94TH ANNUAL adha CONFERENCE
JUNE 14-19, 2017
JACKSONVILLE, FL

LEARN ENGAGE ADVANCE

CE Course Handout

Tricks and Tips for Critically Appraising Science: Distinguishing the Good from the Bad and the Ugly

June 15th, 2017

adha American Dental Hygienists' Association

CRITICAL REVIEWS IN ORAL BIOLOGY & MEDICINE

Dental Flossing and Interproximal Caries: a Systematic Review

P.P. Hujoel[1*,2], J. Cunha-Cruz[3], D.W. Banting[4], and W.J. Loesche[5]

[1]Department of Dental Public Health Sciences and [2]Department of Epidemiology, School of Dentistry, Box 357475. University of Washington, Seattle, WA 98195, USA; [3]Institute of Social Medicine, University of the State of Rio de Janeiro, Brazil; [4]Division of Practice Administration, Schulich School of Medicine & Dentistry, University of Western Ontario, London, ON, Canada; and [5]School of Dentistry, University of Michigan, Ann Arbor; *corresponding author, hujoel@u.washington.edu

J Dent Res 85(4):298-305, 2006

ABSTRACT

Our aim was to assess, systematically, the effect of flossing on interproximal caries risk. Six trials involving 808 subjects, ages 4 to 13 years, were identified. There were significant study-to-study differences and a moderate to large potential for bias. Professional flossing performed on school days for 1.7 years on predominantly primary teeth in children was associated with a 40% caries risk reduction (relative risk, 0.60; 95% confidence interval, 0.48-0.76; p-value, < 0.001). Both three-monthly professional flossing for 3 years (relative risk, 0.93; 95% confidence interval, 0.73-1.19; p-value, 0.32) and self-performed flossing in young adolescents for 2 years (relative risk, 1.01; 95% confidence interval, 0.85-1.20; p-value, 0.93) did not reduce caries risk. No flossing trials in adults or under unsupervised conditions could be identified. Professional flossing in children with low fluoride exposures is highly effective in reducing interproximal caries risk. These findings should be extrapolated to more typical floss-users with care, since self-flossing has failed to show an effect.

KEY WORDS: dental caries, interproximal dental caries, dental floss, dental devices, home care, controlled trials, review, meta-analysis.

Received April 18, 2005; Accepted November 8, 2005

A supplemental appendix to this article is published electronically only at http://www.dentalresearch.org.

INTRODUCTION

Suggestions regarding the benefits of flossing date back to the early 19th century, where the belief was expressed that irritating matter between teeth is the source of dental diseases (Parmly, 1819). Further microbiological work, both in the late 19th century and in the 20th century, implicated dental plaque as the cause of caries. Since plaque build-up at interproximal sites has been reported to be more acidogenic than in other areas of the mouth (Igarashi *et al.*, 1989), and since dental floss has the ability to disrupt and remove some interproximal plaque (Waerhaug, 1981), it appears plausible that the use of dental floss should reduce interproximal caries risk. The goal of this study was to provide a systematic review of the controlled clinical trial evidence on dental floss and interproximal dental caries.

METHODS

Inclusion/Exclusion Criteria

The hypothesis of this study was that dental flossing reduces interproximal caries incidence. The treatment comparisons of interest included flossing *vs.* no flossing, or a comparison of different frequencies of flossing use. Studies where the effect of flossing could not be separated from the effects of other treatments were excluded. The primary study outcome was a measure of caries incidence. There were no restrictions with respect to the study population. Study designs included in this synthesis were limited to controlled clinical trials.

Search Strategy

The literature searches involved MEDLINE, the Cochrane Central Register of Controlled Trials (CENTRAL), the Web of Science, and the controlled-trials database of clinical trials (http://www.controlled-trials.com). Reference lists of potentially relevant reports and review articles were also searched. We used the MEDLINE highly sensitive search strategy for identifying reports of randomized controlled trials (Higgins and Green, 2005). The following strategy was used to search MEDLINE to December, 2004: ("dental caries"[MeSH Terms] OR "tooth demineralization"[MeSH Terms] OR "DMF Index"[MeSH Terms] OR "Approximal caries"[tw] OR "proximal caries"[tw] OR "Interproximal caries"[tw]) AND ("Dental Devices, Home Care"[MeSH Terms] OR floss*[tw]). The search strategies for CENTRAL and Web of Science were equivalent to those used in the MEDLINE search. Attempts to obtain missing information and 'grey' literature were made through contact with selected investigators.

Quality Assessment

A quality assessment of the trials identified by the search strategy was performed, and the following items were assessed as either being utilized in the study or not (Verhagen *et al.*, 1998): random allocation (*adequate* if method of random sequence generation prevents selection bias; *inadequate*, if sequence generation could be related to prognosis; and *unclear* if method of randomization is not reported but the word "random" is used), treatment allocation concealment (*adequate* if patient and investigators cannot foresee assignment, *inadequate* if

patient or investigator can foresee assignment, *unclear* if the method of concealment is not stated), blinding of outcomes assessors, presentation of point estimates with a measure of variability for the primary outcome measure, 'intention to treat' analysis, report of baseline characteristics by treatment group, eligibility criteria, loss to follow-up, and missing values. Blinding of the care provider and the subject was not applicable to our research question.

Topical fluoride exposure was categorized as not recommended (--), recommended to a subgroup or the whole cohort but with no compliance measures (-), recommended and compliance assessed (+), or delivered under supervised conditions (++). Oral hygiene was similarly classified as no instructions provided (--), instructions provided but compliance not measured (-), instructions provided and compliance measured by plaque scores or gingival bleeding scores (+), or provided under supervision (++).

