What is the effect of active ingredients in dentifrice on inhibiting the regrowth of overnight plaque? A systematic review

Valkenburg, C.; Slot, D.E.; van der Weijden, G.A.

DOI
10.1111/idh.12423

Publication date
2020

Document Version
Final published version

Published in
International Journal of Dental Hygiene

Citation for published version (APA):
What is the effect of active ingredients in dentifrice on inhibiting the regrowth of overnight plaque? A systematic review

Cees Valkenburg1,2 | Dagmar Else Slot2 | GA (Fridus) Van der Weijden2

1General Dentist and Clinical Epidemiologist, Hoevelaken, The Netherlands
2Department of Periodontology Academic Centre for Dentistry, Amsterdam (ACTA), University of Amsterdam and Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Correspondence
GA (Fridus) van der Weijden, Department of Periodontology, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, Gustav Mahlerlaan 3004, 1081 LA Amsterdam, The Netherlands.
Email: ga.vd.weijden@acta.nl

Funding information
Work for this paper is funded by the regular academic appointment of Van der Weijden and Slot. They and their research team at ACTA have previously received either external advisor fees, lecturer fees, or research grants from toothbrush and dentifrice manufacturers. Those manufacturers included Colgate, Dentaaid, GABA, Lactona, Oral-B, Procter & Gamble, Sara Lee, Sunstar and Unilever.

Abstract
Objectives: The aim of this systematic review was to establish the adjuvant clinical effect of brushing with a dentifrice containing purported active ingredients as compared to a regular sodium fluoride dentifrice with respect to the inhibition of overnight dental plaque regrowth from studies with human participants.

Methods: MEDLINE-PubMed, EMBASE and Cochrane CENTRAL were searched, up to June 2019. The inclusion criteria were controlled clinical trials with participants aged ≥18 years in good general health. Studies were included that evaluated the effect of toothbrushing with a dentifrice on the inhibition of overnight dental plaque regrowth when an active ingredient was added to the dentifrice as compared to a common sodium fluoride product. Data were extracted from the eligible studies, the risk of bias was assessed, and a meta-analysis was performed where feasible.

Result: Independent screening of 213 unique papers resulted in 10 eligible publications that provided 14 comparisons. Stannous fluoride and triclosan dentifrices were found as the active ingredients. The descriptive analysis indicated that all, but two comparisons demonstrated an additional effect on the active-ingredient dentifrice. The meta-analysis supported and strengthened these findings. It showed that when plaque was scored digitally, a DiffM was -3.15 (95% CI [-4.61; -1.69], P < .001, prediction interval [-5.07; -1.24]). When plaque was scored clinically, the difference of means (DiffM) was -0.33 (95% CI [-0.49; -0.16], P < .001, prediction interval [-0.87; 0.21]).

Conclusion: The results of this review demonstrate moderate-quality evidence that brushing with an active-ingredient dentifrice with stannous fluoride or triclosan does provide an added clinically relevant effect concerning plaque inhibition capabilities that surpass the effect of a regular sodium fluoride dentifrice.

Keywords
dentifrice, oral hygiene, plaque, plaque inhibition, plaque regrowth, systematic review, toothbrushing, toothpaste
1 | INTRODUCTION

Routine toothbrushing is perhaps the single most important step for an individual to take in order to reduce plaque accumulation and to reduce the consequent risk of plaque-associated diseases, such as periodontitis and caries. In addition to toothbrushing, dentifrices can help with stain removal, breath freshening and provide a feeling of cleanliness. Fluoride-containing dentifrices also play an essential role in caries prevention. A recent systematic review concluded that dentifrice does not provide an added effect for the mechanical removal of dental plaque. In terms of plaque removal, toothbrushing is at least as effective as toothbrushing with a dentifrice. In this respect, it seemed that the mechanical action provided by the toothbrush was the main factor in the plaque removal process. However, despite brushing every day, people are typically not effective brushers and live with large amounts of plaque on their teeth. It is here that chemical adjuncts to toothbrushing could be beneficial. Chemicals could prevent bacterial attachment, stop bacterial division and plaque growth, or may even remove plaque.

Although the long-term use of dentifrices with active ingredients intended for patients with gingivitis is associated with the prevention of bacterial biofilm formation, only a few of these products have been systematically evaluated in relation to gingival health. For example, the use of stannous fluoride or triclosan dentifrices resulted in greater gingivitis and plaque reduction than the use of a conventional dentifrice. However, the primary aim of these studies had a focus on plaque removal and not on preventing plaque accumulation on the dentition. An important intervention target for chemotherapeutics is to optimize plaque control by inhibiting overnight plaque regrowth.

Since the use of dentifrices is widespread and available scientific literature suggests that dentifrices reduce plaque regrowth, a further aspect of interest is whether following a brushing exercise dentifrices that contain purported active ingredients reduce overnight plaque regrowth more than regular sodium fluoride dentifrices. This overnight model was not included in a recent SR which demonstrated moderate-quality evidence in a 4-day non-brushing model with dentifrice slurry for a weak inhibitory effect on plaque regrowth in favour of the use of a dentifrice intended for daily use.

