



IN BRIEF

- The reasons for conducting a systematic review
- The advantages and disadvantages of a meta-analysis
- The calculation and evaluation of the effect of interest in a meta-analysis
- An assessment of the compatibility of the trials in a meta-analysis
- A description of some plots used in meta-analysis

Further statistics in dentistry Part 8: Systematic reviews and meta-analyses

A. Petrie¹ J. S. Bulman² and J. F. Osborn³

A **systematic review** of research evidence is an efficient approach to integrating existing information, invariably a multiplicity of published articles, with a view to establishing whether the scientific findings are consistent. If so, it may be possible to draw conclusions and make recommendations about treatment regimens or observed effects which have greater credence than those obtained from individual studies. The systematic review relies on a specified checklist which determines which articles should be included in the review, and how each should be critically appraised to provide relevant information relating to the focus of the review.

FURTHER STATISTICS IN DENTISTRY:

1. Research designs 1
2. Research designs 2
3. Clinical trials 1
4. Clinical trials 2
5. Diagnostic tests for oral conditions
6. Multiple linear regression
7. Repeated measures
8. Systematic reviews and meta-analyses
9. Bayesian statistics
10. Sherlock Holmes, evidence and evidence-based dentistry

SYSTEMATIC REVIEWS

The report of a systematic review is somewhat like that of a research paper; it contains a clear description of the aims, and the material and methods used by the reviewer. The alternative haphazard non-systematic review has no defined rules concerning the process of digesting the mass of information, and is open to abuse.

A systematic review serves various purposes:

- It reduces a large amount of information to a manageable size. This information can be assimilated quickly by healthcare providers, researchers and policy makers. At the initial stage, the systematic review distinguishes between those studies that are essentially unsound and those that provide useful and scientifically worthwhile results in relation to the question of interest.
- By combining the results from various studies which may have been conducted in slightly varying circumstances (eg using different definitions of disease or patient eligibility criteria), it may be possible to determine from the systematic review whether the results are consistent from study to study, and to generalise the results. Furthermore, a systematic review may offer the opportunity to explain any inconsistencies.
- It is usually cheaper and quicker to conduct a systematic review than to embark on a new study.

- It may reduce the delay between research discoveries and the implementation of new effective treatment strategies.
- The systematic review combines information from individual studies so that its overall sample size is greater than that of any one study, and this leads to an increase in the power of the investigation. Thus, the systematic review has a greater chance of eliciting significant treatment effects, which is particularly helpful if the prevalence of the condition is low or if the effect of interest is small.
- The systematic review has an increased sample size compared with any individual study so the estimates of the effects of interest are obtained with increased precision.
- A systematic review limits bias and improves the reliability and accuracy of recommendations because of its formalised and thorough method of investigation.

The **Cochrane Collaboration** (www.update-software.com/ccweb/cochrane/general.htm) is an international network of individuals and institutions which prepares systematic reviews of randomised controlled studies and of observational evidence. It helps to promote the development of systematic reviews by setting explicit standards for systematic reviews. It provides a framework within which scientists of like interests can collaborate, and through its publication, the *Cochrane Database of Systematic Reviews*,

¹Senior Lecturer in Statistics, Eastman Dental Institute for Oral Health Care Sciences, University College London;
²Honorary Reader in Dental Public Health, Eastman Dental Institute for Oral Health Care Sciences, University College London;
³Professor of Epidemiological Methods, University of Rome, La Sapienza
 Correspondence to: Aviva Petrie, Senior Lecturer in Statistics, Biostatistics Unit, Eastman Dental Institute for Oral Health Care Sciences, University College London, 256 Gray's Inn Road, London WC1X 8LD
 E-mail: a.petrie@eastman.ucl.ac.uk

allows electronic access to the latest detailed and highly structured reports on subjects of interest.

