Bayesian network meta-analysis of root coverage procedures: ranking efficacy and identification of best treatment


Abstract
Aims: The aim of this work was to conduct a Bayesian network meta-analysis (NM) of randomized controlled trials (RCTs) to establish a ranking in efficacy and the best technique for coronally advanced flap (CAF)-based root coverage procedures.

Material and Methods: A literature search on PubMed, Cochrane libraries, EMBASE, and hand-searched journals until June 2012 was conducted to identify RCTs on treatments of Miller Class I and II gingival recessions with at least 6 months of follow-up. The treatment outcomes were recession reduction (RecRed), clinical attachment gain (CALgain), keratinized tissue gain (KTgain), and complete root coverage (CRC).

Results: Twenty-nine studies met the inclusion criteria, 20 of which were classified as at high risk of bias. The CAF+connective tissue graft (CTG) combination ranked highest in effectiveness for RecRed (Probability of being the best = 40%) and CALgain (Pr = 33%); CAF+enamel matrix derivative (EMD) was slightly better for CRC; CAF+Collagen Matrix (CM) appeared effective for KTgain (Pr = 69%). Network inconsistency was low for all outcomes excluding CALgain.

Conclusion: CAF+CTG might be considered the gold standard in root coverage procedures. The low amount of inconsistency gives support to the reliability of the present findings.

Key words: coronally advanced flap; gingival recession; network meta-analysis; randomized controlled trials; root coverage

Accepted for publication 9 September 2012

Health-care decisions should be based on the best available evidence. Meta-analysis of randomized controlled trials (RCTs) on direct, head-to-head, comparisons of treatments represents a well-established method for summarising results from a set of different trials and can be considered a reliable approach for evaluating the effectiveness of clinical treatments/interventions (Liberati et al. 2009). In therapy, standard meta-analytical methods summarize evidence from studies which directly compare two types of interventions (i.e. treatment A versus treatment B). When there are multiple treatment options, the number of possible pairwise comparisons increases factorially. Thus, direct comparisons of specific treatments or regimens of interest may not be available and a lack of all possible comparisons among the several treatment options is frequently recognized in the literature.

Network meta-analysis (NM) is a new approach to meta-analysis: direct comparisons of treatments tested in single trials are not summarized in separate meta-analyses, but evidence is combined in a connected network of interventions which
includes both direct and indirect comparisons among different trials. The first attempt to combine direct and indirect inference using data from different sources of evidence was presented by Eddy et al. (1990), who proposed a Bayesian approach to indirect comparison of treatments (confidence profile method – CPM). NM may be considered an evolution of the CPM approach and different models have been recently proposed (Gleser & Olkin 1994, Higgins & Whitehead 1996, Lumley 2002, Psaty et al. 2003, Song et al. 2003, Lu & Ades 2004, 2006, Glenny et al. 2005, Salanti et al. 2008, Buti et al. 2011).

Along with the between-trials heterogeneity (Higgins & Whitehead 1996), NM shows an additional source of variability that is called “inconsistency” or “incoherence.” Inconsistency is the uncertainty due to discrepancies between direct and indirect inference on the pair-wise comparisons (Lumley 2002). For instance, the direct comparison of two treatments may produce a relative effect estimate which is different from that obtained through indirect comparisons (e.g., the direct estimate \( d_{AC} \) does not equal the indirect estimate \( d_{AB} \)) (Salanti et al. 2008). Different methods for evaluating inconsistency of NM have been proposed (Lumley 2002, Lu & Ades 2004, 2006, Dias et al. 2010).

Applications of NM models have been published in medical journals (Bridle et al. 2003, Psaty et al. 2003, Wilby et al. 2005, Lu & Ades 2006, Welton et al. 2009, Orme et al. 2010). In the area of the dental sciences, a literature search has revealed only a few applications of NM (Tu et al. 2010, 2012, Walsh et al. 2010, Faggion et al. 2011).

Several different surgical procedures for the treatment of gingival recessions have been proposed over the past few decades (Roccuzzo et al. 2002, Clauser et al. 2003, Cheng et al. 2007). Cairo et al. (2008) showed that connective tissue grafts (CTGs) or enamel matrix derivative in conjunction with a coronally advanced flap (CAF) enhances the probability of obtaining complete root coverage (CRC) in Miller Class I and II single gingival recessions. Chambrone et al. (2010, 2012) concluded that subepithelial CTGs may be more appropriate in cases where root coverage and gain in keratinized tissue are the main goals.

Nevertheless, these systematic reviews were not able to compare all possible treatment alternatives, to establish a ranking in efficacy and to assess the best treatment choice.

The purpose of this study was to conduct a Bayesian NM with the following specific aims: (1) to combine in a network the published information on comparisons among different root coverage procedures based on CAF or combined techniques to make inference even on comparisons which have not been conducted yet; (2) to establish a Ranking in efficacy of the treatment options; and (3) to identify the Best approach, in terms of recession reduction (RecRed), clinical attachment gain (CALgain), keratinized tissue gain (KTgain), and CRC.

For this purpose, a Bayesian extension of the NM model developed by Lumley (2002) was presented and simple statistical tools for assessing the network inconsistency were applied (Buti 2011).

Material and Methods
The systematic review was conducted according to the PRISMA statement instruments for systematic reviews (Liberati et al. 2009).

Criteria for considering studies for this review

Only RCTs of at least 6 months’ duration were considered for inclusion in this review and then organized by the PICO method (Glossary of Evidence-Based Terms 2007), according to the following points:

(P) Types of participants
Patients with a clinical diagnosis of Miller Class I or II localized gingival recession defect who were surgically treated by root coverage procedures.

(I) Types of interventions
The following surgical procedures for the treatment of single recessions were considered:

- CAF,
- CAF plus connective tissue graft (CAF+CTG),
- CAF plus barrier membrane (CAF+BM),
- CAF plus enamel matrix derivative (CAF+EMD),
- CAF plus acellular dermal matrix (CAF+ADM),
- CAF plus platelet rich plasma (CAF+PRP),
- CAF plus human fibroblast-derived dermal substitute (CAF+HF-DDS),
- CAF plus collagen matrix (CAF+CM).

Randomized controlled trials comparing CAF with multiple combinations (i.e. CAF+CTG+EMD, CAF+BM+CTG, CAF+BM+EMD, etc.) or RCTs comparing variations of the same technique (i.e. CAF with releasing incisions versus CAF without releasing incisions) were not included in this systematic review.

(C) Comparison between interventions

All possible comparisons among the included surgical procedures were investigated.

(O) Type of outcome measures
The following outcome measures were considered:

- Recession reduction (RecRed): change in gingival recession expressed as RecRed (mm) at follow-up visit;
- Clinical attachment level gain (CAL gain): change (mm) in clinical attachment level at follow-up visit;
- Keratinized tissue gain (KT gain): change (mm) in width of keratinized tissue at follow-up visit;
- Complete root coverage (CRC): recession defects that obtained CRC.

Search methods for identification of studies

The identification of RCTs to be included or considered in this NM was conducted via detailed search strategies. Database searching, originally conducted up to and including August 2007 (Cairo et al. 2008), was updated up to June 2012 using three electronic evidence sources:

The Cochrane Central Register of Controlled Trials (Clinical Trials), on 01.06.2012, using the following strategy: “Gingival Recession” [Search All Text] AND “Root Coverage” [Search All Text].