Therefore, the purpose of this paper was to systematically and critically appraise the literature concerning the adjuvant effect of a dentifrice on the inhibition of overnight plaque regrowth.

2 | MATERIALS AND METHODS

This systematic review was prepared and described in accordance with the Cochrane Handbook for Systematic Reviews of Interventions and in the guidelines Transparent Reporting of Systematic Reviews and Meta-Analyses (PRISMA statement).

2.1 | Protocol development

The protocol for this review was developed “a priori” and registered with the International Prospective Register of Systematic Reviews under the registration number CRD42019126734. All post hoc changes were appropriately noted (see Appendix S1).

2.2 | Focused PICOS question

In healthy adults (P), what is the effect of brushing with a dentifrice containing purported active ingredients to inhibit overnight plaque regrowth (I) compared to a regular sodium fluoride dentifrice (C) according the clinical indices of dental plaque (O) using an overnight plaque accumulation model (S)?

2.3 | Search strategy

A structured search strategy was designed to retrieve all relevant studies. As proposed in the Cochrane Handbook the National Library of Medicine, Washington, DC (MEDLINE-PubMed), EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL) were searched from initiation to June 2019 for appropriate papers that answered the focused question. The reference lists of the studies included were hand-searched to identify additional potentially relevant studies, and Google Scholar was used as additional source. No limitations were placed on language or date of publication in the electronic searches of the databases. For details regarding the search terms used, see Table 1.

TABLE 1 Search terms used for PubMed-MEDLINE and Cochrane CENTRAL. The search strategy was customized according to the database being searched. The following strategy was used in the search: ([<intervention: toothpaste> AND <outcome: overnight dental plaque>])

<table>
<thead>
<tr>
<th>Search term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>((&lt;intervention: toothpaste&gt;))</td>
<td>OR and OR terms</td>
</tr>
<tr>
<td>((text words) toothpaste OR dentifrice OR toothpastes OR dentifrices)</td>
<td>OR and OR terms</td>
</tr>
<tr>
<td>AND</td>
<td>OR terms</td>
</tr>
<tr>
<td>(&lt;outcome: overnight dental plaque&gt;)</td>
<td>OR terms</td>
</tr>
<tr>
<td>((text words) overnight OR dental plaque OR plaque removal OR plaque index OR dental plaque OR dental deposits))</td>
<td>OR terms</td>
</tr>
</tbody>
</table>
2.4 | Screening and selection

The titles and abstracts of the studies obtained from the searches were screened independently by two reviewers (CV and DES) to select studies that potentially met the inclusion criteria. No language restrictions were imposed. Based on the title and abstract, the full-text versions of potentially relevant papers were obtained. These papers were categorized (CV and DES) as definitely eligible, definitely not eligible or questionable. Disagreements concerning eligibility were resolved by referring to the original article. If no consensus could be reached, the decision was resolved through arbitration by a third reviewer (GAW). The papers that fulfilled all inclusion criteria were processed for data extraction.

2.5 | The inclusive and exclusive criteria

2.5.1 | Inclusion criteria

The included studies were considered to meet the following criteria:
(a) The study design was either a randomized controlled clinical trial (RCT) or a controlled clinical trial (CCT), and.
(b) the publications were available as full reports.
Population:
(c) The studies were conducted with humans, who were not institutionalized and who were 18 years of age or older,
(d) in good general health (no systematic disorders), and without orthodontic appliances and/or removable prostheses.
Intervention:
(e) The intervention was toothbrushing with an active-ingredient dentifrice.
Comparison:
(f) The control was a standard fluoride dentifrice.
Outcome:
(g) The studies evaluated regrowth of plaque.
(h) Setting: Overnight plaque accumulation model.

2.5.2 | Exclusion criteria

*Chlorhexidine was the active ingredient incorporated in a dentifrice.*20
*Additionally, rinsing with an antiseptic as part of the intervention or control regimen.

For details, see Appendix S2.

2.6 | Assessment of heterogeneity

The following factors were used to evaluate the heterogeneity of the outcomes of the different studies: study design, participant characteristics, study group details, side effects and industry funding.

2.7 | Assessment of methodological quality and risk of bias

All included studies were independently scored for their methodological quality by two reviewers (CV and DES) using the checklist presented in Appendix S3. Disagreement was resolved by consensus, and if disagreement persisted, the decision was resolved through arbitration by a third reviewer (GAW). The assessed items as detailed in Appendix S3 were used to classify a study as having an estimated low, moderate or high risk of bias.21

2.8 | Data extraction

The characteristics of the population, intervention, comparison and outcomes were extracted independently from all studies by two reviewers (CV and DES) using a specially designed data extraction form. Means and standard deviations (SDs) were extracted. Some studies provided standard errors (SEs) of the means. Where possible, the authors calculated standard deviation based on the sample size (SE = SD/√N) and transformed logarithmic value back to the raw scale.22 For those papers that provided insufficient data to be included in the analysis, the first or corresponding author was contacted to request additional data. If no response was received within a reasonable amount of time, the study was not included in the meta-analysis.