META-ANALYSIS

A special form of systematic review is a **meta-analysis** (sometimes called an overview); this is a statistical approach to combining the results from separate but similar studies to provide an overall *quantitative* summary of the effect of interest. A meta-analysis is thus a statistical analysis of a collection of statistical analyses from individual studies. Full details of the theory of meta-analysis may be obtained in Hedges and Olkin (1985).¹ In addition, a paper by Song *et al.* (1997)² provides a useful discussion of how to handle discrepancies in recommendations arising from different meta-analyses of what appear to be the same research question.

In principle, a meta-analysis proffers the advantages of increased power, and increased precision of its estimates, when compared with a single study. In practice, the meta-analysis is open to criticism, essentially on four grounds (Glass *et al.*, 1981):³

1. Because journals rarely publish studies in which the findings are non-significant, published research is biased in favour of significant results. A trial with a significant result is sometimes called a *positive* trial; a *negative* trial is one in which a clinically significant effect is essentially ruled out. This **publication bias** leads to biased meta-analysis results unless the meta-analyst makes a serious attempt to identify and use the results in books, dissertations, unpublished papers presented at professional meetings or located in retrieval systems for unpublished papers (such as SIGLE produced by the European Association for Grey Literature), etc.
2. The studies included in the meta-analysis may differ in respect of features such as design, outcome measure, measuring technique, definition of variables and subjects, and duration of follow-up. Such **clinical heterogeneity** needs to be explored carefully as it may affect the overall conclusions and the clinical implications of the review. Generally, a meta-analysis of clinical trials is restricted to include only those trials that are *randomised*. Additional requirements of blind or objective assessment of response, ideally with analysis by intention-to-treat and complete follow-up, are sometimes imposed (Peto, 1987).⁴ Such trials are less likely to lead to biased results than those which do not possess these attributes.
3. The studies included in the meta-analysis may vary in their quality, and it has been shown (Jahad, 1996)⁵ that a meta-analysis which comprises studies of high quality, as opposed to poor quality, tends to be less enthusiastic about an intervention. However, it can be argued that poorly designed or 'bad' studies should be included in the meta-analysis because of the inclusiveness of the method and the subjective nature of the considera-

tions which might lead to their exclusion. The problem of including both 'good' and 'bad' studies can be handled empirically by conducting separate analyses for groups of studies of similar quality, and examining whether the results differ for poorly and well designed studies. Sometimes, the results from all the studies are combined by assigning weights to the studies according to their relative quality, but this approach can be criticised on grounds of the arbitrariness of the assignment.

4. The results included in the meta-analysis may not be independent. This situation arises when a multivariate study provides more than one test of significance relevant to the hypothesis that the meta-analysis is examining. Also, non-independence of the results may arise when the studies are conducted by the same investigator at different times, or by different investigators who have communicated with each other and modified their studies on the basis of earlier results. Furthermore, some trials are published more than once.

EXAMPLE

Meta-analyses in dentistry are not very common. The example used in this paper, a meta-analysis by van Rijkom *et al.* (1998),⁶ can be criticised but is, nevertheless, thorough and accessible. The authors used a meta-analysis to estimate the overall caries inhibiting effect of fluoride gels applied to the permanent teeth of children aged 6 to 15 years. Each of the 19 studies included in the analysis, referenced by study number at the end of this paper, was obtained from a MEDLINE search of the published literature of English and German studies. All these studies satisfied various selection inclusion and exclusion criteria, and their follow-up periods were between 1.5 and 3 years (median 3 years). In particular, each of the chosen studies was a randomised controlled trial in which the effect of the fluoride gel treatment was compared with no treatment or placebo treatment. In fact, some of the 19 studies were independent substudies of a larger study which had been split into two to reflect differences in general fluoride regimen. The inhibiting effect of the treatment was expressed for each study by the prevented fraction (PF); this was calculated as the difference in the incidence between the decayed, missing and filled surfaces (DMFS) in the control group (lc) and the incidence in the experimental group (le), divided by that in the control group [ie $PF = (lc - le)/lc$]. The absolute difference between the incidences in the two groups was standardized (ie divided by lc) since the PF was assumed to be less sensitive to experimental circumstances, such as the age range of the study population and the duration of the study, than $(lc - le)$.