The Embase, on 01.06.2012, using the following strategy: gingival AND recession AND root AND coverage AND (controlled clinical trial)/lim OR [randomized controlled trial]/lim AND [humans]/lim

Limited to: controlled clinical trial from Evidance Based Medicine; randomized controlled trial from Evidance Based Medicine; humans.

Hand searching included a complete search of the Journal of Clinical Periodontology, the Journal of Periodontology, the Journal of Periodontal Research, and the International Journal of Periodontics and Restorative Dentistry up to 01.06.2012.


Selection of studies, assessment of validity, data extraction, and management

Details regarding screening of titles, abstracts, and full texts of articles published until August 2007 were previously reported (Cairo et al. 2008). The identification of studies conducted from August 2007 to June 2012 was similarly performed by two independent reviewers (J. B. and M. L.) who screened titles, abstracts, and full texts of the articles. Disagreement between the two reviewers was solved by discussion with the attendance of another author (M. N.).

At this point, data were independently extracted and entered into a computer by two review authors (J. B. and M. L.) using specifically designed data-collection forms. Patient characteristics, treatments, clinical outcomes, and study quality were systematically registered. When clinical data on one or more of the outcome variables were not published/reported in the original article, the authors of the RCT were contacted and kindly asked to send their raw data (non-published data) for inclusion in the statistical model. In case of missing data or lack of answer by the authors after two contact attempts, RCTs were considered not eligible for inclusion in this review.

When several articles were published for the same sample at different follow-up durations, the article reporting the longest follow-up duration was considered; however, if the number of withdrawals and drop-outs were ≥30% of the original sample, data from the previous article with a shorter follow-up were used.

Risk of bias assessment

The quality assessment of the included studies conducted up to and including August 2007 (for details see Cairo et al. 2008) was updated up to June 2012. This evaluation was performed independently and in duplicate form by two review authors (J. B. and M. L.).

Bayesian network meta-analysis (NM)

A Bayesian NM model was constructed for each of the four outcome variables RecRed, CALgain, KTgain, and CRC allowing for the inclusion of all the possible treatment comparisons. Direct comparisons of treatments as well as indirect comparisons were analysed in the same framework. The NM model applied in this review is a Bayesian extension of the hierarchical random-effects model proposed by Lumley (2002) for networks of two-arm trials (Buti 2011):

\[
Y_{jhb} \sim N(d_{hb} + u_j + w_{hk}, \sigma_{jhb}^2),
\]

\[
u_j \sim N(0, \tau^2),
\]

\[w_{hk} \sim N(0, \tau^2_w),\]

where: \( j = \text{study} \); \( k,b = \text{treatments} \); \( Y_{jhb} = \text{estimate of the effect of treatment} \ k \text{ when compared with} \ b \text{ in the} \ j \text{th trial} \); \( \sigma_{jhb} = \text{estimated standard error of} \ Y_{jhb}; \ d_{hb} = \text{average effect of treatment} \ k \text{ when compared with} \ b \); \( u_j = \text{random term from a Normal distribution with mean equal to} 0 \text{ and variance} \tau^2 \text{ for the} j \text{th trial} \). This term captures the heterogeneity among trials; \( w_{hk} = \text{random term from a Normal distribution with mean equal to} 0 \text{ and variance} \tau^2_w \text{ for the comparison between treatment} k \text{ and} b \). This term captures the inconsistency of this specific comparison with the rest of the evidence.

The random terms in the model were assumed to be mutually independent. The relative treatment effects \( d_{hb} \) were expressed in terms of mean differences for the continuous outcomes (RecRed, CALgain, and KTgain) and in terms of log odds ratios for the dichotomous outcome (CRC).

Non-informative vague priors for all the hyper-parameters in the model were specified so that whatever information is contained in the prior, the data should dominate it: Normal distributions with large variance \( N(0, 10^4) \) for the effect measures \( d_{hb} \) and Inverse Gamma distributions \( IG(10^{-3}; 10^{-3}) \) for \( \tau^2 \) and \( \tau^2_w \) (Lumley 2002, Lu & Ades 2004, 2006). Sensitivity analyses were performed by specifying on \( \tau \) and \( \tau^2_w \) Uniform distributions ranging from 0 to \( a \) (where \( a = 0.5, 1, 2 \)).

Single Bayesian meta-analysis (SM)

Bayesian pair-wise meta-analyses were performed permitting a comparison with Bayesian NM models. A random-effects model was specified in all cases except for those involving only two studies, when a fixed-effects model was preferred due to the little information on the heterogeneity variance. Non-informative priors were used throughout the analyses.

Estimation method and software

A Markov Chain Monte Carlo approach was used to obtain a sample

© 2012 John Wiley & Sons A/S
from the joint posterior distribution of the parameters for both SM and NM models. For SM models, three chains of 20,000 values were generated and a 6000-run burn-in was applied. For NM models, three chains of 3500,000 iterations were generated, then one sample of the five was used after a 50,000-run burn-in. The convergence of the chains was checked by the method of Gelman & Rubin (1992). The marginal distributions of the parameters of interest were summarized by the posterior mean, the median, and the 90% credibility interval, i.e. the 5th and 95th percentiles of the simulated values (Sterne & Davey Smith 2001). All the analyses were performed using WinBUGS software, version 1.4.3 (Spiegelhalter et al. 2003) and R software (Comprehensive R Archive Network - http://CRAN.R-Project.org).

Ranking and best

The Bayesian NM model allowed for establishing an efficacy Ranking among the tested treatments. The posterior distribution of the rank of each treatment was obtained and its median was calculated.

The probability that each of the root coverage procedures included in the analysis was the Best was also estimated (Lu & Ades 2004, 2006).

Network inconsistency

The between-trials heterogeneity was quantified by the variance $\tau^2$ as in standard pair-wise meta-analyses (Higgins & Thompson 2002, Edwards et al. 2009); a fixed-effect model resulted if $\tau^2 = 0$. The network inconsistency was expressed by the variance $\tau^2_w$. The sources of variability were evaluated by the posterior distribution of the proportion of variability due to the between-trials heterogeneity ($\widehat{I}^2$) and the posterior distribution of the proportion of variability due to the inconsistency ($\widehat{I}^2_w$) over the total variance (Higgins & Thompson 2002). Moreover, the inconsistency variance ($\tau^2_w$) was compared with the between-trials heterogeneity ($\tau^2$) and expressed in terms of probability that $\tau^2_w$ was larger than $\tau^2$, Pr($\tau^2 > \tau^2$) (Lu & Ades 2006).

The assessment of inconsistency affecting each single comparison was conducted on the basis of the posterior distribution of the random terms $w_{bh}$. When this distribution was centred far from 0, discrepancy between direct and indirect inference on the effect of treatment $k$ in respect to $b$ was present.

Results

Search results

The search results are presented in Fig. 1. The electronic search in MEDLINE (by PubMed), in the Cochrane Collaboration and in the Embase databases provided, respectively, 332, 142, and 125 articles published up to June 2012. The hand searching identified seven articles which could not be found during the electronic search. After the combination of the search results, 339 records were identified. Subsequently, after reading titles and abstracts, 248 articles were selected. Finally, full-text reading, manuscript review, and application of inclusion criteria were allowed to select 31 articles up to June 2012.