2.9 | Data analysis

As a summary, a descriptive data presentation was used for all studies. For studies that had multiple treatment arms and for which data from the control group were compared with more than one other group, the number of participants (n) in the control group was divided by the number of comparisons. The data are presented, and the modifications of the original indices23-25 are provided. The difference of means (DiffM) between brushing with and without an active-ingredient dentifrice was calculated using a “random-effects” model with an “inverse variance” method as proposed by DerSimonian and Laird.26 For MA with more than two comparisons, 95% predictive intervals were calculated to quantify potential treatment effects in a future clinical setting.27

Heterogeneity was tested using the chi-square test and the I² statistic with 95% confidence intervals around I².17,28 If possible, tests for small-study effects were conducted in order to detect potential publication bias. Therefore, regression tests and their modifications as proposed by Egger et al29 and Sterne et al30 were used as well as non-parametric tests and their modifications, as proposed by Begg et al31 and the trim-and-fill method and their modifications, as proposed by Duval and Tweedie32 and Peters et al33. The adjusted treatment effect estimate as based on the Copa selection model was calculated.34,35 Trial sequential analysis (TSA) was carried out which allow for an estimate of the risk of a type I
error. The required information size (RIS) and the trial sequential monitoring boundaries (TSMB) for benefit or futility were calculated. The RIS was calculated based on a type I error risk of \( \alpha = 5\% \) and a type II error risk of \( \beta = 0.20 \), with a statistical test power of 80%. RIS was accounted for heterogeneity and multiple comparisons. The Lan-DeMets version of the O’Brien-Fleming function was used for calculating the TSMBs. TSA software version 0.9.5.10 Beta (Copenhagen Trial Unit, Copenhagen, Denmark) was used.

Inflation bias or “p-hacking” was tested with a P-curve analysis. Post hoc sensitivity analysis was conducted to evaluate the influence of a single study on the overall effect estimate by stepwise omitting, one by one, each of the studies included in the meta-analysis and re-evaluating the summary effect estimates. Post hoc analysis was conducted on study design. Computations for the MA were performed using R (https://www.r-project.org) with the packages meta and metafor.

In order to judge the clinical relevance of study results, “distribution-based” methods were used. The clinical relevance was scored as not clinically relevant, potentially clinically relevant or clinically relevant based on the relationship among the mean difference of the variable, minimal important differences (MIDs) and effect sizes (ES). The MID was determined by multiplying the effect size of the difference obtained between groups considered as important (0.2 or 0.5 ES according to Cohen) by the pooled baseline standard deviation between the two groups.

2.10 | Grading the “body of evidence”

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system was used to rank the evidence. Two reviewers (DES and GAW) rated the quality of the evidence and the strength and direction of the recommendations according to the...
following aspects: risk of bias, consistency of results, directness of evidence, precision and publication bias, and magnitude of the effect.

3 | RESULTS

3.1 | Search and selection results

The search of the MEDLINE-PubMed and Cochrane CENTRAL databases resulted in 270 unique papers (for details, see Appendix S1). Screening of the titles and abstracts resulted in 18 papers, for which full reports were obtained. After detailed reading of the full reports, 11 studies were excluded at this stage. Manual searching of the reference lists of the selected papers provided one additional relevant paper. Additional searching in Google Scholar revealed two additional suitable papers. In total, 10 eligible publications describing 14 comparisons were included in this systematic review. For details, see Figure 1.

3.2 | Study characteristics and heterogeneity

The included studies exhibited considerable heterogeneity. Information regarding the study characteristics is presented in detail in Appendix S2A. The demographic characteristics are summarized in Appendix S2B.

With the exception of two studies, all reported to be randomized. Only two studies had a parallel design. The others all used a crossover design. All the studies included evaluated overnight plaque scores in the morning before brushing. In one study, the period of no oral hygiene extended to 12 hours and, in two other studies, to 24 hours. A fourth study used a power brush. The studies were carried out in several countries around the world, which included India, China, the UK, and the United States.

A total of 456 participants provided data for this review. Participants ranged from experienced employee dentifrice panel to non-dental populations.

In the studies, there were several dropouts. Two participants did not complete the study protocol and 15 participants were ineligible due to migration, unforeseen health events and other unspecified reasons. One study did not describe dropouts. The time for brushing varied in the studies. One study instructed participants to brush for one minute twice per day and two studies instructed participants to brush for two minutes twice per day. The studies did not allow any additional oral hygiene products but in four studies, floss users could continue to floss their posterior teeth only. In three studies, participants were explicitly instructed to brush only their lingual surfaces in the evening prior to the next appointment and in one study, the participants were requested to swish the intra-oral dentifrice slurry over the facial surfaces for 30 seconds. Compliance in using the dentifrice was not monitored in any of the studies. One study provided, during the initial study visit, a full dental prophylaxis to remove all supragingival plaque and calculus.