THE EFFECT OF INTEREST IN A META-ANALYSIS

Explaining the effect of interest

Suppose that the meta-analysis comprises k studies, and that θ_i denotes the value of an

Meta-analysis

A meta-analysis integrates the quantitative findings from separate but similar studies and provides a numerical estimate of the overall effect of interest



appropriate measure of the **effect of interest** in the i^{th} study. In a clinical trial, this will usually be the effect of the experimental treatment relative to the control treatment. If the outcome variable is quantitative, θ_i is typically the difference between the experimental and control treatment means in the population or some standardized version of this difference. If the outcome is binary, for example 'success' or 'failure', θ_i is often the logarithm of the odds ratio or relative risk. In the fluoride gel example, the effect of interest is the prevented fraction (PF), expressed as a percentage.

There are two approaches to combining the information in a meta-analysis. The **parametric approach** is usually adopted; this assumes that the effect of interest in each study is Normally distributed. Note that both the difference between the means and its standardized difference are Normally distributed for Normally distributed data; similarly, the logarithm of the relative risk, equal to the difference in the logarithms of the two risks, is approximately Normally distributed. The PF in this example is assumed to be approximated Normally distributed. The parametric approach focuses on combining the results from the k studies, estimating the overall effect of interest, with its confidence interval, testing its significance and interpreting these results. Occasionally, a **non-parametric approach** is used which makes no distributional assumptions about the effect of interest. However, the non-parametric methods often require the raw data from each study, which can limit their use, and they are not described here.

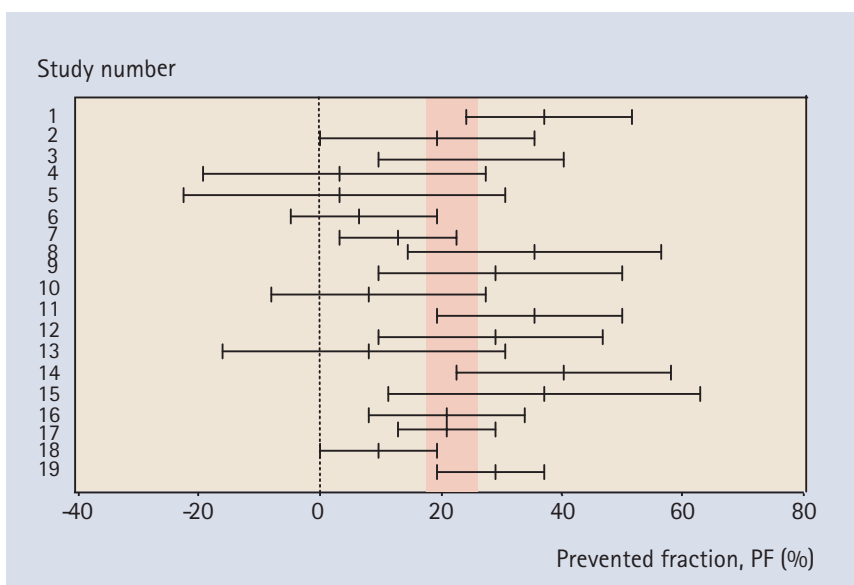


Fig. 1 Caries-inhibiting effect of fluoride gel treatment (PF with 95% confidence intervals) in 19 studies. The shaded area shows the 95% confidence interval for the pooled PF.⁶

Displaying the effect of interest

Initially, it is helpful to **display** the quantitative results from each study both in tabular and diagrammatic form. The table includes the relevant information on each trial; for example, the sample size, baseline patient characteristics, information on inclusion criteria and withdrawal rates, and the effects of interest, such as the odds ratio. Table 1 summarises some of the more important features of the selected studies in the fluoride gel example.