Study characteristics

Included studies

A total of 31 articles were detected and included in this study: 27 articles...
up to August 2007 (for details see Cairo et al. 2008) and four articles published between August 2007 and June 2012 which reported additional information on the following comparisons:

- In Mahajan et al. (2007), the comparison between CAF with CAF+ADM was evaluated and outcomes for RecRed and KTgain were included in the present NM;
- Abolfazli et al. (2009) compared CAF+CTG versus CAF+EMD: individual patient data (IPD) for RecRed, CALgain, KTgain, and CRC were obtained by the authors and re-analysed;
- McGuire & Scheyer (2010) compared CAF+CTG with CAF+CM: IPD for all the outcome variables were obtained by the authors and re-analysed;
- Cardaropoli et al. (2012) compared CAF+CTG with CAF+CM: IPD for RecRed, CALgain, KTgain, and CRC were obtained by the authors. However, 4 of the 18 subjects in the samples were excluded from NM because they participated in a parallel study with more than one defect of gingival recession per patient. Hence, IPD of the remaining 14 recession defects were re-analysed.

Excluded studies

A total of 53 studies did not fulfil the inclusion criteria: 40 studies (in agreement with Cairo et al. 2008) up to August 2007; 13 studies, published between August 2007 and June 2012, were excluded for the following reasons:

- Three studies considered comparisons of combinations of more than two techniques (Keceli et al. 2008, Cardaropoli & Cardaropoli 2009, Rasperini et al. 2011);
- Three studies reported comparisons with or between surgical techniques not investigated in the present systematic review (Jankovic et al. 2010, Lafzi et al. 2011, Nazareth & Cury 2011);
- One study reported a comparison between variations of a same surgical technique (Mazzocco et al. 2011);
- Two studies were not RCTs (Jha- veri et al. 2010, Nickles et al. 2010);
- Three studies evaluated more than one site for each technique in each patient (De Souza et al. 2008, Haghighati et al. 2009, Moslemi et al. 2011);
- One study showed inadequate statistical analysis (Banihashem-rad et al. 2009).

Methodological quality and risk of bias of the included studies

The quality of assessment for the 29 included studies is illustrated in Table 1. Twenty trials were classified as at high risk of bias, while nine as at low.

Results of the Analysis

The results of the analysis are reported separately for: (A) the continuous outcome variables (RecRed, CALgain, KTgain) and (B) the dichotomous outcome variable (CRC).

Recession reduction (RecRed), clinical attachment gain (CALgain), and keratinized tissue gain (KTgain)

Data from 26 of the 29 studies (for a total 31 articles) included in this systematic review were available for the analysis of RecRed, CALgain, and KTgain. Romagna-Genon (2001), McGuire & Nunn (2003) and Wilson et al. (2005) did not provide data for RecRed, CALgain, and KTgain; in the case of Côrtes et al. (2004, 2006), 6-month follow-up data were used for the analysis of RecRed, CALgain, and KTgain; in the case of Leknes et al. (2005) and Amarante et al. (2000), the 12-month follow-up data were considered as reported in Cairo et al. (2008). The treatment alternatives considered for the analysis were seven:

- CAF
- CAF+CTG
- CAF+BM
- CAF+EMD
- CAF+ADM
- CAF+PRP
- CAF+CM

Nine direct comparisons were based on data from RCTs (Fig. 2a):

- CAF versus CAF+CTG 2 RCTs
- CAF versus CAF+BM 2 RCTs
- CAF versus CAF+EMD 5 RCTs
- CAF versus CAF+ADM 3 RCTs for RecRed and KTgain; 2 RCTs for CAL gain
- CAF versus CAF+PRP 1 RCT
- CAF+CTG versus CAF+BM 6 RCTs
- CAF+CTG versus CAF+EMD 1 RCT
- CAF+CTG versus CAF+ADM 4 RCTs
- CAF+CTG versus CAF+CM 2 RCTs

Twelve comparisons were never directly tested in RCTs (Fig. 2b):

- CAF versus CAF+CM
- CAF+CTG versus CAF+PRP
- CAF+BM versus CAF+EMD
- CAF+BM versus CAF+ADM
- CAF+BM versus CAF+PRP
- CAF+EMD versus CAF+ADM
- CAF+EMD versus CAF+PRP
- CAF+EMD versus CAF+CM
- CAF+ADM versus CAF+CM
- CAF+PRP versus CAF+CM.

Seven independent SM and one NM model were conducted for each outcome variable (Table 2). The SM summarized evidence from the direct comparisons (Fig. 2a). Meta-analyses for CAF versus CAF+PRP and CAF+CTG versus CAF+EMD were not performed as there was only one available study for each of these pairwise comparisons. The NM model made it possible to include all the 21 potential pair-wise comparisons: the seven direct comparisons analysed by SM models, the CAF versus CAF+PRP and the CAF+CTG versus CAF+CM comparisons which could not be analysed by SM and the 12 indirect comparisons never tested in single trials (Fig. 2b).