The dentifrices used in the studies varied between the studies with regard to the following: percentage fluoride compound, percentage stannous fluoride and the difference in accurately describing the ingredients. One study used a stannous fluoride prototype dentifrice. All the other dentifrice products were marketed at the time of the individual studies. Information on dentine abrasivity (RDA) was lacking in all studies.

In the included papers, two different indices for plaque scoring were used. Seven studies used the digital plaque image analysis (DPIA) index, which evaluates the percentage of visible tooth area that is covered with plaque. Three studies used the Quigley-Hein Index modified by Turesky et al., which scores plaque on a 0 to 5 point scale. The Turesky et al. index was scored full mouth and the DPIA index scored the facial aspect of the 12 front teeth.

3.3 | Industry funding

All studies mentioned the utilization of commercially available oral hygiene products (dentifrice, toothbrush). Seven studies mentioned funding by the Procter and Gamble Company, and one study mentioned funding by the Colgate-Palmolive Company. Two studies mentioned no funding, but (some) authors were employees of the Procter and Gamble Company.

3.4 | Side effects

Three included papers did mention the recording of side effects but did not observe any adverse events or side effects.

3.5 | Methodological quality and assessment of bias

To estimate the potential risk of bias, the methodological qualities of the included studies were used, as assessed in the checklist presented in Appendix S3. The procedures for allocation concealment were not described in any of the selected studies. Two studies provided explicit information on sample size calculation and power. Blinding to the product was described in all the selected studies with the exception of one study. Blinding of the examiner to the product however was unclear. Based on a summary of the proposed criteria, the estimated potential risk of bias was low for six studies, moderate for two studies, and high for two studies.

3.6 | Study outcomes results

Appendix S4 presents the results of the data extraction. For plaque scores, the Turesky et al. modification of the Quigley and Hein...
<table>
<thead>
<tr>
<th>Measurement moment</th>
<th>Included studies</th>
<th>Effect size (random)</th>
<th>Prediction interval$^b$</th>
<th>Heterogeneity</th>
<th>For details, see</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$I^2$   95% CI</td>
<td>$P$-value$^b$</td>
</tr>
<tr>
<td><strong>Baseline (Turesky)</strong></td>
<td></td>
<td></td>
<td></td>
<td>95% PI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DiffM</td>
<td>95% CI</td>
<td>$P$-value</td>
<td>95% PI</td>
</tr>
<tr>
<td>Baseline (Turesky)</td>
<td>Six comparisons</td>
<td>ALL</td>
<td>0.07</td>
<td>(-0.01; 0.15)</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td>He et al$^{62, 2x}$</td>
<td>Tcs</td>
<td>0.09</td>
<td>(-0.02; 0.21)</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>Prasad et al$^{15}$ 2x</td>
<td>Sn</td>
<td>0.05</td>
<td>(-0.07; 0.17)</td>
<td>0.422</td>
</tr>
<tr>
<td>END (Turesky)</td>
<td>Six comparisons</td>
<td>ALL</td>
<td>-0.33</td>
<td>(-0.49; -0.16)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>He et al$^{62, 2x}$</td>
<td>Tcs</td>
<td>-0.43</td>
<td>(-0.72; -0.13)</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Prasad et al$^{15}$ 2x</td>
<td>Sn</td>
<td>-0.23</td>
<td>(-0.42; -0.04)</td>
<td>.016</td>
</tr>
<tr>
<td>END (DPIA)</td>
<td>Seven comparisons</td>
<td>ALL</td>
<td>-3.15</td>
<td>(-4.61; -1.69)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: ALL, Triclosan and stannous dentifrice combined; CI, Confidence interval; DiffM, difference of means; $I^2$, Heterogeneity test, $I^2$ statistic; PI, Prediction interval; Sn, Stannous dentifrice; Tcs, Triclosan dentifrice.

$^a$ ≥ 3 comparisons.

$^b$ The number of decimals to which the annotations have been rounded off is 2.
Plaque Index was assessed clinically, and outcomes according to the DPIA index\textsuperscript{25} were scored digitally.

### 3.7 Descriptive analysis

Appendix S5 provides a descriptive summary of the significant difference between an active-ingredient dentifrice compared to a regular sodium fluoride dentifrice on overnight dental plaque regrowth, as reported by the original authors. Twelve comparisons out of 14 demonstrated a significant difference between interventions in favour of the use of dentifrice with purported active ingredients, while one comparison with a stannous fluoride–containing dentifrice\textsuperscript{59} and one comparison with a triclosan dentifrice\textsuperscript{64} demonstrated no significant difference in comparison with a regular dentifrice.

### 3.8 Meta-analysis

All studies except one\textsuperscript{25} provided information on sample size, outcomes and standard errors/deviations. No additional data were obtained after contacting the authors. A random meta-analysis could be performed, but the studies were separately analysed based on the index used. A subgroup analysis was performed by dentifrice ingredient. Table 2 presents the outcomes.

