002;8:67–76.
39. Perinetti G, Baccetti T, Contardo L, Di Lenarda R. Gingival crevicular fluid alkaline phosphatase activity as a non-invasive biomarker of skeletal maturation. *Orthod Craniofac Res* 2011;14:44–50.
40. Perinetti G, Baccetti T, Di Leonardo B, Di Lenarda R, Contardo L. Dentition phase and chronological age in relation to gingival crevicular fluid alkaline phosphatase activity in growing subjects. *Prog Orthod* 2011;12:100–6.
41. Ren Y, Vissink A. Cytokines in crevicular fluid and orthodontic tooth movement. *Eur J Oral Sci* 2008;116:89–97.
42. Yamaguchi M. RANK/RANKL/OPG during orthodontic tooth movement. *Orthod Craniofac Res* 2009;12:113–9.