Statistical Analysis

For each trial, the number of surfaces at risk and the number of new interproximal caries lesions were derived or estimated based on published data. Since the timing of caries events was typically not reported, caries risks rather than caries rates are reported. Differences in follow-up times across studies typically have a negligible impact on the results (Guevara *et al.*, 2004), a fact which was confirmed to be accurate for summary estimates reported here.

Both the relative risk (RR, or a ratio of the surface caries risk in the flossing group to the surface caries risk in the control group) and the risk difference (RD, or a difference in the surface caries risks between the flossing and the control groups) and their respective standard deviations were calculated. The standard errors of the RR and RD were calculated according to standard formulae for 2x2 tables (Rothman *et al.*, 1998), multiplied by the square root of the variance inflation factor to account for the dependence of sites within patients. We estimated the variance inflation factors for each trial by dividing the appropriate standard error of a patient-based *t* test (one- or two-sample *t* tests as described in the APPENDIX) by the standard error of a surface-based *t* test. Three split-mouth trials (Granath *et al.*, 1979; Wright *et al.*, 1979, 1980) did not report sufficient information for variances to be calculated. For two split-mouth trials (Wright *et al.*, 1979, 1980), the variance was derived from the McNemar statistic and an estimate of the correlation of matched pairs within patients (Mäkinen *et al.*, 1996). For the remaining split-mouth trial (Granath *et al.*, 1979), the variance was derived from the reported caries rate (Marinho *et al.*, 2004) and an estimate of the split-mouth correlation from an external dataset (Mäkinen *et al.*, 1995).

The results of the studies included in the review were summarized by general variance-based methods, with the weight for each study being the inverse of the variance (fixed-effects model) (Petitti, 1994). The summary RR and RD are equal to the sum of the weights times the risk estimate for each study, divided by the sum of the weights (Petitti, 1994). The weighted grand mean difference was calculated, and the I^2 statistic was used to test the hypothesis of the homogeneity of the effect. An I^2 statistic larger than 50% is considered moderate to high heterogeneity (Higgins *et al.*, 2003) and an indication to use random-effects models (DerSimonian and Laird, 1986). The effect of study characteristics such as fluoride, oral hygiene, or caries risk on flossing effectiveness was estimated by meta-regression (Thompson and Higgins, 2002).

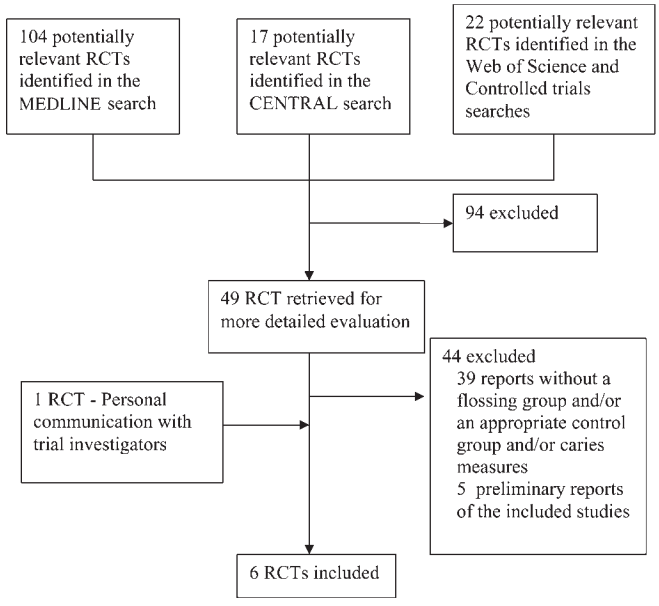


Figure 1. Flow diagram of the selection process of controlled trials on flossing and interproximal caries.