The analysis of the available data from the modification of the Quigley and Hein (Q&H) Plaque Index by Turesky et al\textsuperscript{24} included three studies, which resulted in six comparisons. End scores did provide a significant difference of means in favour of the use of dentifrice with stannous fluoride or triclosan (DiffM $-0.33$; $P < .001$; 95% CI: $[-0.49; -0.16]$). Appendix S6A,B present the forest plot of baseline and end scores using the Q&H index by Turesky et al\textsuperscript{24}

The overnight plaque indices were scored digitally using the DPIA index.\textsuperscript{25} Also, a significant difference (DiffM $3.15$; 95% CI $[-4.61; -1.69]$, $P < .001$) was observed in favour of the active ingredients triclosan or stannous fluoride. See Appendix S7 for the funnel plot.

### 3.9 Statistical heterogeneity

The percentage of variance in the meta-analysis attributable to study heterogeneity was high for the studies assessing the Q&H index by Turesky et al\textsuperscript{24} index ($I^2 = 78\%$; 95% CI: 52%–90%); $P < .01$) and low for studies that assessed plaque digitally using the DPIA index\textsuperscript{25} ($I^2 = 0\%$; 95% CI: 0%–45%); $P = .78$).

### 3.10 Publication bias detection

The test for funnel plot asymmetry, based on rank correlation\textsuperscript{31} or linear regression method,\textsuperscript{29} was not significant ($P = .46$ and $P = .55$). Contour-enhanced funnel plots and plots with trimfill\textsuperscript{33,68} are presented in Appendix S8A,B. Since most of the missing studies are located in regions of high significance, publication bias is unlikely to be the underlying cause of asymmetry.\textsuperscript{68}

A Copas selection model analysis was conducted to investigate, and attempt to correct for, selection/publication bias in the meta-analysis.\textsuperscript{34,35} Adjusting for selection bias, the Copas selection model estimated that the pooled adjusted standardized mean difference was $-0.72(P < .0001; 95\% \text{ CI: } [-1.00; -0.44])$ and equalled the random-effects model estimate of $-0.72 (P < .0001; 95\% \text{ CI: } [-1.01; -0.42])$. Although tests and funnel plots suggest that publication bias is not likely, it could not be ruled out. See Appendix S8C for the results of the Copas selection model analysis.

### 3.11 Trial sequential analysis

Appendix S10 presents the results of the trial sequential analyses (TSA) per index used. TSA of this MA showed that the effect was conclusive and reliable, and that additional data are unlikely to affect the summary effect.\textsuperscript{51}

### 3.12 Post hoc sensitivity analysis study design

Three post hoc sensitivity analyses of the crossover trials using stannous fluoride were performed in order to confirm the robustness of the results of the MA.\textsuperscript{69} A within-patient correlation of 0.5 was assumed because information of the required matched outcome data was not available.\textsuperscript{17,70} The sensitivity analysis of the crossover trials with correlation coefficients of 0, 0.25 and 0.5 is in agreement with the results of the MA. See Appendix S11 for the results of the post hoc sensitivity analysis.

### 3.13 Additional analysis

The results of the influence or sensitivity analysis by calculating pooled estimates omitting one study at a time showed that no single study significantly influenced the pooled DiffMs. See Appendix S9 for supporting information. Concerning the inflation bias indicated the P-curve analysis evidential value and no indication for p-hacking, data-mining or “selective reporting”.\textsuperscript{42,43,71,72} See Appendix S8D for the P-curve plot.

### 3.14 Clinical significance assessment

Because of the availability of sufficient data, calculation of the clinical significance or relevance was possible for three studies\textsuperscript{58,59,62} with six comparisons. The final clinical relevance judgement was estimated to be clinically relevant for all but one.\textsuperscript{59} When the studies with stannous fluoride or triclosan were pooled, the judgement was clinically relevant for the pooled triclosan experiments and...
TABLE 3  Estimated evidence profile regarding the effect on the inhibition of plaque regrowth of the adjunctive use of an active dentifrice with brushing

<table>
<thead>
<tr>
<th>Determinants of quality</th>
<th>Plaque scores</th>
<th>Subanalysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Triclosan dentifrice</td>
</tr>
<tr>
<td>Study design (Appendix S2A)</td>
<td>8 RCTs and 2 CCTs</td>
<td>4 RCTs</td>
</tr>
<tr>
<td>Risk of bias (Appendix S3)</td>
<td>Low to high</td>
<td>Low to moderate</td>
</tr>
<tr>
<td>Consistency</td>
<td>Rather consistent</td>
<td>Rather consistent</td>
</tr>
<tr>
<td>Directness</td>
<td>Generalizable</td>
<td>Generalizable</td>
</tr>
<tr>
<td>Precision (Table 1)</td>
<td>Precise</td>
<td>Precise</td>
</tr>
<tr>
<td>Reporting bias</td>
<td>Possible</td>
<td>Possible</td>
</tr>
<tr>
<td>Magnitude of the effect (Table 1)</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Strength and direction of the recommendation</td>
<td>Moderate-quality evidence in favour of</td>
<td>Moderate-quality evidence in favour of</td>
</tr>
<tr>
<td>Overall recommendation</td>
<td>With the interest to inhibit overnight regrowth of plaque, consider a dentifrice product that contains either triclosan or stannous fluoride</td>
<td></td>
</tr>
</tbody>
</table>

potentially clinically relevant in the case of the stannous fluoride experiments. See Appendix S12 for the results of the clinical relevance assessment.

