

## Systematic review of water fluoridation

Marian S McDonagh, Penny F Whiting, Paul M Wilson, Alex J Sutton, Ivor Chestnutt, Jan Cooper, Kate Misso, Matthew Bradley, Elizabeth Treasure, Jos Kleijnen



### Abstract

**Objective** To review the safety and efficacy of fluoridation of drinking water.

**Design** Search of 25 electronic databases and world wide web. Relevant journals hand searched; further information requested from authors. Inclusion criteria were a predefined hierarchy of evidence and objectives. Study validity was assessed with checklists. Two reviewers independently screened sources, extracted data, and assessed validity.

**Main outcome measures** Decayed, missing, and filled primary/permanent teeth. Proportion of children without caries. Measure of effect was the difference in change in prevalence of caries from baseline to final examination in fluoridated compared with control areas. For potential adverse effects, all outcomes reported were used.

**Results** 214 studies were included. The quality of studies was low to moderate. Water fluoridation was associated with an increased proportion of children without caries and a reduction in the number of teeth affected by caries. The range (median) of mean differences in the proportion of children without caries was -5.0% to 64% (14.6%). The range (median) of mean change in decayed, missing, and filled primary/permanent teeth was 0.5 to 4.4 (2.25) teeth. A dose-dependent increase in dental fluorosis was found. At a fluoride level of 1 ppm an estimated 12.5% (95% confidence interval 7.0% to 21.5%) of exposed people would have fluorosis that they would find aesthetically concerning.

**Conclusions** The evidence of a beneficial reduction in caries should be considered together with the increased prevalence of dental fluorosis. There was no clear evidence of other potential adverse effects.

### Introduction

In the white paper, *Saving Lives: Our Healthier Nation*, the UK government highlighted the commonly held belief that there is strong evidence that water fluoridation improves and considerably reduces inequality in dental health.<sup>1</sup> The government also acknowledged that "the extensive research linking water fluoridation to improved dental health was mostly undertaken a few years ago," and as a result this study was commissioned to provide a comprehensive systematic

review of the safety and efficacy of fluoridation of the public water supply.

We focused on the two main objectives: the effects of fluoridation of drinking water supplies on the incidence of caries and whether fluoridation has negative effects. The full report is available elsewhere.<sup>2</sup>

### Methods

#### Search strategy

We searched 25 specialist databases, including Medline, Embase, TOXLINE, and Current Contents (Science Citation Index) from inception of the database to February 2000. In addition, we hand searched Index Medicus (1945-63) and Excerpta Medica (1955-73). Further searches included the world wide web and bibliographies of all included studies. We sought additional references from individuals and organisations through a dedicated web site for this review ([www.york.ac.uk/inst/crd/fluorid.htm](http://www.york.ac.uk/inst/crd/fluorid.htm) includes the full report) and through members of a specifically designated advisory panel. Published and unpublished studies in any language were included. Full details of the search strategy are reported elsewhere.<sup>2</sup>

#### Inclusion criteria

We applied two types of inclusion criteria. The first was the level of evidence, based on the risk of bias. Studies were classified into the levels of evidence. Evidence rated below level B (moderate quality evidence, moderate risk of bias) was not considered in the evaluation of efficacy. In the assessment of safety all levels of evidence were considered. If a study met only one or two of three criteria for a given level of evidence, it was assigned the next level down. Details of both types of inclusion criteria can be found on the *BMJ's* website.