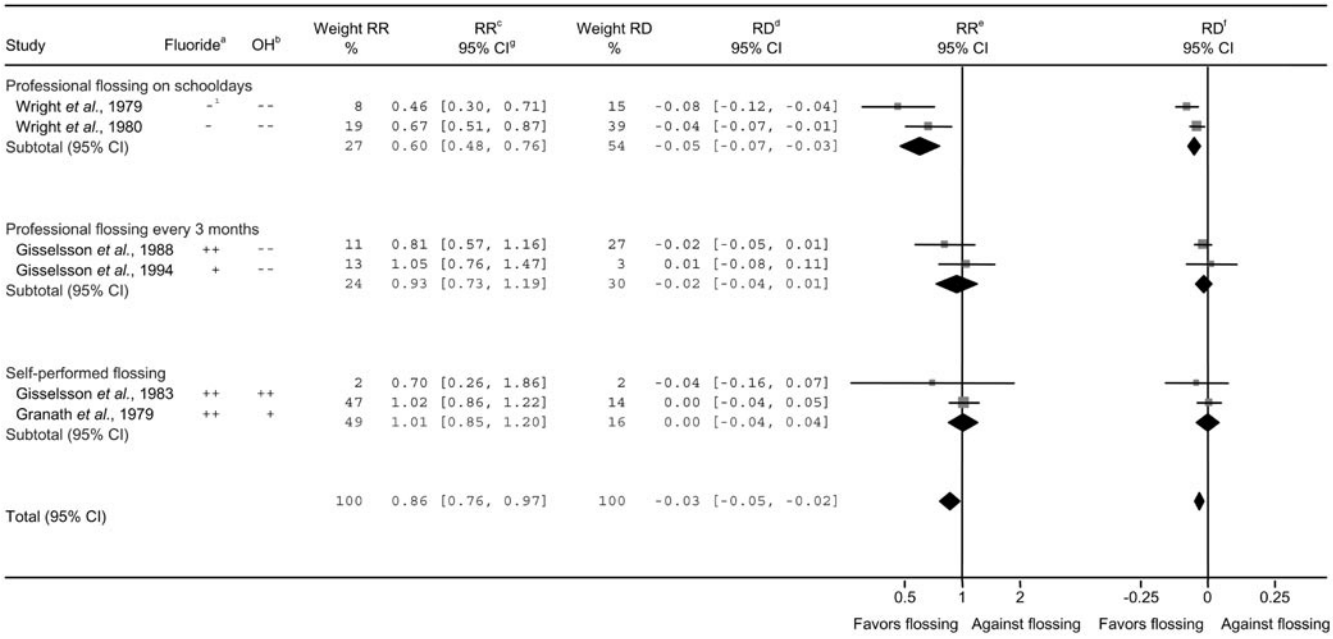
RESULTS

Study Identification and Characteristics (Fig. 1)

The MEDLINE search identified 104 controlled trials, the CENTRAL search identified 17 controlled trials, and the other sources identified 22 trials. One trial could not be located in MEDLINE or in CENTRAL, but was identified by contacting trial investigators (Wright *et al.*, 1980). Trials were excluded from this review for the following reasons: no assessment of the effect of flossing, absence of caries outcomes, no control group that would have allowed flossing effects to be estimated, or preliminary reports of the included studies (Fig. 1). Six randomized clinical trials of the effect of flossing on interproximal caries, involving 808 participants followed for 1.7 to 3 yrs, were included in this review (Table 1). The ages of the study participants ranged from 4 to 13 yrs. Flossing was professionally performed on school days in two studies (typically 5 days *per* week for a 10-month period or approximately 121-150 flossings days *per* year) (Wright *et al.*, 1979, 1980), professionally performed once every 3 months in two studies (Gisselsson *et al.*, 1988, 1994), supervised on school days in one study (Granath *et al.*, 1979), and unsupervised in one study (Gisselsson *et al.*, 1983).

Study Quality

The quality assessment of the controlled trials revealed generally poor reporting of the studies and the presence of a moderate to high risk of bias (Table 2). Two trials used the word "random" to describe the allocation method of proximal surfaces within individuals to treatments, but the allocation concealment method was not described (Wright *et al.*, 1979, 1980). One study did not use the word "random" to describe the allocation method, but did describe that children were ranked according to baseline disease severity and systematically (non-randomly) distributed to experimental and control groups (Gisselsson *et al.*, 1983). In three studies, treatments were



[a] Fluoride Topical fluoride exposure was categorized as not recommended (--), recommended to a subgroup or the whole cohort but with no compliance measures (-), recommended and compliance assessed (+), or delivered under supervised conditions (++).

[b] Oral hygiene was similarly classified as no instructions provided (--), instructions provided but compliance not measured (-), instructions provided and compliance measured by plaque scores or gingival bleeding scores (+), or provided under supervision (++).

[c] Test for overall effect: Z = 2.54 (P = 0.01) and test for heterogeneity: Chi^2 = 16.77, df = 5 (P = 0.005), I^2 = 70.2%.

[d] Test for overall effect: 3.88 (P = 0.0001) and test for heterogeneity: Chi^2 = 9.37, df = 5 (P = 0.10), I^2 = 46.6%.

[e] RR: Relative Risk.

[f] RD: Risk Difference.

[g] CI: Confidence intervals.

[h] The publication suggests a score of (--), while a personal communication suggests a score of (-).

Figure 2. Flossing and interproximal dental caries—fixed-effects meta-analysis and Forrest plot of the relative risks and risk differences.**(AQ)**

assigned not to individuals, but to groups of individuals, such as school classes (Granath *et al.*, 1979; Gisselsson *et al.*, 1988, 1994). The method of assignment of flossing to group was based on birth-year (Gisselsson *et al.*, 1994), by lot (Granath *et al.*, 1979), and by allocation (the word "random" not specified) (Gisselsson *et al.*, 1988). In all four studies reporting information on baseline caries rates, the experimental units assigned to flossing were healthier (fewer cavities) at baseline (Granath *et al.*, 1979; Gisselsson *et al.*, 1983, 1988, 1994). When summarized across the four studies, the difference in baseline caries was significant (p-value < 0.05). No statistical comparison of baseline caries rates by treatment group could be obtained in two studies (Wright *et al.*, 1979, 1980), preventing an assessment of the successfulness of the randomization.

The outcome assessors were not blinded in one study (Gisselsson *et al.*, 1983). No studies reported confidence intervals for the outcome point estimates. It was unclear if any of the reported studies provided an intent-to-treat analysis. Eligibility criteria were provided for all six studies. The three split-mouth trials did not report adjustment for within-patient correlation of split-mouths or class clustering (Granath *et al.*, 1979; Wright *et al.*, 1979, 1980). Two studies reported industry support for the conduct of the study (Gisselsson *et al.*, 1988; Wright *et al.*, 1980). For one study, this support included materials, costs, and salaries (Banting, 2005), while for the other study the support was limited to materials (Birkhed, 2005). The four remaining studies reported government or professional organization support (Granath *et al.*, 1979; Wright *et al.*, 1979; Gisselsson *et al.*, 1983, 1994).