3.15  | Evidence profile

The data gathered are indirect as the model of interest is a research model for a proof of principle. However, the data are rather consistent and precise. Table 3 summarizes the various aspects that were used to rate the quality of the evidence as proposed by the GRADE working group.

Eight out of 10 studies included in this review were RCTs, which are widely considered the gold standard of study design when assessing effectiveness, assuming that they are methodologically sound. The risk of bias, caused by methodological limitations, varied among the studies from low to high. Restricting in sensitivity analyses the meta-analyses to low or moderate risk of bias studies, the results were found to be robust. The unexplained heterogeneity in the meta-analyses was low for the studies using the digital DPIA index and moderate to considerable for the studies using the Turesky et al index, so it is reasonable to be moderate confident in the results presented. The results are consistent in different human populations in different geographic areas and are therefore considered generalizable. The intervention effects were consistent across studies, and no significant differences were found except for two comparisons. Tests and funnel plots for publication bias suggest that publication bias is not likely. The strength of a recommendation based on the quality of the evidence emerging from this review is estimated to be moderate concerning the usage of triclosan and weak concerning the usage of stannous fluoride in comparison with a regular fluoride dentifrice on the inhibition of overnight plaque regrowth.

4  | DISCUSSION

The prevention of dental caries and periodontal diseases centres on dental plaque control. In this context, chemical agents could represent a valuable complement to mechanical plaque control. Over recent decades, studies of various agents have provided information on their efficiency in controlling or inhibiting plaque growth. However, differentiating between dentifrices in terms of their antiplaque properties also requires assessment of their ability to inhibit plaque regrowth, which is commonly measured as overnight plaque accumulation. The purpose of this systematic review (SR) was to establish to what extent a dentifrice inhibits overnight plaque regrowth.

In this SR, dentifrices containing the active ingredients stannous fluoride or triclosan were significantly more effective at inhibiting overnight plaque regrowth than regular dentifrices containing sodium fluoride. The effect of stannous fluoride was found to extend over a 24-hour period.

In the studies with the DPIA index, no baseline scores were available, but all the participants received both interventions. Therefore, the differences as revealed in the meta-analysis concerning the end scores demonstrated true differences in outcomes of the investigated products.

Colgate® Total® and Crest® Pro-Health® are currently the only two dentifrices with purported antiplaque properties accepted by the ADA. Claims that chemotherapeutic products...
control or modify plaque may be made only if it can also be demonstrated that there is a significant effect against gingivitis. The criteria set are that studies should show a statistically significant proportional reductions of 20% or more in indices, referring to a comparison between the active therapy and the control at the end of the study. In the present SR, the criteria of sufficient proportional reduction were only met by five out of 14 experiments. This contrasts with the studies using the digital Q&H index, where one out of five studies did not meet this criterion. The weighted mean proportional reduction for the Q&H index by Turesky et al was 12.9% and for the digital DPIA index 25.0%. Several studies indicate that this difference may find its origin in the fact that computer-based plaque analyses are more precise, more objective and more sensitive than classic plaque indices. The Q&H index by Turesky et al is a 0-5 integer assessment of the plaque on labial, buccal and lingual surfaces of each individual tooth. For example, if a particular tooth area is assessed as a score of one and a toothbrush removes 50% of the plaque at this site, the resultant is still one. In order for the index to be zero, the plaque must be completely removed. It is envisioned that DPIA will overcome this and other problems. Another noteworthy difference is that DPIA performs a partial plaque measurement (the facial surfaces of 12 anterior teeth). Data from a large cross-sectional study demonstrate that an efficient, partial mouth plaque measurement at visible sites (19% of total) was comparable to whole mouth plaque scores. This is in concordance with earlier findings from Bentley & Disney.

All the studies in this SR are in some way related to the industry. Correlations between funding by industry and study outcomes are frequently observed in the literature. Studies that report positive or significant results are more likely to be published and statistically significant outcomes have higher odds of being fully reported. On the other hand, especially from renowned manufactures, the quality of the research is high because the procedures are ensured according to the criteria of good clinical practice. Moreover, several studies concluded that positive conclusions in dentifrice trials are not associated with conflict of interest or funding.

Stannous fluoride and triclosan are agents known to have antimicrobial properties. The exact ingredients of the dentifrices of the included studies in this SR were not always clear. There were also differences in ppm fluoride levels in the comparisons. This may be a concern in the comparisons of dentifrices because fluoride and sodium lauryl sulphate (SLS) also have antibacterial potential. Different formulations of the same active agents may have different effect sizes. Moreover, the compositions of a dentifrice product changes in time. The current formulation of Crest® Pro-Health® has since 2005 incorporated stabilized stannous fluoride and an ingredient for whitening benefits, sodium hexametaphosphate. In combination with zinc citrate, triclosan does not seem to be as effective as when it is formulated with Gantrex. Which effect versus a control was demonstrated to be non-significant. The complex compositions of dentifrices should be considered when evaluating individual ingredients.