#### Data extraction and assessment of study quality

Inclusion criteria were assessed independently by at least two reviewers. Extraction of data from studies and assessment of validity was independently performed by two reviewers and checked by a third reviewer. Disagreements were resolved through consensus. We assessed study validity formally using a published checklist modified for this review.<sup>3</sup> Each item on the checklist was given one point, with a total of eight points possible for all study designs except case-control studies, which could attain a total of nine points.<sup>2</sup>

Editorial by Hausen

NHS Centre for Reviews and Dissemination, University of York, York YO10 5DD

Marian S McDonagh  
research fellow

Penny F Whiting  
research fellow

Paul M Wilson  
research fellow

Kate Misso  
information officer

Matthew Bradley  
research fellow

Jos Kleijnen  
professor

Dental Public Health Unit, Dental School, University of Wales College of Medicine, Cardiff CF14 4XY

Ivor Chestnutt  
senior lecturer

Jan Cooper  
research fellow

Elizabeth Treasure  
professor

continued over

*BMJ* 2000;321:855-9



Additional material comprising criteria for inclusion, members of the advisory panel, references (w1 etc) for included studies, and meta-regression table can be found on the *BMJ's* website

Department of  
Epidemiology and  
Public Health,  
University of  
Leicester, Leicester  
LE1 6TP

Alex J Sutton  
lecturer in medical  
statistics

Correspondence to:  
M McDonagh  
msm7@york.ac.uk

## Outcome measures

Studies that estimated the effect of fluoridation on caries investigated two main outcomes at baseline and at the final examination. These were decayed, missing, and filled primary/permanent teeth and the proportion of children without caries. The measure of effect used for the analysis was the difference of the change in prevalence of caries from baseline to the final examination in the fluoridated area compared with the control area in children of the same age.

To allow investigation of the effect of baseline levels of caries, we took the outcome measure from the final survey data for the meta-regressions of caries studies. The outcomes used were the data on effect size (mean

difference) for decayed, missing, and filled primary/permanent teeth and the data on difference in risk for the proportion of children without caries. This was done because correlation between the mean difference of the change in incidence of caries and baseline caries may lead to a spurious association. The median risk difference was used to calculate the number needed to treat for the proportion of children without caries.

Several indices are used to classify enamel opacities, including fluorosis. Dental fluorosis was defined here as any score other than normal on each index used. As the importance of a fluorosis score at the lowest level of each index is debatable, a second method was selected. This method describes the number of people who have dental fluorosis that may cause "aesthetic concern to the patient." The level at which fluorosis was judged to cause aesthetic concern was taken from a survey of 12 year old children in the United Kingdom<sup>10</sup> and corresponded to a tooth surface index of fluorosis score of two or more, a Thylstrup and Fejerskov index score of three or more, or Dean's classification of "mild" or worse. Studies that used other indices could not be included in this analysis. Full details of indices can be found elsewhere.<sup>2</sup>

## Analysis

Where the data were in a suitable format we plotted measures of effect and 95% confidence intervals. Heterogeneity was investigated by visual examination of plots and statistically with the Q statistic.<sup>4</sup> If we found significant heterogeneity we conducted meta-regression. Random effects models were adopted throughout to combine study results.<sup>5</sup> Meta-regression was used to explore the influence of study characteristics on outcome in an attempt to try to explain any heterogeneity between studies.<sup>4</sup> Stata version 6.0 (Stata Corporation, US) was used for this analysis.<sup>6</sup>