Overall Summary (Fig. 2)

Differences in interproximal caries rates in the flossing and control groups were found to be not statistically significant in four of the six reported trials and statistically significant in two studies. For these latter studies, we imputed p-values of 0.0003 (Wright *et al.*, 1979) and 0.006 (Wright *et al.*, 1980). The p-values were imputed because the reported values in the study did not take into account the clustering of matched surface-pairs within patients. Four trials reported a decreased relative caries risk, ranging from 19% to 54%, associated with flossing (Wright *et al.*, 1979, 1980; Gisselsson *et al.*, 1983, 1988), while two trials reported a slightly increased relative caries risk associated with flossing (Granath *et al.*, 1979; Gisselsson *et al.*, 1994). Substantial quantitative heterogeneity was found on the multiplicative scale (I^2 = 70%, p-value < 0.001) and less so on the additive scale (I^2 = 47%; p-value = 0.10). Using a fixed-effects model, ignoring heterogeneity, on a relative risk basis, we found a 14% lower risk of caries on flossed surfaces over an approximately two-year period (Risk Ratio = 0.86; 95% confidence interval (CI) = 0.76, 0.97; p-value = 0.01). On an additive scale, flossing was associated with an absolute caries risk reduction of 3% (95% CI = -0.05, -0.02; p-value < 0.001). A random-effects model, which may be more appropriate considering the substantial heterogeneity, identified marginally

Table 1. Characteristics of Controlled Trials on Flossing and Interproximal Caries

Study	Wright *et al.* (1979)	Granath *et al.* (1979)	Wright *et al.* (1980)	Gisselsson *et al.* (1983)	Gisselsson *et al.* (1988)	Gisselsson *et al.* (1994)
Study Design	Split-mouth	Split-mouth	Split-mouth	Parallel	Parallel	Parallel
Caries risk in flossing group	25/374 (0.07)	313/1245 (0.25)	50/624 (0.08)	40/414 (0.10)	331/4199 (0.08)	263/927 (0.28)
Caries risk in control group	54/374 (0.14)	303/1229 (0.25)	75/624 (0.12)	40/289.5* (0.14)	373/3836 (0.10)	487/1810 (0.27)
Flossing intervention	Professional flossing on test quadrants on school days	Supervised 'simple' flossing on school days. Waxed floss was forced up and down once through each proximal contact on the experimental side	Professional flossing on test quadrants on school days	Unsupervised flossing. Instruction to use floss every evening. Every 3rd week, their performance on the method was controlled, and re-instruction was given if necessary	Professional flossing every 3rd month	Professional flossing every 3rd month
Compliance	Supervised	Supervised	Supervised	Unsupervised—Only eight children reported flossing at least every other day	Supervised	Supervised
Control groups for flossing intervention	Contralateral side in same patient received no professional flossing.	Dental assistants checked that only the experimental side was flossed.	Contralateral side in same patient received no professional flossing.	No flossing instructions	No professional flossing	No professional flossing
Outcome assessment	DFS and dfs indices of permanent and primary posterior teeth based on clinical and radiographic examinations. Only proximal lesions penetrating the dentino-enamel junction were counted	DFS index of permanent posterior teeth (excluding second molars) based on clinical and radiographic diagnostic criteria. No information on caries severity.	DFS and dfs indices of permanent and primary posterior teeth based on clinical and radiographic examinations. Only proximal lesions penetrating the dentino-enamel junction were counted.	DFS index of incisors and 1st molars based on clinical diagnostic criteria only. Koch caries criteria (Koch, 1967)	DFS of all 56 interproximal surfaces based on clinical and radiographic diagnostic criteria. Enamel caries lesions were included in the score	def of interproximal surfaces based on clinical and radiographic diagnostic criteria. Both enamel and dentinal proximal lesions were counted.
Participants enrolled	118	Not reported	188	40	163[a]	197[a]
Participants analyzed	88	140	147	35	148	174
Participants' age (yrs) at start of study	5.8	12-13	1st grade	10-11	11	4
Study duration (yrs)	1.7	2	1.7	2	3	3
Fluoride exposure	Drinking water < 0.1 ppm, no information on use of fluoride products, no oral hygiene procedures or instructions	Fluoridated toothpaste and brushing every 6th wk with 0.2% NaF solution	Drinking water < 0.1 ppm, self-reported use of fluoride and non-fluoride toothpaste, no oral hygiene procedures or instructions	Weekly 0.2% NaF solution and bi-annual topical fluoride application	Drinking fluoridated water 0.3 ppm, weekly rinsing with 0.2% NaF	Drinking fluoridated water 0.2 ppm; recommended to use 250 ppm fluoride toothpaste and fluoride tablets

* (Mean number of interproximal surfaces examined *per* child - mean number of carious interproximal surfaces at baseline) multiplied by the number of children of control group, or (24-4.7)*15.

[a] Estimated based on the assumption of similar dropout rates across treatment groups.

statistically significant association on a relative risk scale (Risk ratio = 0.79; 95% confidence interval, 0.61, 1.01; p-value = 0.06) and a statistically significant association on a risk difference scale (Risk difference, -0.03; 95% CI = -0.06,-0.01; p-value = 0.02).

Effectiveness Modifiers

Flossing's effectiveness depended significantly on whether studies reported the administration of topical fluorides and the assessment of compliance (p-value < 0.001). The four studies reporting delivery of topical fluoride or assessment of fluoride compliance had a slightly increased caries risk associated with flossing (Granath *et al.*, 1979; Gisselsson *et al.*, 1983, 1988, 1994), while the two studies reporting no assessment of topical fluoride exposure or compliance measures showed a highly significant flossing benefit (Wright *et al.*, 1979, 1980). Oral hygiene measures were not significantly related to the effectiveness of flossing (p-value = 0.31). The flossing effectiveness depended significantly on the caries risk in the population, with high caries risk translating into non-effectiveness of flossing (p < 0.002). The interproximal caries risk was approximately 25% in two studies (Granath *et al.*,