Table 1 Summary of study results arranged according to application method and application frequency⁶

Study no [‡]	Fluoride regimen	Applic. Freq. Times/year	Method [†]	Control group			Experimental group			PF (%)	95% CI (%)
				Incident DMFS	SEM	n [‡]	Incident DMFS	SEM	n [‡]		
1	1	1	1	7.26	0.46	105	5.46	0.44	103	37	23 to 52
2	2	1	1	3.82	0.29	311	3.14	0.23	278	18	0 to 35
3	2	1	1	8.61	0.61	170	6.51	0.50	182	24	9 to 40
4	1	1	1	2.15	0.17	170	2.07	0.20	148	4	-20 to 28
5	2	1	1	1.91	0.17	153	1.83	0.20	163	4	-23 to 31
6	2	2	1	7.19	0.32	311	6.68	0.32	296	7	-5 to 19
7	1	2	1	8.52	0.37	316	7.39	0.32	315	13	3 to 24
8	2	2	1	8.15	0.87	78	5.28	0.66	115	35	14 to 57
9	2	2	1	4.40	0.38	103	3.08	0.37	108	30	9 to 51
10	1	2	1	3.24	0.22	207	2.94	0.21	224	9	-8 to 27
11	2	4	1	4.21	0.32	166	2.74	0.26	145	35	19 to 51
12	1	90	1	2.20	0.20	220	1.57	0.16	337	29	9 to 48
13	1	6	2	4.26	0.37	148	3.94	0.35	161	8	-15 to 30
14	1	30	2	4.40	0.42	57	2.63	0.32	63	40	22 to 59
15	2	30	2	3.44	0.38	43	2.18	0.38	38	37	11 to 63
16	1	360	3	6.86	0.37	242	5.40	0.34	253	21	8 to 34
17	1	360	3	8.34	0.28	449	6.60	0.24	451	21	13 to 29
18	1	360	3	7.70	0.29	380	6.94	0.27	378	10	0 to 20
19	1	360	3	6.39	0.23	416	4.56	0.22	460	29	20 to 37

^{*}Fluoride regimen: 1 = assumed non-fluoride toothpaste, 2 = assumed fluoride toothpaste or non-fluoride toothpaste including fluoride water

[†]Method: 1 = tray > 1%, 2 = brush > 1%, 3 = brush < 1%

[‡]n: study size [‡]These study numbers are referenced at the end of the paper

The most usual pictorial representation (Fig. 1) is sometimes called a 'forest plot'. It shows the estimated effect of interest (in the fluoride gel example it is the PF but might, in other circumstances, be the standardized difference in means or the odds ratio) for each of the separate studies in the meta-analysis. The confidence intervals for the true effect in each case, as well as the overall estimated effect (and related confidence interval) from the pooled data from all the studies, are also indicated. An explanation of the method used to calculate the overall estimated effect is given in the section entitled 'calculating the effect of interest'. A vertical line, known as the 'line of no effect', is sometimes drawn in the diagram. It represents equal effectiveness of the treatments (for example, it would correspond to a value of zero for PF or a difference in means, or unity if the effect of interest were the odds ratio). In the fluoride gel example, only five of the confidence intervals for the true PFs cross the line of no effect, whereas twelve of the confidence intervals are to the right of it; this suggests that fluoride gel is an effective inhibitor of caries.

It is possible to get some idea of whether the estimates of the effects from the different studies are compatible by 'eye-balling' the forest plot. If the confidence intervals of the effects overlap, then the trials are likely to be compatible, whereas if there is no overlap, then they are incompatible. The confidence intervals in Fig. 1 show considerable overlap, suggesting that the results of the different studies are likely to be compatible. However, it should be noted that the estimated PFs show substantial variation, so this conclusion should be viewed with caution.