RecRed

The results from the NM model qualitatively agreed with those from SM, even if the variability was larger (Table 2). When compared with CAF alone, the greatest mean differences for RecRed were achieved by the combined CAF+CTG treatment.
<table>
<thead>
<tr>
<th>Study</th>
<th>Tr. comparison</th>
<th>CRC</th>
<th>RecRed</th>
<th>CAL gain</th>
<th>KT gain</th>
<th>F-UP</th>
<th>Study design</th>
<th>All. conc.</th>
<th>Ex. blind.</th>
<th>Drop-outs</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>da Silva et al. (2004)</td>
<td>CAF + CTG vs. CAF</td>
<td>0.80 (1.15)</td>
<td>0.44 (0.27)</td>
<td>0.32 (0.38)</td>
<td>0.76 (0.29)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>No</td>
<td>No drop-out</td>
<td>High</td>
</tr>
<tr>
<td>Cortellini et al. (2009)</td>
<td>CAF + CTG vs. CAF</td>
<td>0.93 (0.45)</td>
<td>0.52 (0.23)</td>
<td>0.40 (0.22)</td>
<td>0.70 (0.25)</td>
<td>6</td>
<td>RCT, parallel study</td>
<td>Adequate</td>
<td>Yes</td>
<td>2 drop-outs</td>
<td>Low</td>
</tr>
<tr>
<td>Lins et al. (2003)</td>
<td>CAF vs. BM vs. CAF</td>
<td>–0.40 (0.29)</td>
<td>–0.50 (0.27)</td>
<td>0.50 (0.41)</td>
<td></td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>No</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Amarante et al. (2000) and Leknes et al. (2005)</td>
<td>CAF vs. CAF</td>
<td>–0.54 (0.65)</td>
<td>–0.20 (0.21)</td>
<td>–0.20 (0.24)</td>
<td>0.10 (0.15)</td>
<td>12</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>Yes</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Modica et al. (2000)</td>
<td>CAF vs. BM vs. CAF</td>
<td>0.00 (1.00)</td>
<td>0.90 (0.43)</td>
<td>0.90 (0.43)</td>
<td>0.60 (0.22)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>Yes</td>
<td>No drop-out</td>
<td>High</td>
</tr>
<tr>
<td>Spuhr et al. (2005)</td>
<td>CAF vs. EMD vs. CAF</td>
<td>1.32 (0.50)</td>
<td>0.28 (0.23)</td>
<td>0.49 (0.27)</td>
<td>0.317 (0.20)</td>
<td>24</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>No</td>
<td>No drop-out</td>
<td>Low</td>
</tr>
<tr>
<td>Del Pizzo et al. (2005)</td>
<td>CAF vs. EMD vs. CAF</td>
<td>1.61 (1.55)</td>
<td>0.07 (0.25)</td>
<td>0.20 (0.22)</td>
<td>0.53 (0.26)</td>
<td>24</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>Yes</td>
<td>7 drop-outs</td>
<td>Low</td>
</tr>
<tr>
<td>Castellanos et al. (2006)</td>
<td>CAF vs. EMD vs. CAF</td>
<td>0.91 (0.44)</td>
<td>0.97 (0.56)</td>
<td>0.86 (0.30)</td>
<td></td>
<td>12</td>
<td>RCT, parallel study</td>
<td>Inadequate</td>
<td>No</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Pilloni et al. (2006)</td>
<td>CAF vs. EMD vs. CAF</td>
<td>2.57 (0.94)</td>
<td>0.93 (0.24)</td>
<td>0.74 (0.25)</td>
<td>0.19 (0.07)</td>
<td>18</td>
<td>RCT, parallel study</td>
<td>Inadequate</td>
<td>Yes</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Woodyard et al. (2004)</td>
<td>CAF vs. ADM vs. CAF</td>
<td>3.09 (1.21)</td>
<td>1.23 (0.38)</td>
<td>0.48 (0.41)</td>
<td>0.48 (0.41)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>Yes</td>
<td>No drop-out</td>
<td>High</td>
</tr>
<tr>
<td>Córtes et al. (2004) and Córtes et al. (2006)</td>
<td>CAF vs. CAF</td>
<td>0.00 (1.29)</td>
<td>0.08 (0.16)</td>
<td>0.19 (0.27)</td>
<td>0.23 (0.29)</td>
<td>24</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>No</td>
<td>No drop-out</td>
<td>High</td>
</tr>
<tr>
<td>Mahajan et al. (2007)</td>
<td>CAF vs. CAF</td>
<td>1.00 (0.48)</td>
<td>0.43 (0.23)</td>
<td></td>
<td></td>
<td>6</td>
<td>RCT, parallel study</td>
<td>Inadequate</td>
<td>Unclear</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Huang et al. (2005)</td>
<td>CAF vs. PRP vs. CAF</td>
<td>0.22 (0.86)</td>
<td>–0.20 (0.35)</td>
<td>–0.50 (0.58)</td>
<td>–0.30 (0.34)</td>
<td>6</td>
<td>RCT, parallel study</td>
<td>Adequate</td>
<td>Yes</td>
<td>1 drop-out</td>
<td>Low</td>
</tr>
<tr>
<td>Jepsen et al. (1998)</td>
<td>CAF vs. BM vs. CAF</td>
<td>0.00 (0.82)</td>
<td>–0.01 (0.23)</td>
<td>–0.03 (0.27)</td>
<td>–0.93 (0.49)</td>
<td>12</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>No</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Zucchelli et al. (1998)</td>
<td>CAF vs. BM vs. CAF</td>
<td>–1.39 (0.61)</td>
<td>0.60 (0.21)</td>
<td>0.10 (0.22)</td>
<td>–2.45 (0.17)</td>
<td>12</td>
<td>RCT, parallel study</td>
<td>Inadequate</td>
<td>Yes</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Trombelli et al. (1998)</td>
<td>CAF vs. CTG</td>
<td>–2.58 (1.08)</td>
<td>0.90 (0.26)</td>
<td>0.60 (0.27)</td>
<td>1.00 (0.34)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>No</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Borghetti et al. (1999)</td>
<td>CAF vs. CTG</td>
<td>0.00 (0.73)</td>
<td>0.00 (0.37)</td>
<td>0.15 (0.44)</td>
<td>0.61 (0.29)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>No</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Tatakis &amp; Trombelli (2000)</td>
<td>CAF vs. CTG</td>
<td>–1.95 (1.51)</td>
<td>–0.40 (0.24)</td>
<td>–0.20 (0.22)</td>
<td>–0.60 (0.33)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Inadequate</td>
<td>Yes</td>
<td>1 drop-out</td>
<td>High</td>
</tr>
<tr>
<td>Wang et al. (2001)</td>
<td>CAF vs. CTG</td>
<td>0.00 (0.82)</td>
<td>0.20 (0.27)</td>
<td>0.40 (0.27)</td>
<td>0.40 (0.25)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>Yes</td>
<td>No drop-out</td>
<td>Low</td>
</tr>
<tr>
<td>Romagna-Genon (2001)</td>
<td>CAF vs. CTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>No</td>
<td>1 drop-out</td>
<td>High</td>
</tr>
<tr>
<td>McGuire &amp; Nunn (2003)</td>
<td>CAF vs. CTG</td>
<td>0.84 (0.83)</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>Yes</td>
<td>3 drop-outs</td>
<td>Low</td>
</tr>
<tr>
<td>Abolfazli et al. (2009)</td>
<td>CAF vs. CTG</td>
<td>0.41 (0.91)</td>
<td>–1.17 (0.39)</td>
<td>–2.29 (0.26)</td>
<td>–1.25 (0.33)</td>
<td>24</td>
<td>RCT, split mouth</td>
<td>Unclear</td>
<td>Yes</td>
<td>No inform.</td>
<td>High</td>
</tr>
<tr>
<td>Aichelmann-Reidy et al. (2001)</td>
<td>CAF vs. CTG</td>
<td>–0.76 (0.55)</td>
<td>–0.50 (0.29)</td>
<td>–0.10 (0.28)</td>
<td>–0.40 (0.52)</td>
<td>6</td>
<td>RCT, split mouth</td>
<td>Adequate</td>
<td>Yes</td>
<td>No drop-out</td>
<td>Low</td>
</tr>
</tbody>
</table>
then by CAF + ADM (0.46, 90% CrI: [−0.17; 1.16]). The CAF + CTG combination resulted to be clearly better than CAF + BM (−0.59, 90% CrI: [−1.25; −0.03]).

The NM model permitted comparisons, which had never been directly tested involving the most effective treatments: CAF + ADM versus CAF + CM, CAF + ADM versus CAF + EMD, and CAF + CM versus CAF + EMD. However, the estimated differences did not result to be either statistically or clinically significant.

The percentage of total variability due to the between-trials heterogeneity was equal to 36%, while the network inconsistency was lower ($I^2_w = 22\%$), with $Pr(t^2_w > \tau^2) = 0.40$.

The posterior distributions of the $w_{bk}$ terms were substantially centred around zero, indicating low inconsistency in the single pair-wise comparisons. However, a slight amount of skewness was observed when CAF was compared with CAF + EMD and in the comparison CAF + CTG versus CAF + EMD (Fig. 4).