Table 2. Quality of Controlled Trials on Flossing and Interproximal Caries

Study	Wright *et al.* (1979)	Granath *et al.* (1979)	Wright *et al.* (1980)	Gisselsson *et al.* (1983)	Gisselsson *et al.* (1988)	Gisselsson *et al.* (1994)
Random allocation mechanism	Unclear	Adequate	Unclear	Inadequate	Inadequate	Inadequate
Reported allocation concealment	Unclear	Unclear	Unclear	Inadequate	Unclear	Inadequate
Details of allocation	Proximal surfaces were randomly assigned to flossed and control groups by quadrant	Each of the six classes was divided into two halves by lot; half flossed on the left side and the other half on the right side	Proximal surfaces were randomly assigned to floss and control groups by quadrant	Alternate allocation of individual subjects according to baseline DFS scores	Nine school classes were allocated to three treatments	Control group was born in same city but in the year that preceded and followed the experimental group
Blinding of outcome assessors	Yes	Yes	Yes	Not reported	Yes	No
Reported loss to follow-up	Yes	No	Yes	Yes	Yes	Yes
N (%) of drop-outs	30 (25.4)	Not reported	41 (21.8)	5 (12.5)	15 (9.2)	23 (11.4)
Intention to treat analysis	No	Unclear	No	Partly	No	No
Reported baseline characteristics	No	No	No	Yes	Yes	Yes
Reported eligibility criteria	Yes	Yes	Yes	Yes	Yes	Yes
Reported point estimates with correct confidence intervals	No	No	No	No	No	No
Reported missing values	Yes	Yes	Yes	Yes	Yes	Yes

1979; Gisselsson *et al.*, 1994), and ranged from 10% to 14% in the four remaining studies.

Subgroup Analyses

The results were stratified according to flossing frequency and method. A summary of the two studies in children using professional flossing on predominantly primary teeth performed on school days during 1.7 yrs identified a reduced caries risk on the difference scale (Risk Difference= -0.05; 95% CI = -0.07, -0.03; p-value < 0.001; I^2 = 55%) and on the ratio scale (Relative Risk = 0.60; 95% CI = 0.48, 0.76; p-value < 0.001; I^2 = 50%). For the two studies that performed professional flossing once every 3 mos for 3 yrs, no reduced caries risk was observed on the difference scale (Risk Difference = -0.02; 95% CI = -0.04, 0.01; p-value = 0.32; I2 = 0%) or on the ratio scale (Relative Risk = 0.93; 95% CI = 0.73, 1.19; p-value = 0.58; I2 = 10%). The two studies using self-performed flossing over a two-year period similarly did not identify a reduced caries risk on either the difference scale (Risk Difference = 0.00; 95% CI = -0.04, 0.04; p-value = 0.96; I2 = 0%) or the ratio scale (Relative Risk, 1.01; 95% CI = 0.85, 1.20; p-value = 0.93; I2 = 0%).

Sensitivity Analyses

The two split-mouth studies that reported a statistically significant flossing effect did not take into account the within-patient correlation of matched surface-pair observations. Sensitivity analyses were used to impute at what level of within-patient clustering of caries events the reported flossing benefits would become non-significant. If the within-patient correlation of matched surface-pairs was larger than 0.96 in the first study (Wright *et al.*, 1979) and larger than 0.39 in the second study (Wright *et al.*, 1980), the results would not be statistically significant.

DISCUSSION

A systematic review of the evidence on flossing indicates that professional flossing performed in first-grade children on school days reduced caries risk by 40%. This benefit was identified in predominantly the primary teeth in children who, it is assumed, had comparatively poor oral hygiene and minimal exposure to fluoride. When professional flossing was performed on a three-monthly basis, there was no evidence of a benefit, suggesting that infrequent flossing may be ineffective when it comes to caries control. When flossing was self-performed by young adolescents, even under supervision on school days, there was also no evidence of benefit, which may be due to the presence of fluorides, poor flossing techniques, or other reasons. No evidence on the effectiveness of floss in adults or under real-world clinical conditions could be identified. In particular, there was no evidence that flossing is effective in the presence of topical fluorides.

Of the six trials that evaluated the effect of flossing on interproximal caries risk, two trials reported that professional flossing reduced caries risk (Wright *et al.*, 1979, 1980). The strengths of these two studies include the large observed relative risk reduction of 40%, the statistical significance of the

combined results, and the possibility that flossing benefits were underestimated because there was no professional flossing performed on weekends and the extended summer break, and, possibly, because parental flossing of control teeth diluted the professional flossing effect. Weaknesses include that both studies were not truly independent, since they were conducted by the same investigators, that financial support may have biased study findings (Wright *et al.*, 1980), that minimal data were available on oral hygiene and fluoride exposure, and that a difference of 54 caries lesions in two studies combined is a small number on which to base universal flossing recommendations. If the benefits of flossing clustered within mouths more than we estimated, or if the baseline randomization was biased, the statistical significance of a flossing benefit could come into question.

Four studies failed to identify a flossing benefit. In three of the studies, apparent straightforward reasons can be identified to explain the lack of a flossing effect. A sample size of 20 flossers, only eight of whom actually reported using floss more than every other day, doomed, in all likelihood, one study's ability to find a flossing effect (Gisselsson *et al.*, 1983). Three-monthly flossing in two other studies may have been too infrequent to provide a benefit (Gisselsson *et al.*, 1988, 1994). The enigma is one split-mouth study showing a high caries rate, but no identifiable anti-caries benefit from supervised flossing (Granath *et al.*, 1979). Like the other studies, this study had weaknesses, including no information on dropouts, insufficient details on statistical analyses, and no information on group randomization. The lack of an effect is puzzling, however. Possibly, moving the floss once through the contact point, as opposed to wrapping the floss around the tooth's proximal surface and then moving it up and down to disrupt or remove the interproximal plaque, was not effective against caries. An alternative explanation is that young adolescents may have been more conscientious regarding their oral hygiene, and the increased frequency of topical fluoride exposure through toothbrushing may have eliminated any benefit of flossing.