Checking for compatibility between the trials

A more formal approach to determining incompatibility is to perform a statistical test, described in detail in the section entitled 'Testing for homogeneity'. If, on the basis of the test result, the observed effects are more dispersed than would be expected on the basis of chance alone, **statistical heterogeneity** is said to be present, ie the estimated effects exhibit considerable variation and are incompatible. Statistical heterogeneity may be caused by clinical heterogeneity, methodological differences or it may be related to unknown trial characteristics. The presence of statistical heterogeneity is indicated if the test of statistical *homogeneity* (homogeneity implies that the effects are equal) is significant. If the test is not significant, this does not imply that there must be statistical homogeneity. A non-significant result implies only that there is no evidence to reject the null hypothesis of homogeneity, and not that there is evidence to accept it. It should be pointed out that the test of statistical homogeneity has low power and, therefore, may fail to produce a statistically significant result unless there is marked heterogeneity. Whether or not the test is significant, it is important to provide an estimate of the extent to which

there is statistical heterogeneity. Then, if this estimate indicates that there might be substantial statistical heterogeneity, the aspects of clinical heterogeneity which may be causing it should be investigated.

Calculating the effect of interest

The **overall estimate of effect** is usually taken as a weighted average of the estimates from the k individual studies in the meta-analysis. Thus the estimate of the overall effect is:

$$\hat{\theta} = \frac{\sum \hat{\theta}_i w_i}{\sum w_i}$$

where $\hat{\theta}_i$ is the estimated effect and w_i the weight of the i^{th} study, and the sums extend over all k studies. Usually the weights are chosen to be inversely related to the variances of the estimated effects and this is approximately the same as choosing weights which are proportional to the sample size, so that the larger studies are given more weight than the smaller ones.

- A **fixed-effects** estimation method can be used if there is no evidence of statistical heterogeneity and/or if the meta-analysis comprises only a small number of studies. This approach assumes that the separate studies are the only ones that are of interest, and that the underlying true effect in each study is the same and equal to θ , ie $\theta_1 = \theta_2 = \dots = \theta_k = \theta$. The variance of the estimated effect in each study comprises only the random variation in that study. When estimating the assumed common effect, θ , the weight attached to each θ_i is the reciprocal of its variance, so that more weight is given to the more precise estimate (ie that with a narrower confidence interval).
- If statistical heterogeneity is believed to be present, the **random-effects** estimation method may be appropriate. This approach assumes that the k separate studies are a random sample from a larger population of studies, and there is a population effect of interest, θ , about which the effects of the individual studies vary. Then $\hat{\theta}$ is an estimate of θ which is now the mean of the effects of interest obtained from all the studies in the population. As in the fixed-effects method, the weight for the i^{th} study using the random-effects method is chosen to be the reciprocal of the variance of the estimated effect for that study. However, unlike the fixed-effects method, the random-effects method incorporates both the random variation within the study and the heterogeneity between the different studies into this variance. It produces a wider confidence interval for the overall estimate than the fixed-effects method, as would be expected from an estimate that reflects the heterogeneity of the estimates. It is important to remember that if there is statistical heterogeneity, caution must be adopted when interpreting the

Heterogeneity between studies



- **Statistical heterogeneity** exists when the quantitative results from the studies in a meta-analysis exhibit considerable variation.
- **Clinical heterogeneity** exists when the studies differ in features such as the design, outcome measure, definition of variables and subjects, etc.

overall estimate of effect, however it is derived, and the reasons for the heterogeneity should be investigated.

Note that for both the fixed-effects and random-effects approaches, an approximate 95% **confidence interval** for the overall estimate of effect, θ , is given by $\hat{\theta} \pm 1.96\sqrt{1/\sum w_i}$. The w_i is the reciprocal of the variance of the estimated effect in the i^{th} study. It is a measure of the random variation within the study for the fixed-effects approach, but includes the variation between the k estimated effects as well for the random-effects approach.