The Ranking of treatments by effectiveness was the following:

1. CAF + CTG (posterior median rank = 1.92); 2. CAF + ADM (3.12); 3. CAF + CM (3.39); 4. CAF + EMD (3.65); 5. CAF + BM (4.96); 6. CAF (5.45); and 7. CAF + PRP (5.50) (Table 3, Fig. 3).

The surgical procedures with the highest probability (Pr) of being the Best treatments were the combined CAF + CTG treatment (Pr = 40%) and CAF + CM treatment (Pr = 25%) (Table 3).

**CALgain**

A certain discrepancy between the NM and SM results was detected: in particular, the effect measures for the comparison CAF versus CAF + BM and CAF versus CAF + EMD were of opposite sign. When compared with CAF alone, the greatest mean differences for CALgain were achieved by the combined treatment CAF + CTG (0.74, 90% CrI: [0.15; 1.25]), and then by CAF + ADM (0.46, 90% CrI: [−0.17; 1.16]). The CAF + CTG combination resulted to be clearly better than CAF + BM (−0.59, 90% CrI: [−1.25; −0.03]).

The percentage of total variability due to the between-trials heterogeneity was negligible ($I^2 = 1\%$), while a large amount of network inconsistency was found ($I^2_w = 88\%$) along with a very high value of $Pr(t^2_w > \tau^2) = 0.99$. The median of
the posterior distributions of the \( w_{jk} \) terms for the comparisons CAF versus CAF+EMD and CAF+CTG versus CAF+EMD was placed far from 0 and, to a lower degree, even the median of the posterior distributions for CAF versus CAF+BM and CAF+CTG versus CAF+BM (Fig. 4); this denotes the presence of relevant inconsistency on these specific comparisons. Due to the large inconsistency of the network, the results of both direct and indirect comparisons were affected by large variability.

The Ranking of treatments by effectiveness was the following: 1. CAF+CTG (2.27); 2. CAF+ADM (3.43); 3. CAF+BM (3.66); 4. CAF+CM (4.02); 5. CAF (4.34); 6. CAF+PRP (5.08); and 7. CAF+EMD (5.19) (Table 3, Fig. 3).

The surgical procedures with the highest probability (Pr) of being the Best treatments were CAF+CTG (Pr = 33%) and CAF+CM (Pr = 19%) (Table 3).

**KTgain**

The results from the NM model qualitatively agreed with those of SM, even if the variability was larger (Table 2). On the basis of NM model, the combined CAF+CM (1.41, 90% CrI: [0.48; 2.33]) and CAF+CTG (1.18, 90% CrI: [0.70; 1.66]) treatments showed the greatest mean differences for KTgain when compared with CAF alone, while CAF+EMD achieved 0.36 (90%CrI: [−0.17; 0.86]).

The difference in effectiveness between CAF+CM and CAF+CTG combinations was not either clinically or statistically significant (0.23, 90%CrI: [−0.57; 1.01]). A clear adjunctive benefit in combining collagen matrix to CAF rather than EMD, ADM, PRP, or BM was found. The network inconsistency was low (\( I^2 = 0.06\% \)) and its relevance was scarce with respect to the between-trials heterogeneity \( Pr(\tau^2 > \tau^2) = 0.13 \). Only the posterior distribution of the \( w_{jk} \) terms relative to the CAF versus CAF+CTG comparison was slightly skewed with respect to 0 (Fig. 4).

Fig. 2. Network of the seven treatments for recession reduction, clinical attachment gain, and keratinized tissue gain variables without indirect comparisons (a), with both direct and Indirect comparisons (b); and network of the eight treatments for CRC variable without indirect comparisons (c), with both direct and Indirect comparisons (d): Solid lines refer to direct comparisons (the width of the lines is proportional to the number of randomized controlled trials (RCTs) included for each comparison) while dotted lines refer to those comparisons that have not been tested directly in RCTs.
Table 2. Results of the Bayesian network meta-analysis and single Bayesian meta-analyses for RecRed, CALgain, KTgain and CRC.