An important limitation of the current evidence on flossing is the inability to establish whether flossing provides a benefit above and beyond brushing with a fluoridated toothpaste (Rule *et al.*, 1984; Conti *et al.*, 1988). In the two studies where toothbrushing compliance at home was not assessed, and where no toothbrushing was performed under supervised conditions at school, a flossing benefit was observed (Wright *et al.*, 1979, 1980). In the four studies where topical fluoride compliance was assessed or delivered under supervised conditions, flossing was non-effective (Granath *et al.*, 1979; Gisselsson *et al.*, 1983, 1988, 1994). This stark contrast between studies suggests that topical fluoride exposure may attenuate or eliminate the effectiveness of flossing. The effect of flossing was also limited to studies with low caries rates, raising the possibility that flossing, just like other caries-preventive agents, may be less effective in high-risk populations (Seppä *et al.*, 1991; Forgie *et al.*, 2000; Hausen *et al.*, 2000; Kallestal, 2005). Nonetheless, these are leaps of faith, since many other factors, including flossing frequency and technique, also differed between the studies.

Several observations suggest that the actual fluoride exposure in the two studies reporting a significant flossing benefit was low. The two studies were conducted on young (5-6 yrs old) children who are typically assumed to be "unable to perform satisfactory oral hygiene themselves" (Poulsen *et al.*, 1976), and for whom "no other oral hygiene procedures or instruction was provided" (Wright *et al.*, 1979, 1980). This assumption of lack-of-oral hygiene appears to be validated, since children using and not using fluoridated toothpaste had approximately the same number of caries lesions in one of the two studies (Wright *et al.*, 1980). While the lack of a fluoride effect may have been a chance finding due to low power, or due to self-selection bias, it may also possibly have been due to lack of adequate brushing. In addition, the children were living in an area where fluoride levels in the water were low. As a result, there is a possibility that flossing may be effective in a situation where oral hygiene is poor and where exposure to fluorides is minimal. This assumption of poor oral hygiene and consequent low-fluoride exposure is plausible and is supported, in part, by the reported data, but remains an assumption nonetheless, since information on the actual oral hygiene levels and toothpaste characteristics in these two studies was not reported (Wright *et al.*, 1979, 1980).

Weaknesses of this systematic review relate to the inclusion/exclusion criteria, the focus of the research question, the statistical uncertainties present in calculating summary estimates, and lack of consideration for the caries lesion severity in the analysis. A rigorous pre-analysis definition of study inclusion/exclusion criteria appeared impossible because of the authors' familiarity with the trials, and the high likelihood that apparently reasonable pre-synthesis established study inclusion criteria would have excluded most, if not all, of the available evidence. Additionally, this systematic review focused on the effectiveness of floss, not flossing and brushing. A systematic review on the latter topic would have included a study where the effect of brushing *and* flossing on interproximal caries risk was evaluated (Horowitz *et al.*, 1977), and where, in the absence of fluoride, a marginally significant effect of plaque removal on interproximal caries risk was observed. The three split-mouth studies on flossing and caries (Granath *et al.*, 1979; Wright *et al.*, 1979, 1980) failed to report information that would have allowed variances to be estimated. Unless data from these split-mouth trials can be resurrected, we are unable to determine how (in)accurate our imputed values were. Finally, for this synthesis, caries lesions limited to the outer enamel, caries lesions into the dentin, and caries lesions in primary and permanent tooth were all considered to be similar events. The validity of this assumption has been insufficiently evaluated and must be considered when the findings are interpreted.

An additional weakness of the systematic review is the lack of power. This has two consequences. First, possibly, real-world flossing has a modest impact on interproximal caries lesions, and only large studies will be able to identify these benefits reliably. The second consequence of the lack of adequately powered clinical trials on self-performed flossing is the lack of safety information. None of the six controlled trials assessed safety. It cannot be excluded that non-professional simple flossing (moving the floss through the interproximal contact point only), such as performed by the young adolescents in one study (Granath *et al.*, 1979), increases the caries risk by 22% (22% is the upper 95% confidence interval observed in the study on self-flossing). Flossing may cause harm by disrupting from 2 to 3.5 mm of the epithelial cuff around the teeth (Waerhaug, 1981), and by damaging both

tooth and periodontal structures (Ratcliff, 1966; Gow and Kelleher, 2003). Daily imperfect flossing may select for colonization by floss-resistant cariogenic strains that penetrate the tooth or the white-spot lesion, may transmit infections from one interproximal site to another, or may enhance pathogenic maturation of inaccessible or unremoved plaque (Loesche, 1993). Granath indicated that some have postulated that the viability of bacteria in a plaque follows a gradient, the outermost bacteria being most viable. Ineffective flossing stirs plaque around and might therefore be harmful if the less viable plaque is removed (Granath, 2005). The possibility for harmful effects of non-professional flossing should be assessed in future trials.

A scientific double-standard exists in the evaluation of dental drugs and devices. Some dental devices, such as toothbrushes or floss, have largely escaped the rigorous scientific evaluation that is required for drugs. While the Food and Drug Administration and the American Dental Association indicate that dental floss may reduce caries risk (Food and Drug Administration, 2005), there are not two independent, randomized controlled trials demonstrating that self-performed flossing can reduce caries risk. The Council on Scientific Affairs of the American Dental Association suggests (Acceptance Program Guidelines, 2003) that interdental cleaning devices should be evaluated "under unsupervised conditions" and "by the average patient", conditions under which the effect of floss on caries has not been evaluated. There have been no trials showing that flossing prevents caries in adults in real-world clinical situations.