The authors in the fluoride gel example used two approaches to investigate heterogeneity. Firstly, they believed that the large overlap in the confidence intervals in the forest plot was an indication that there was no evidence of statistical heterogeneity. However, although there was considerable overlap in the confidence intervals, the estimates from the different studies showed substantial variation. Secondly, they used a multiple regression analysis to determine whether there were factors influencing the caries-inhibiting effect of fluoride gel application. This analysis showed no significant influence of the covariables that were thought could be relevant, namely, 'application frequency', 'application methods' (tray/brush), 'baseline caries prevalence' and 'general fluoride regimen'. Thus, they concluded that all studies could be regarded as equally effective, and the overall effect could be estimated using a fixed effects model. The weight for each study was chosen to be the inverse of the variance of the prevented fraction, PF; this gave an estimated overall PF of 22%. The 95% confidence interval for the true PF was 18% to 25%, the shaded area in Fig. 1; this excludes zero and suggests that fluoride gel was an effective inhibitor of caries in children of this age.

HYPOTHESIS TESTS IN META-ANALYSIS

There are two hypothesis tests that are of crucial importance in a meta-analysis, one which tests for homogeneity of the effects of interest and the other which tests the significance of the overall treatment effect.

Testing for homogeneity

The test of the null hypothesis that the studies are homogenous with respect to their effects of interest should be performed initially, rather than relying solely on the subjective opinion obtained from the forest plot. The test of this hypothesis, sometimes called that of 'combinability', is usually based on the magnitude of the test statistic

$$Q = \sum w_i (\hat{\theta}_i - \hat{\theta})^2$$

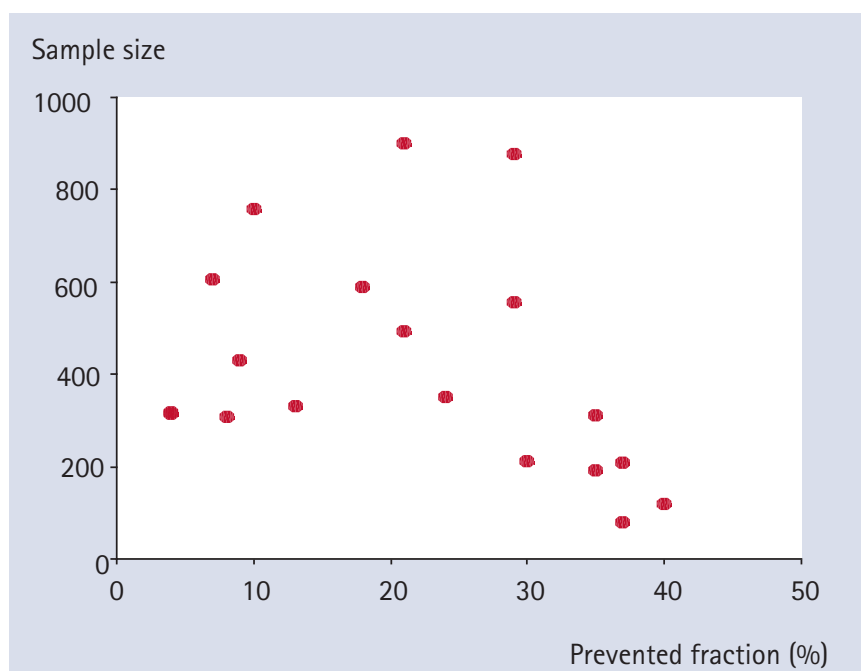
which is assumed to follow a chi-squared distribution with $(k-1)$ degrees of freedom. As homogeneity is assumed under the null hypothesis in this test, the fixed-effects and the random-

effects approaches are not distinguished, and the weight, w_i , is the same as that used in the fixed-effects approach, i.e. it is the reciprocal of the variance of the effect in the i^{th} study ($i = 1, 2, 3, \dots, k$), where that variance is a measure only of the random variation within the study.