<table>
<thead>
<tr>
<th>Tr. Comparison</th>
<th>N RecRed, CALgainKTgain</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bayesian network meta-analysis (NM)</td>
<td>Single Bayesian network meta-analysis (SM)</td>
<td>Bayesian network meta-analysis (NM)</td>
<td>Single Bayesian network meta-analysis (SM)</td>
<td>Bayesian network meta-analysis (NM)</td>
<td>Single Bayesian network meta-analysis (SM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Est.</td>
<td>90% CI</td>
<td>Est.</td>
<td>90% CI</td>
<td>Est.</td>
<td>90% CI</td>
<td>Est.</td>
<td>90% CI</td>
<td>Est.</td>
<td>90% CI</td>
</tr>
<tr>
<td>CAF vs. CAF-CTG</td>
<td>2</td>
<td>0.68</td>
<td>0.15; 1.25</td>
<td>0.49</td>
<td>0.20; 0.78</td>
<td>0.74</td>
<td>-0.00; 2.03</td>
<td>0.38</td>
<td>0.07; 0.69</td>
<td>1.18</td>
<td>0.70; 1.66</td>
</tr>
<tr>
<td>CAF vs. CAF-BM</td>
<td>2</td>
<td>0.08</td>
<td>-0.58; 0.72</td>
<td>-0.27</td>
<td>-0.55; 0.01</td>
<td>0.23</td>
<td>-1.32; 1.90</td>
<td>-0.33</td>
<td>-0.63; -0.04</td>
<td>0.05</td>
<td>-0.51; 0.62</td>
</tr>
<tr>
<td>CAF vs. CAF-EMD</td>
<td>5</td>
<td>0.32</td>
<td>-0.41; 0.92</td>
<td>0.57</td>
<td>0.24; 0.91</td>
<td>-0.41</td>
<td>-2.02; 1.11</td>
<td>0.54</td>
<td>0.26; 0.84</td>
<td>0.36</td>
<td>-0.17; 0.86</td>
</tr>
<tr>
<td>CAF vs. CAF-ADM</td>
<td>3$^*$</td>
<td>0.46</td>
<td>-0.17; 1.16</td>
<td>0.65</td>
<td>-0.17; 1.67</td>
<td>0.35</td>
<td>-1.24; 1.87</td>
<td>0.28</td>
<td>-0.09; 0.65</td>
<td>0.34</td>
<td>-0.21; 0.89</td>
</tr>
<tr>
<td>CAF vs. CAF-PRP</td>
<td>1</td>
<td>-0.17</td>
<td>-1.28; 0.96</td>
<td>-0.20</td>
<td>-1.17; 0.77</td>
<td>-0.56</td>
<td>-2.87; 1.61</td>
<td>-0.50</td>
<td>-0.75; 1.75</td>
<td>-0.29</td>
<td>-1.37; 0.79</td>
</tr>
<tr>
<td>CAF vs. CAF-HF-DDS</td>
<td></td>
<td>0.42</td>
<td>-0.64; 1.54</td>
<td>0.13</td>
<td>-2.42; 2.77</td>
<td>1.41</td>
<td>0.40; 2.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAF-CTG vs. CAF-BM</td>
<td>6</td>
<td>-0.59</td>
<td>-1.25; -0.03</td>
<td>-0.38</td>
<td>-0.62; -0.12</td>
<td>-0.50</td>
<td>-2.04; 1.16</td>
<td>-0.05</td>
<td>-0.29; 0.20</td>
<td>-1.13</td>
<td>-1.62; -0.63</td>
</tr>
<tr>
<td>CAF-CTG vs. CAF-EMD</td>
<td>1</td>
<td>-0.36</td>
<td>-1.19; 0.33</td>
<td>-0.17</td>
<td>-2.19; -0.15</td>
<td>-1.14</td>
<td>-2.78; 0.42</td>
<td>-0.29</td>
<td>-3.13; 0.04</td>
<td>-0.82</td>
<td>-1.46; 0.57</td>
</tr>
<tr>
<td>CAF-CTG vs. CAF-ADM</td>
<td>4</td>
<td>-0.22</td>
<td>-0.86; 0.42</td>
<td>-0.39</td>
<td>-1.16; 0.28</td>
<td>-0.38</td>
<td>-1.90; 1.32</td>
<td>-0.38</td>
<td>-1.00; 0.18</td>
<td>-0.85</td>
<td>-1.49; 0.29</td>
</tr>
<tr>
<td>CAF-CTG vs. CAF-PRP</td>
<td></td>
<td>-0.85</td>
<td>-2.10; 0.40</td>
<td>-1.30</td>
<td>-3.98; 1.25</td>
<td>-1.47</td>
<td>-2.65; -0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAF-CTG vs. CAF-HF-DDS</td>
<td>1</td>
<td>-0.26</td>
<td>-1.17; 0.69</td>
<td>-0.35</td>
<td>-0.61; -0.10</td>
<td>-0.58</td>
<td>-2.73; 1.70</td>
<td>-0.64</td>
<td>-1.01; -0.27</td>
<td>0.23</td>
<td>-0.57; 1.01</td>
</tr>
<tr>
<td>CAF-BM vs. CAF-EMD</td>
<td>2</td>
<td>0.24</td>
<td>0.09; 0.65</td>
<td>-0.64</td>
<td>-2.88; 1.51</td>
<td>0.31</td>
<td>-0.44; 1.04</td>
<td>0.26</td>
<td>-0.32; 0.87</td>
<td>0.28</td>
<td>-0.41; 0.96</td>
</tr>
<tr>
<td>CAF-BM vs. CAF-ADM</td>
<td>2</td>
<td>0.38</td>
<td>0.40; 1.25</td>
<td>0.12</td>
<td>-1.95; 3.22</td>
<td>0.53</td>
<td>-0.34; 1.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAF-BM vs. CAF-PRP</td>
<td>2</td>
<td>-0.25</td>
<td>-1.53; 1.07</td>
<td>-0.80</td>
<td>-3.45; 1.82</td>
<td>-0.34</td>
<td>-1.36; 0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAF-BM vs. CAF-HF-DDS</td>
<td>2</td>
<td>0.34</td>
<td>-0.75; 1.48</td>
<td>-0.08</td>
<td>-2.62; 2.57</td>
<td>1.36</td>
<td>0.40; 2.23</td>
<td>0.58</td>
<td>0.08; 4.04</td>
<td>0.32</td>
<td>-0.84; 2.30</td>
</tr>
<tr>
<td>CAF-EMD vs. CAF-ADM</td>
<td>1</td>
<td>0.14</td>
<td>-0.66; 1.12</td>
<td>0.76</td>
<td>-1.25; 2.85</td>
<td>0.43</td>
<td>-0.73; 0.54</td>
<td>0.33</td>
<td>0.05; 2.24</td>
<td>0.11</td>
<td>-0.25; 0.66</td>
</tr>
<tr>
<td>CAF-EMD vs. CAF-PRP</td>
<td>2</td>
<td>-0.49</td>
<td>-1.76; 0.89</td>
<td>-0.15</td>
<td>-3.06; 2.64</td>
<td>-0.65</td>
<td>-1.84; 0.55</td>
<td>0.45</td>
<td>0.37; 1.17</td>
<td>0.26</td>
<td>0.38; 0.84</td>
</tr>
<tr>
<td>CAF-EMD vs. CAF-HF-DDS</td>
<td>2</td>
<td>0.10</td>
<td>-1.01; 1.39</td>
<td>0.56</td>
<td>-2.04; 3.27</td>
<td>0.15</td>
<td>0.05; 2.07</td>
<td>0.35</td>
<td>0.04; 2.51</td>
<td>0.64</td>
<td>0.11; 0.93</td>
</tr>
<tr>
<td>CAF-ADM vs. CAF-HF-DDS</td>
<td>2</td>
<td>-0.63</td>
<td>-1.93; 0.65</td>
<td>-0.92</td>
<td>-3.77; 1.80</td>
<td>-0.62</td>
<td>-1.03; 0.49</td>
<td>0.35</td>
<td>0.04; 2.51</td>
<td>1.04</td>
<td>0.20; 1.97</td>
</tr>
<tr>
<td>CAF-PRP vs. CAF-HF-DDS</td>
<td>2</td>
<td>-0.04</td>
<td>-1.17; 1.10</td>
<td>-0.20</td>
<td>-2.63; 2.54</td>
<td>1.08</td>
<td>0.10; 2.04</td>
<td>1.94</td>
<td>0.03; 0.71</td>
<td>1.05</td>
<td>0.20; 1.97</td>
</tr>
</tbody>
</table>
CAF + CTG therapy (Pr = 28%) (Table 3).

Complete Root Coverage (CRC)

Twenty-five of the 29 studies included in this systematic review were available for the analysis of CRC. Romagna-Genon (2001), Lins et al. (2003), Castellanos et al. (2006) and Mahajan et al. (2007) did not provide data for this variable; in the case of Coêrtes et al. (2004, 2006), 12-month follow-up data were used for the analysis of CRC; in the case of Leknes et al. (2005) and Amarante et al. 2000, the 12-month follow-up data were considered as reported in Cairo et al. (2008).

The treatment alternatives considered for the analysis were eight: CAF, CAF + CTG, CAF + BM, CAF + EMD, CAF + ADM, CAF + PRP, CAF + HF-DDS, CAF + CM. Up to eight treatment options, a total of 28 comparisons were possible. Ten direct comparisons were based on data from RCTs (Fig. 2c):

- CAF versus CAF + CTG: 2 RCTs
- CAF versus CAF + BM: 1 RCT
- CAF versus CAF + EMD: 4 RCTs
- CAF versus CAF + ADM: 2 RCTs
- CAF versus CAF + PRP: 1 RCT
- CAF versus CAF + HF-DDS: 6 RCTs
- CAF versus CAF + CM: 2 RCTs

Eighteen comparisons had never been directly tested in RCTs (Fig. 2d):

- CAF versus CAF + HF-DDS
- CAF versus CAF + CM
- CAF + CTG versus CAF + PRP
- CAF + BM versus CAF + EMD
- CAF + CTG versus CAF + ADM
- CAF + CTG versus CAF + HF-DDS
- CAF + CTG versus CAF + CM

The treatment alternatives considered for the analysis were eight: CAF, CAF + CTG, CAF + BM, CAF + EMD, CAF + ADM, CAF + PRP, CAF + HF-DDS, CAF + CM. With eight treatment options, a total of 28 comparisons were possible. Ten direct comparisons were based on data from RCTs (Fig. 2c):