The advocacy of flossing as a caries-preventive tool hinges in large part on apparent common sense. Since dental plaque is cariogenic, and since dental floss disrupts and removes some interproximal plaque (Waerhaug, 1981), flossing should reduce caries risk. Such a common-sense argument represents the lowest level of scientific evidence (Sackett *et al.*, 2000). Common sense was wrong in claiming that knee debridement relieves osteoarthritic knee pain (Moseley *et al.*, 2002), that optic nerve decompression prevents vision loss (Ischemic Optic Neuropathy Decompression Trial: twenty-four-month update, 2000), or that internal mammary artery ligation improves cardiovascular outcomes (Cobb *et al.*, 1959). Several trials have also failed to support the common-sense argument that dental plaque removal lowers caries risk (Horowitz *et al.*, 1977; McKee *et al.*, 1977; Silverstein *et al.*, 1977; Agerbaek *et al.*, 1978; Ashley and Sainsbury, 1981), which led to the hypothesis that a mutans streptococci infection cannot be controlled by mechanical means (Loesche, 1993). One should be careful to justify flossing based on common-sense arguments, especially when other caries-preventive interventions are supported by higher levels of evidence.

In summary, the controlled trial evidence on flossing and dental caries is challenging to interpret because of the inconsistent results across trials, the difficulty in extrapolating results of two trials conducted in children who differ substantially from typical floss-users, and the poor to moderate scientific quality of some of the reported studies. There have been no evaluations in real-world clinical situations, and, as a result, clinical recommendations have to be based on a level of evidence which would be questionable if flossing were a drug. The current low-level evidence is consistent with the hypothesis that regular and meticulous flossing can drastically lower interproximal caries risk in young children with poor toothbrushing habits and low fluoride exposure. Better toothbrushing and/or enhanced topical fluoride exposure may attenuate or eliminate this flossing effect. The dental professional should determine, on an individual patient basis, whether professional-quality flossing is an achievable goal, and to what extent a recommendation to floss may decrease the exposure time to caries interventions that are supported by better evidence. Factorial designs, where the effects of novel fluoride toothpastes and flossing devices are evaluated simultaneously, may provide a relatively low-cost opportunity to determine what fraction, if any, of interproximal cavities can be prevented by dental floss in a fluoridated world.

REFERENCES

Acceptance Program Guidelines (2003). Dental floss or other interdental cleaners. ADA American Dental Association. Council on Scientific Affairs 2003. http://www.ada.org/prof/resources/positions/standards/guide_floss.pdf.

Agerbaek N, Poulsen S, Melsen B, Glavind L (1978). Effect of professional toothcleansing every third week on gingivitis and dental caries in children. *Community Dent Oral Epidemiol* 6:40-41.

Ashley FP, Sainsbury RH (1981). The effect of a school-based plaque control programme on caries and gingivitis. A 3-year study in 11 to 14-year-old girls. *Br Dent J* 150:41-45.

Banting D (2005). Personal communication. London, ON, Canada.

Birkhed D (2005). Flossing studies (personal communication).

Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA (1959). An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med* 260:1115-1118.

Conti AJ, Lotzkar S, Daley R, Cancro L, Marks RG, McNeal DR (1988). A 3-year clinical trial to compare efficacy of dentifrices containing 1.14% and 0.76% sodium monofluorophosphate. *Community Dent Oral Epidemiol* 16:135-138.

DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Control Clin Trials* 7:177-188.

Food and Drug Administration (2005). 21CFR872.6390. In: Code of Federal Regulations [http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=872.6390].

Forgie AH, Paterson M, Pine CM, Pitts NB, Nugent ZJ (2000). A randomised controlled trial of the caries-preventive efficacy of a chlorhexidine-containing varnish in high-caries-risk adolescents. *Caries Res* 34:432-439.

Gisselsson H, Björn AL, Birkhed D (1983). Immediate and prolonged effect of individual preventive measures in caries and gingivitis susceptible children. *Swed Dent J* 7:13-21.

Gisselsson H, Birkhed D, Björn AL (1988). Effect of professional flossing with chlorhexidine gel on approximal caries in 12- to 15-year-old schoolchildren. *Caries Res* 22:187-192.

Gisselsson H, Birkhed D, Björn AL (1994). Effect of a 3-year professional flossing program with chlorhexidine gel on approximal caries and cost of treatment in preschool children. *Caries Res* 28:394-399.

Gow AM, Kelleher MG (2003). Tooth surface floss loss: unusual interproximal and lingual cervical lesions as a result of bizarre dental flossing. *Dent Update* 30:331-336.

Granath LE (2005). Answers to questions regarding dental flossing for caries control (personal communication). Seattle.

Granath LE, Martinsson T, Matsson L, Nilsson G, Schroder U, Söderholm B (1979). Intraindividual effect of daily supervised flossing on caries in schoolchildren. *Community Dent Oral Epidemiol* 7:147-150.

Guevara JP, Berlin JA, Wolf FM (2004). Meta-analytic methods for pooling rates when follow-up duration varies: a case study. *BMC Med Res Methodol* 4:17.

Hausen H, Karkkainen S, Seppä L (2000). Application of the high-risk strategy to control dental caries. *Community Dent Oral Epidemiol* 28:26-34.