It is interesting to note that, even though the confidence intervals for the PF's from the separate studies overlap (Fig. 1), the chi-squared test of homogeneity in the fluoride gel example gives a result which is marginally significant at the 5% level. This, together with the view that the different studies have estimated PFs which show considerable variation, should perhaps be an indication that combining the estimates of PF is questionable and that the overall estimate of the PF should be interpreted with caution.

In addition to the overall test which investigates heterogeneity (in fact, it tests homogeneity), it is possible to test for *funnel plot asymmetry* which assesses bias. Details may be obtained from Eggar *et al.*, (1997).⁷ A **funnel plot** (Fig. 2) is a scatter plot of the sample size against the treatment effect estimate generated in an individual study. The sample size can be replaced by the precision of the estimated effect. Since the precision of the estimated treatment effect increases as the sample size of the component study increases, the results from small studies would be expected to show a wide scatter at the bottom of the graph, with the spread decreasing (narrowing to produce a funnel effect) at the top of the graph for the larger studies. The funnel plot will often be skewed and asymmetrical if bias is present, as demonstrated in Fig. 2. The lower left corner of the funnel is somewhat empty (ie lacking publications), indicating that some studies on small sample sizes with small effects are probably missing. The effect of this publication bias on the overall PF, however, is likely to be marginal, because the weight of such unpublished low-power studies is small. It is

Fig. 2 Funnel plot of the sample size for each study plotted against the estimated prevented fraction, PF (%) for the 19 studies included in the meta-analysis



possible to measure the degree of asymmetry in a funnel plot, but the approach is of limited value if only a few trials are included (remembering that the unit of analysis is the randomised trial and not its patients).

Testing the treatment effect

The null hypothesis that the true effect of interest is zero (ie $H_0: \theta_1 = \theta_2 = \dots = \theta_k = 0$ in the fixed-effects method, or $H_0: \theta = 0$ in the random-effects method) is tested using the test statistic:

$$U = \frac{(\sum \hat{\theta}_i w_i)^2}{\sum w_i}$$

which follows the chi-squared distribution with one degree of freedom. Each w_i is the reciprocal of the variance of the estimated effect of the i^{th} study. In the fixed-effects approach, this variance is a measure only of the random variation within the study; in the random-effects method, the variance comprises both the random variation within the study and the variation between the estimated effects in the k studies. The test for the overall PF in the fluoride gel example gives a highly significant result ($P < 0.001$) indicating

that the overall effect, estimated by a PF of 22% (95% confidence interval equal to 18% to 25%), is significantly different from zero. This implies that the fluoride gel is an effective inhibitor of caries in children of this age. Note that, as observed previously, zero lies outside the 95% confidence interval for the overall PF, as expected if the test result is significant.

The authors would like to thank Dr James Lewsey for his comments on earlier drafts of the manuscript.

- Hedges LV, Olkin I. *Statistical Methods for Meta-analysis*. Academic Press: New York, 1985.
- Song F, Landes D P, Glenny A-M, Sheldon T A. Prophylactic removal of impacted third molars: an assessment of published reviews. *Br Dent J* 1997; **182**: 339-346.
- Glass G, McGaw B, Smith M J. *Meta-analysis in Social Research*. Beverly Hills CA: Sage, 1981.
- Peto R. Why do we need systematic overviews of randomised trials? *Stat Med* 1987; **6**: 233-240.
- Jahad A R, McQuay H J. Meta analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996; **49**: 235-243.
- van Rijkom H M, Truin G J, van't Hof M A. A meta-analysis of clinical studies on the caries-inhibiting effect of fluoride gel treatment. *Caries Res* 1998; **32**: 83-92.
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; **315**: 629-634.