- CAF versus CAF + CTG: 2 RCTs
- CAF versus CAF + BM: 1 RCT
- CAF versus CAF + EMD: 4 RCTs
- CAF versus CAF + ADM: 2 RCTs
- CAF versus CAF + PRP: 1 RCT
- CAF versus CAF + HF-DDS: 6 RCTs
- CAF versus CAF + CM: 2 RCTs

Eighteen comparisons had never been directly tested in RCTs (Fig. 2d):

- CAF versus CAF + HF-DDS
- CAF versus CAF + CM
- CAF + CTG versus CAF + PRP
- CAF + BM versus CAF + EMD
- CAF + CTG versus CAF + ADM
- CAF + CTG versus CAF + HF-DDS
- CAF + CTG versus CAF + CM

For the comparison CAF vs. CAF+ADM the RCTs with direct comparison between the two treatment groups were only two for CALgain variable as the study by Mahajan et al. (2007) did not provide data on these outcomes. "Est." is the mean of the posterior distribution under NM model except for $ \tau^2 $ and $ s^2_w $ where "Est." is the median; "90% CrI" is the 90% Credibility Interval for the Est. A positive value of "Est." is to be interpreted as a difference in efficacy in favor of the second treatment when compared to the first, as shown in the column "Tr. Comparison" (i.e. in the CAF vs. CAF+CTG comparison in the NM model for RecRed, CAF+CTG shows a mean difference of 0.68 mm greater than CAF). "N" is the number of trials with direct evidence for the treatment comparison.

"OR" is the odds ratio. A value of the "OR" > 1 is to be interpreted as a difference in efficacy in favor of the second treatment when compared to the first, as shown in the column "Tr. Comparison" (i.e. in the CAF vs. CAF+CTG comparison in the NM model, CAF+CTG shows a mean OR of 2.37 greater than CAF). "N" is the number of included studies for each direct treatment comparison.

$ \tau^2 $ (heterogeneity) and $ s^2_w $ (inconsistency) are expressed as the median of the posterior distribution under NM model; for CRC $ \tau^2 $ and $ s^2_w $ are expressed as LogOR of the median of the posterior distribution under NM model. $ \hat{F}^2 $ is the proportion of total variability due to the between-trasb heterogeneity. $ \hat{F}_w^2 $ is the proportion of total variability due to the inconsistency.

Table 1. (continued)

<table>
<thead>
<tr>
<th>Tr. Comparison</th>
<th>N RecRed, CALgain/CTGain</th>
<th>RecRed</th>
<th>CALgain</th>
<th>KTgain</th>
<th>N CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bayesian network meta-analysis (NM)</td>
<td>Single Bayesian meta-analysis (SM)</td>
<td>Bayesian network meta-analysis (NM)</td>
<td>Single Bayesian meta-analysis (SM)</td>
<td>Bayesian network meta-analysis (NM)</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>90% CrI</td>
<td>Est.</td>
<td>90% CrI</td>
<td>Est.</td>
</tr>
<tr>
<td>CAF+PRP vs. CAF+CM</td>
<td>0.39</td>
<td>0.00</td>
<td>2.17</td>
<td>0.72</td>
<td>2.46</td>
</tr>
<tr>
<td>CAF+HF-DDS vs. CAF+CM</td>
<td>0.12</td>
<td>0.03</td>
<td>0.32</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>$ \tau^2 $</td>
<td>0.36</td>
<td>0.05</td>
<td>0.68</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>$ s^2_w $</td>
<td>0.07</td>
<td>0.01</td>
<td>0.05</td>
<td>1.07</td>
<td>0.28</td>
</tr>
<tr>
<td>$ \hat{F}^2 $</td>
<td>0.22</td>
<td>0.00</td>
<td>0.82</td>
<td>0.88</td>
<td>0.63</td>
</tr>
<tr>
<td>Pr($ \hat{F}^2 &gt; \tau^2 $)</td>
<td>0.40</td>
<td>0.00</td>
<td>0.99</td>
<td>0.09</td>
<td>0.13</td>
</tr>
</tbody>
</table>

© 2012 John Wiley & Sons A/S
• CAF + BM versus CAF + ADM
• CAF + BM versus CAF + PRP
• CAF + BM versus CAF + HF-DDS
• CAF + BM versus CAF + CM
• CAF + EMD versus CAF + ADM
• CAF + EMD versus CAF + PRP
• CAF + EMD versus CAF + HF-DDS
• CAF + EMD versus CAF + CM
• CAF + ADM versus CAF + PRP
• CAF + ADM versus CAF + HF-DDS
• CAF + ADM versus CAF + CM
• CAF + PRP versus CAF + HF-DDS
• CAF + PRP versus CAF + CM
• CAF + HF-DDS versus CAF + CM

Seven independent SM and one NM model were conducted for CRC outcome variable (Table 2). The SM summarized evidence from the direct comparisons (Fig. 2c). Meta-analyses for CAF versus CAF+BM, CAF versus CAF+PRP, and CAF+CTG versus CAF+HF-DDS were not performed as there was only one available study for each of these pair-wise comparisons. The NM model made it possible to include all the 28 potential pair-wise comparisons (Table 2): the seven direct comparisons analysed by SM models, the CAF versus CAF+BM, CAF versus CAF+PRP, and CAF+CTG versus CAF+HF-DDS comparisons which could not be analysed by SM and the 18 indirect comparisons never tested in single trials (Fig. 2d).

CRC

The NM results were consistent with those obtained under SM. On the basis of the NM model, the combined CAF+EMD therapy, as compared with CAF alone, achieved the greatest additional benefit in terms of CRC (OR = 3.91, 90% CI: [1.76; 9.48]); this was followed by CAF+HF-DDS (OR = 2.51, 90% CI: [0.06; 90.56]) and by CAF+CTG (OR = 2.37, 90% CI: [1.18; 5.09]).

Even if the CAF+EMD combination resulted more effective than the CAF+CTG (OR = 1.65, 90% CI: 0.66; 4.15), this result was not statistically significant.

The percentage of total variability due to the between-trials heterogeneity was negligible (I^2 = 2%) as well as the percentage due to the network inconsistency (I^2 = 2%). The value of Pr(\tau^2 > \tau^2) was equal to 0.54. The median of the posterior distribution of the w_hk terms for the comparison CAF+ADM versus CAF and CAF+CTG versus CAF was located slightly far from 0 (Fig. 4).

The Ranking of treatments by effectiveness for CRC was the following: 1. CAF+EMD (1.93); 2. CAF+CTG (2.96); 3. CAF+HF-DDS (3.72); 4. CAF+ADM (4.30); 5. CAF+PRP (4.94); 6. CAF (5.83); 7. CAF+BM (5.87); and 8. CAF+CM (6.45) (Table 3, Fig. 3).

The combined technique with the highest probability of being the Best treatment was CAF+EMD (Pr = 43%); this was followed by CAF+HF-DDS (Pr = 39%) (Table 3).

Discussion

Network meta-analysis represents a new method to combine evidence from direct and indirect comparisons among a set of different trials in a unique network of treatments. This statistical approach is of paramount importance in evaluating the efficacy of several treatments dealing with the same clinical condition (Glenny et al. 2005). Therefore, it has been considered a useful method for analysing the performance of different root coverage procedures.