Higgins JP, Green S, editors (updated May 2005). Cochrane handbook for systematic reviews of interventions 4.2.5. In: The Cochrane Library,

Issue 3, 2005. Chichester, UK: John Wiley & Sons, Ltd.
Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ* 327:557-560.
Horowitz AM, Suomi JD, Peterson JK, Lyman BA (1977). Effects of supervised daily dental plaque removal by children: II. 24 months' results. *J Public Health Dent* 37:180-188.
Igarashi K, Lee IK, Schachtele CF (1989). Comparison of in vivo human dental plaque pH changes within artificial fissures and at interproximal sites. *Caries Res* 23:417-422.
Ischemic Optic Neuropathy Decompression Trial: twenty-four-month update (2000). *Arch Ophthalmol* 118:793-798.
Kallestal C (2005). The effect of five years' implementation of caries-preventive methods in Swedish high-risk adolescents. *Caries Res* 39:20-26.
Koch G (1967). Effect of sodium fluoride in dentifrice and mouthwash on incidence of dental caries in school children. *Odontol Rev* 18.
Loesche WJ (1993). Dental caries: a treatable infection. Ann Arbor, MI: Automated Diagnostic Documentation Inc.
Mäkinen KK, Bennett CA, Hujoel PP, Isokangas PJ, Isotupa KP, Pape HR Jr, *et al.* (1995). Xylitol chewing gums and caries rates: a 40-month cohort study. *J Dent Res* 74:1904-1913.
Mäkinen KK, Hujoel PP, Bennett CA, Isotupa KP, Mäkinen PL, Allen P (1996). Polyol chewing gums and caries rates in primary dentition: a 24-month cohort study. *Caries Res* 30:408-417.
Marinho VC, Higgins JP, Sheiham A, Logan S (2004). One topical fluoride (toothpastes, or mouthrinses, or gels, or varnishes) versus another for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev* 1:CD002780.
McKee DP, Faine RC, Murphy RF (1977). The effectiveness of a dental health education program in a nonfluoridated community. *J Public Health Dent* 37:290-299.
Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, *et al.* (2002). A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 347:81-88.
Parmly LS (1819). A practical guide to the management of the teeth; comprising a discovery of the origin of caries, or decay of the teeth. Philadelphia: Collins & Croft.
Petitti DB (1994). Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine. New York: Oxford University Press.
Poulsen S, Agerbaek N, Melsen B, Korts DC, Glavind L, Rølla G (1976). The effect of professional toothcleansing on gingivitis and dental caries in children after 1 year. *Community Dent Oral Epidemiol* 4:195-199.
Ratcliff PA (1966). Review of literature on periodontal therapy—for World Conference on Periodontal Research. Ann Arbor, MI: The University of Michigan.
Rothman KJ, Greenland S, Sonis J (1998). Modern epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven.
Rule JT, Smith MR, Truelove RB, Macko DJ, Castaldi CR (1984). Caries inhibition of a dentifrice containing 0.78% sodium monofluorophosphate in a silica base. *Community Dent Oral Epidemiol* 12:213-217.
Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB (2000). Evidence-based medicine: how to practice and teach EBM. Edinburgh: Churchill Livingstone.
Seppä L, Hausen H, Pollanen L, Karkkainen S, Helasharju K (1991). Effect of intensified caries prevention on approximal caries in adolescents with high caries risk. *Caries Res* 25:392-395.
Silverstein SJ, Gold S, Heilbron D, Nelms D, Wycoff S (1977). Effect of supervised deplaquing on dental caries, gingivitis, and plaque (abstract). *J Dent Res* 56(Spec Iss A):A85.
Thompson SG, Higgins JP (2002). How should meta-regression analyses be undertaken and interpreted? *Stat Med* 21:1559-1573.
Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, *et al.* (1998). The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 51:1235-1241.
Waerhaug J (1981). Healing of the dento-epithelial junction following the use of dental floss. *J Clin Periodontol* 8:144-150.
Wright GZ, Banting DW, Feasby WH (1979). The Dorchester dental flossing study: final report. *Clin Prev Dent* 1:23-26.
Wright GZ, Feasby WH, Banting D (1980). The effectiveness of interdental flossing with and without a fluoride dentifrice. *Pediatr Dent* 2:105-109.

AMSTAR – a measurement tool to assess the methodological quality of systematic reviews.

1. Was an 'a priori' design provided?
The research question and inclusion criteria should be established before the conduct of the review.

Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a "yes."

□ Yes
□ No
□ Can't answer
□ Not applicable

---

2. Was there duplicate study selection and data extraction?
There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.

Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other's work.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

3. Was a comprehensive literature search performed?
At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.

Note: If at least 2 sources + one supplementary strategy used, select "yes" (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).

□ Yes
□ No
□ Can't answer
□ Not applicable

---

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.

Note: If review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

5. Was a list of studies (included and excluded) provided?
A list of included and excluded studies should be provided.

Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no."

□ Yes
□ No
□ Can't answer
□ Not applicable

---

6. Were the characteristics of the included studies provided?
In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.

Note: Acceptable if not in table format as long as they are described as above.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

7. Was the scientific quality of the included studies assessed and documented?
'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.

Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

□ Yes
□ No
□ Can't answer
□ Not applicable

---

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.

Note: Might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

9. Were the methods used to combine the findings of studies appropriate?
For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I^2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).

Note: Indicate "yes" if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

10. Was the likelihood of publication bias assessed?
An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).

Note: If no test values or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

□ Yes
□ No
□ Can't answer
□ Not applicable

---

11. Was the conflict of interest included?
Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

Note: To get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.

□ Yes
□ No
□ Can't answer
□ Not applicable

Shea *et al. BMC Medical Research Methodology* 2007 **7**:10 doi:10.1186/1471-2288-7-10

Additional notes (in italics) made by Michelle Weir, Julia Worswick, and Carolyn Wayne based on conversations with Bev Shea and/or Jeremy Grimshaw in June and October 2008 and July and September 2010.