Recently, the FDA has banned triclosan and certain other antiseptic chemicals. Products containing triclosan should now be subject to a premarket review. The US FDA, the European Commission and several national health authorities have reviewed Colgate Total with triclosan on several occasions and have approved Colgate Total as a safe and effective medicinal dentifrice up to the approved level of 0.3%. However, its effectiveness as an antimicrobial agent, the risk of antimicrobial resistance and its possible role in hormonal developmental disruption remain controversial. Beginning of 2019, Colgate has changed its formulation and has removed triclosan and has now a completely new formulation with L-arginine and zinc.

A chlorhexidine dentifrice can also be effective for plaque control. However, the inclusion of chlorhexidine in a dentifrice formulation can pose problems because chlorhexidine can be inactivated by ingredients such as flavours and anionic detergents. The side effects and tooth discoloration are an obstacle to the generalized use of chlorhexidine products and may have a negative impact on patient compliance, which limits its usefulness in daily practice.

4.1 Prediction intervals

Besides the difference in means (DiffM) and 95% confidence intervals, we also calculated 95% prediction intervals. The prediction interval can help understand the uncertainty about whether or not an intervention works. A prediction interval quantifies the dispersion of effect estimates of the interventions. In the experiments using the DPIA index, the effect of a new study will be within an interval of −5.07 and −1.24 with 95% confidence. For the studies using the Q&H index the effect of a new study will be within an interval of −0.87 and 0.21 with 95% confidence. In the latter case, the estimated probability that the true effect of the use of a dentifrice with triclosan or stannous fluoride in comparison with a regular dentifrice will be null or higher in a new study is 94%.

4.2 Influence analysis

The leave-one-out method can be used in a random-effects context to informally investigate the influence of specific studies by assessing whether these studies have a very high influence on the overall results of the meta-analysis effect sizes. The plot highlights influential studies, as when they are left out of the analysis, the overall estimate will be notably distorted. When a sensitivity analysis shows that the overall result is not affected to a large extent, the results of the meta-analysis give more confidence. In this review, the results of the sensitivity analyses showed that no single study significantly...
influenced the pooled estimates. See Appendix S9 for supporting information and plots.

### 4.3 | Inflation bias analysis

Inflation bias, also known as “p-hacking” or “selective reporting,” is assumed to occur when researchers try out several statistical analyses and/or data eligibility specifications and then selectively report those that produce significant results.\(^{42,43,71,72}\) The P-curve is a plot of the distribution of p-values reported in a set of scientific studies. Comparisons between ranges of p-values have been used to evaluate fields of research in terms of the extent to which studies have genuine evidential value, and the extent to which they suffer from bias in the selection of variables and analyses for publication.\(^{97}\) For details, see Appendix S8D.

### 4.4 | Trial sequential analysis

Most systematic reviews with meta-analyses are underpowered.\(^{98,99}\) TSA is a cumulative random-effects meta-analysis method that estimates a “required information size” (ie required meta-analysis sample size) using the same framework as sample size calculation for an individual RCT, but additionally accounting for heterogeneity and multiple comparisons when new RCTs are added. TSA has demonstrated the usefulness in establishing firm conclusions from a meta-analysis.\(^{60,100}\) The TSA of the Q&H index by Turesky et al\(^{24}\) index showed that the statistical evidence was conclusive.\(^{99,101}\) The same conclusion can be drawn with the TSA of the DPIA index \(^{25}\) for the studies with stannous fluoride. Therefore, TSA suggests that the statistical evidence of these meta-analyses is firm for both products. The conclusion of sufficient statistical power is supported by the P-curve in Appendix S8D-2.

### 4.5 | Post hoc sensitivity analysis study design

In a crossover trial, each participant serves as his/her own control. Between-patient variation is removed from the treatment comparison resulting in a smaller number of patients to achieve the same statistical power. Using a crossover design results in a gain in precision in all trials.\(^{44}\) A correlation coefficient describes how similar different assessments of the interventions are within a participant.

Because the results of crossover trials are generally similar to those of parallel-arm trials,\(^{102}\) the results of the crossover trials included in this MA were treated as parallel-arm trials. However, treatment-period interaction and carry-over effects of crossover trials may jeopardize the validity of such simple inferences. Nevertheless, the results of the sensitivity analysis of the crossover trials with correlation coefficients of 0.5, 0.25 and 0 were in agreement with the results of the MA. For details, see Appendix S11.