The purpose of this study was to use a Bayesian NM model which could allow to make inference not only on direct treatment comparisons as evaluated in previous systematic reviews (Cairo et al. 2008) but also on comparisons never directly tested in RCTs. The direct comparisons were only 9 from a total of 21 potential pair-wise comparisons for RecRed, CALgain, and KTgain, and they were only 10 from a total of 28 potential pair-wise comparisons for CRC.

Network meta-analysis made it possible to investigate treatment comparisons never directly tested in RCTs: 12 indirect comparisons for RecRed, CALgain, and KTgain and 18 for CRC. This is relevant from a clinical point of view: for instance, the indirect comparison of CAF + BM versus CAF+PRP showed that the first technique was better in terms of RecRed, while CAF+ADM achieved more KTgain than CAF+PRP. Both these comparisons had never been tested in RCTs.

Another important advantage of the Bayesian NM is the possibility of obtaining a Ranking of the treatments included in the analysis and assessing the probability for each treatment to be the Best. Traditional meta-analytical methods do not permit grading treatments by effectiveness, while this is a fundamental issue in the clinical decision-making process, especially in the presence of a large number of available treatment options.

While the statistical combination of direct and indirect evidence is a powerful and welcome addition to evidence synthesis techniques, it also raises the issue of properly assessing a source of variability emerging in NM models: the inconsistency. The simple statistical tools proposed in this study have been developed with the aim to quantify the amount of inconsistency, both in the network and in the specific pair-wise comparisons. Inconsistency is strongly related to the discrepancy between the SM and the NM results for the comparisons informed by data. For RecRed, KTgain, and CRC, the present NM provided results qualitatively in agreement with those achieved under the single Bayesian meta-analyses. On the contrary, relevant discrepancies between the SM and NM results were observed for CALGain. Accordingly, network inconsistency (I^2) was not a concern for RecRed, KTgain, and CRC, while it was large for CALgain. The inspection of the posterior distributions of the w_hk terms made it possible to argue that the inconsistency affected in particular the evidence cycles involving the comparisons among the CAF, CAF+CTG, CAF+BM, and CAF+EMD treatments. Due to this large inconsistency, the results for CALgain were judged as quite unreliable. However, CALgain might be considered a variable of not primary relevance in evaluating root coverage procedures.

From this analysis, the relative performance of CAF alone, CAF+PRP, and CAF+BM was lower in terms of RecRed, KTgain, and CRC (Oates et al. 2003, Chambrone et al. 2010) than other CAF-based combined techniques. The CAF+CTG procedure resulted the best or the sec-
second to best treatment for all the considered outcomes. The probability that CAF+CTG was the Best treatment was 40% in terms of RecRed and 28% in terms of KTgain; CAF+EMD resulted better than CAF+CTG in terms of CRC, even if the uncertainty around the estimated OR was large enough. The CAF+CM procedure surprisingly appeared to be the worse treatment in terms of CRC and the most effective in terms of KTgain, with a probability to be the Best equal to 69%; however, these conflicting results might be due to a lack of information about this specific root coverage technique, which was investigated in only two recent studies (McGuire & Scheyer 2010, Cardaropoli et al. 2012) classified as at high risk of bias (Table 1).

Along with ranking and Best probabilities, the size of the relative effects should be carefully considered for clinical purposes. For instance, CAF+CTG, which ranked first for RecRed, presented a slight estimated difference of only −0.22 mm when compared with CAF+ADM; this difference might be considered of scarce clinical impact.

The sensitivity analysis (not reported) showed that changing the structure of the variance priors, the qualitative conclusions about the treatments comparison did not change. However, the variations in the width of the credibility intervals for the estimated differences and ORs were sometimes relevant.

It should be noted that the formulation of the NM model proposed here

---

**Fig. 3.** Plots of Ranking in efficacy for recession reduction, clinical attachment gain, keratinized tissue gain, and complete root coverage.

**Fig. 4.** Boxplots of the posterior distributions of the $w_{jk}$ terms for recession reduction, clinical attachment gain, keratinized tissue gain, and complete root coverage: each box ranges from the lower to the upper quartile with the internal line indicating the median; the whiskers provide two extreme values for the distribution so that data exceeding these limits can be considered outlying. Outliers are not reported.
can also be extended to multi-arm trials (Lumley 2002). However, only two arms studies met the inclusion criteria of this review. The attention was focused on CAF technique-based procedures. Multiple combinations of techniques were excluded due to difficulties in detecting the weight of a single therapy in the overall clinical outcome. Results from recent systematic reviews were in agreement with the outcomes of the present search. In particular, Chambrone et al. (2008, 2010, 2012) found significantly greater gain in keratinized tissue for enamel matrix protein compared with CAF and showed that subepithelial CTG, matrix grafts, and EMD were superior to CAF in achieving CRC, with CTG showing the best predictability. Accordingly, the results of this NM showed that CAF+CM might represent a valid approach among several root coverage procedures considering the role of KT in long-term maintenance of a healthy periodontium. Most of the RCTs included in this NM, were classified as at high risk of bias (20 of 29) (Table 1). This aspect should be taken into consideration when interpreting the present findings and could in principle affect the validity of the study conclusions. However, the low amount of inconsistency which characterized the network for RecRed, KTgain, and CRC gives support to the overall reliability of the results.

Conclusions

- CAF+CTG was the most effective procedure for root coverage in terms of RecRed and CALgain and could be considered the gold standard in the treatment of Miller Class I and II gingival recessions.
- CAF+EMD ranked first for CRC. However, the estimated OR between CAF+EMD and CAF+CTG was affected by large uncertainty.
- CAF+CM resulted to be the best treatment in terms of KTgain along with CAF+CTG, but further studies are required to properly assess the benefit of this and other treatment alternatives such as CAF+ADM.
- The high degree of inconsistency detected for the CALgain variable focused on the questionable role of this outcome in evaluating the efficacy of root coverage procedures.

Acknowledgements

Suggestions made by Dr. Anne-marie Glenny and Prof. Helen Worthington helped to develop the models presented in this article to a great extent. A special thanks to Prof. Graham Dunn for his help in this research and to Prof. Kevin Seymour for his scientific support.

References


Cardaropoli, D., Tamagnone, L., Roffredo, A. & Gaveglio, L. (2012) Treatment of gingival recession defects using coronally advanced flap with a porcine collagen matrix compared to...


Address: Jacopo Buti via R. Giuliani 144/A 50141 Florence Italy E-mail: ruijack@libero.it

Clinical Relevance

Scientific rationale for the study: The clinical rationale was to establish a ranking of root coverage procedures and to identify the best choice among several treatment options. To pursue this aim, a network meta-analysis model was developed and simple tools for evaluating network inconsistency were proposed. Principal findings: Coronally advanced flap, in combination with connective tissue grafts and in combination with enamel matrix derivative proved to be the most effective root coverage treatments. Practical implications: CAF+CTG, CAF+EMD, and CAF+CM can be considered very effective combination techniques for treating gingival recessions. Further studies are needed to properly assess the effectiveness of procedures alternative to CAF+CTG.