Studies, labelled by study number, included in the meta analysis of van Rijkom et al. (1998)

- Bryan FT, Williams J E. The cariostatic effectiveness of a phosphate-fluoride gel administered annually to school children: final results. *J Public Health Dent* 1970; **30**: 13-16.
- Cons N C, Janerich DT, Senning R S. Albany topical fluoride study. *JADA* 1970; **2**: 263-266.
- Horowitz H S, Doyle J. The effect on dental caries of topical applied acidulated phosphate-fluoride: Results after three years. *JADA* 1971; **82**: 359-365.
- Szwejdja L F. Fluorides in community programs: Results after two years from a fluoride gel applied topically. *J Public Health Dent* 1971; **31**: 241-242.
- Szwejdja L F. Fluorides in community programs: A study of four years of various fluorides applied topically to the teeth of children in fluoridated communities. *J Public Health Dent* 1972; **32**: 25-31.
- Mainwaring P J, Naylor M N. A three-year clinical study to determine the separate and combined caries-inhibiting effects of sodium monofluoro-phosphate toothpaste and an acidulated phosphate-fluoride gel. *Caries Res* 1978; **12**: 202-212 (Sub-study with non-fluoride toothpaste).
- Mainwaring P J, Naylor M N. A three-year clinical study to determine the separate and combined caries-inhibiting effects of sodium monofluoro-phosphate toothpaste and an acidulated phosphate-fluoride gel. *Caries Res* 1978; **12**: 202-212 (Sub-study with fluoride toothpaste).
- Cobb H B, Rozier R G, Bawden J W. A clinical study of the caries preventive effects of an APF solution and an APF thixotropic gel. *Pediatr Dent* 1980; **2**: 263-266.
- Hagan P P, Bawden J W. The caries-preventive effects of full- and half-strength topical and acidulated phosphate fluoride. *Pediatr Dent* 1985; **7**: 185-191.
- Olivier M, Brodeur J-M, Simard P L. Efficacy of APF treatments without prior toothcleaning targeted to high-risk children. *Community Dent Oral Epidemiol* 1992; **20**: 38-42.
- Trubman G J, Crellin J A. Effect on dental caries of self-application of acidulated phosphate fluoride paste and gel. *JADA* 1973; **86**: 153-157.
- Englander H R, Sherrill L T, Miller B G, Carlos J P, Mellberg J R, Senning R S. Incremental rates of dental caries after repeated topical sodium fluoride applications in children with lifelong consumption of fluoridated water. *JADA* 1971; **82**: 354-358.
- Heifeetz S B, Horowitz H S, Driskill W S. Two-year evaluation of a self-administered procedure for the topical application of acidulated phosphate-fluoride: Final report. *J Public Health Dent* 1970; **30**: 7-12.
- Marthaler T M, König K G, Mühlemann H R. The effect of a fluoride gel use for supervised tooth-brushing 15 or 30 times per year. *Helv Odont Acta* 1970; **14**: 67-77 (Sub-study with non-fluoride toothpaste).
- Marthaler T M, König K G, Mühlemann H R. The effect of a fluoride gel use for supervised tooth-brushing 15 or 30 times per year. *Helv Odont Acta* 1970; **14**: 67-77 (Sub-study with fluoride toothpaste).
- Howat A P, Holloway P J, Davies T G H. Caries prevention by daily supervised use of a MPF gel dentifrice. *Br J Dent* 1978; **145**: 233-235.
- Fogels H R, Alman J E, Meade I J, O'Donnell J P. The relative caries-inhibiting effects of a stannous fluoride dentifrice in a silica gel base. *JADA* 1979; **99**: 456-459.
- Abrams R G, Chambers D W. Caries-inhibiting effect of a stannous fluoride silica gel dentifrice: A three-year clinical study. *Clin Prev Dent* 1980; **2**: 22-27.
- Rule J T, Smith M R, Truelove R B, Macko D J, Castaldi C D. Caries inhibition of a dentifrice containing 0.78% sodium monofluorophosphate ion in a silica base. *Community Dent Oral Epidemiol* 1984; **12**: 213-217.