### 4.6 | Clinical significance assessment

Statistical significance analysis provides only a dichotomous answer. It may or may not be statistically significant and does not offer an indication of how important the result of the study is.\(^{51,103}\) A result can be clinically relevant but might be neglected if statistical significance was not attained due to small sample sizes and high intersubject variability. Clinical relevance or clinical significance assessment indicates whether the results are meaningful or not. In the absence of normative values for the outcomes of interest, other methods must be sought for assessing clinical relevance, such as the effect size (ES),\(^{49}\) minimal important difference (MID)\(^{50,51}\) and clinically relevant judgement.\(^{51,52}\) Assessing and reporting the clinical relevance of the outcome in addition to the analysis of statistical significance can help to simplify the transfer of knowledge from research into practice.\(^{53}\)

The clinical significance assessment as proposed\(^{51,52}\) could be performed on the studies with baseline information. These were only studies that used the Q&H index by Turesky et al\(^{24}\). A mean difference between groups higher than the MID can be considered clinically relevant.\(^{51,52}\) In the present review, five comparisons showed a clinically relevant result and one comparison showed no clinically relevant result. For details, see Appendix S12.

### 4.7 | Limitations related to the evidence that emerges from this review

Several limitations were identified for this review.

- While there is an emerging evidence base in public health, the data in support can often be difficult to find. Indexing of journals in MEDLINE has assisted those conducting systematic reviews to more easily identify published studies. However, information technology and the processes associated with indexing are not infallible. Studies may not be correctly marked by study design which may mean they are missed in the electronic searching process.\(^{104}\)
- The more resources searched, the higher the yield, and thus time and consequently the costs required to conduct a systematic review. While there is an abundance of evidence to suggest how extensive a search for randomized controlled trials (RCTs) should be, it is neither conclusive nor consistent.\(^{105}\)
- Another limitation may be the use of published research papers only. The authors of this review did not have the resources to obtain data that are kept "on file" by the various dentifrice manufacturers. This is known as the "file drawer problem," as a form of publication bias.\(^{106,107}\)
- Due to the focused question of this SR, no long-term studies were involved. As a representative of home-use, longer-duration studies of antimicrobial properties of dentifrice are required.\(^{108}\)
- The compliance of the participants to the given protocols may be considered as an important factor in the study outcomes. None of the studies mentioned that compliance was evaluated.
• Various toothbrush types were used in the studies included, and therefore, evaluation of the added benefit of the dentifrice between studies might be influenced by this diversity.

• The populations selected for studies of dental plaque assessment, in most cases, would be individuals with mild to moderate gingivitis. The question is whether it corresponds to the average person in the population. It is quite conceivable that some people with significant plaque formation benefit substantially more from a dentifrice with active ingredients than individuals do with little plaque formation.

• The clinically subjective indices are limited because inconsistent application of the index, especially in long-term clinical trials, often leads to greater variation in the data. Also lack of sensitivity of the scale may require larger study populations to define averages.

• All the included studies became available during the last two decades. However, in the majority of cases, the manner of reporting did not follow current standards, such as CONSORT 2010 and TIDieR 2014. This limitation is also reflected in the results of the risk of bias assessment. This systematic review reinforces the importance of correct and complete reporting and adherence to standards, particularly the new TIDieR checklist regarding the description and replication of interventions.

5 | CONCLUSION

This systematic review demonstrates, based on existing data, that brushing with a dentifrice with a purported active ingredient to inhibit plaque regrowth, such as stannous fluoride or triclosan, provides a significant and clinically relevant effect that surpasses the effect of a regular sodium fluoride dentifrice.

6 | CLINICAL RELEVANCE

6.1 | Scientific rationale for the study

Dentifrice does not provide an added effect for the mechanical removal of dental plaque. The question is whether purported active ingredients in dentifrices may inhibit dental plaque regrowth more effectively than a regular sodium fluoride dentifrice.

6.2 | Principal findings

Active ingredients in dentifrice such as stannous fluoride or triclosan do provide an inhibiting effect on overnight plaque scores that surpass the effect of a regular sodium fluoride dentifrice.

6.3 | Practical implications

Dentifrice does not significantly contribute to the mechanical removal of plaque but may serve as a carrier for active ingredients. The use of a dentifrice with the specific ingredient's stannous fluoride or triclosan inhibits overnight plaque regrowth more than a regular fluoride dentifrice.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Joost Bouwman, the head librarian of the ACTA, Franc van der Vlugt, the head librarian of the KNMT mediacentrum and Stephen M. Siegel, who helped retrieve full-text papers. We thank the authors Dr J. C. Jakobsen and Dr Lifeng Lin for additional information and the anonymous reviewers for their time and valuable comments. The authors are also grateful to the following authors for their responses, time and effort in searching for additional data: Dr T. He and Dr FS Panagakos.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ETHICAL APPROVAL

Ethical approval was not required.

ORCID

Cees Valkenburg https://orcid.org/0000-0002-6656-6934
Dagmar Else Slot https://orcid.org/0000-0001-7234-0037
GA (Fridus) Van der Weijden https://orcid.org/0000-0002-5075-8384

REFERENCES


33. Lemieux J, Beaton DE, Hogg-Johnson S, Bordeleau LJ, Goodwin PJ. Three methods for minimally important difference: no relationship was found with the net proportion of patients improving. J Clin Epidemiol. 2007;60:448-455.


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.