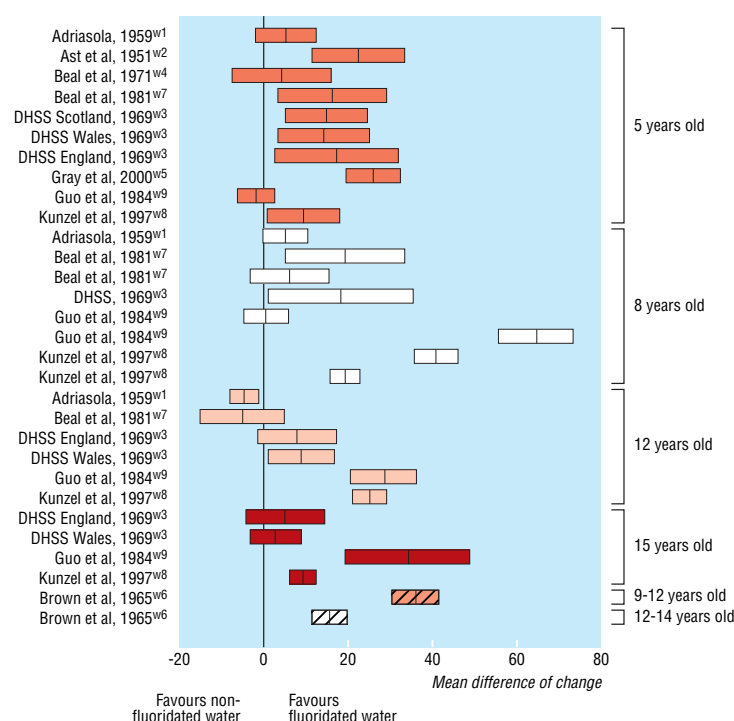
We used multi-level regression analysis to combine studies and investigate the association of water fluoride concentration with the prevalence of dental fluorosis (the analysis was conducted separately for all fluorosis and fluorosis of aesthetic concern) and used a multilevel model to combine studies. Each area with a different fluoride concentration under observation within a study was included separately in the model. The log (odds) of having fluorosis was modelled as a function of fluoride concentration. The analysis was carried out with the MIXED procedure within SAS (SAS Institute, US). Full details of methods used in the analyses, including all factors investigated in meta-regressions can be found elsewhere.<sup>2</sup>

## Results

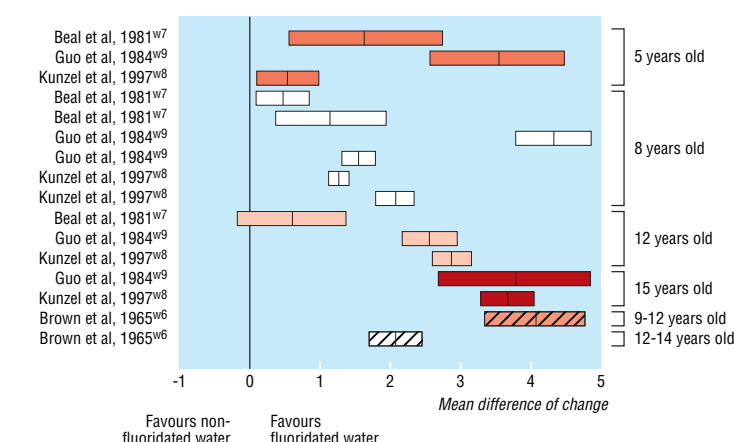
We included 214 studies; none was of evidence level A (high quality, bias unlikely). The study designs used included 45 controlled before-after studies, 102 cross sectional studies, 47 ecological studies, 13 cohort (prospective or retrospective) studies, and seven case-control studies. Summaries of individual study designs and full details on findings are available elsewhere.<sup>2</sup>

## Positive effects

Twenty six studies of the effect of water fluoridation on dental caries met the inclusion criteria. All but three of the studies included were controlled before-after studies. Of the three remaining, two used prospective



**Fig 1** Change in proportion (%) of children without caries in fluoridated compared with non-fluoridated areas (mean difference and 95% confidence interval)



**Fig 2** Change in decayed, missing, and filled teeth for primary/permanent teeth (mean difference and 95% confidence interval)

cohort designs and the other a retrospective cohort design. The controlled before-after studies assessed different groups of children of the same age (12 years) at the baseline (before fluoridation) and final (after fluoridation) surveys. All studies were of evidence level B (moderate), and the mean validity score was 5 (range 3.5 to 6.8) out of 8.

Figures 1 and 2 show estimates of the effect of fluoridation on the change in decayed, missing, and filled teeth and on the change in children without caries compared with control children for studies in which fluoridation was initiated after the baseline survey.<sup>w1-9</sup> Individual studies contributed more than one age group to the results. There was significant heterogeneity among the included studies ( $P < 0.001$ ).

The range (median) of the mean difference in the proportion (%) of children without caries was -5.0% to 64% (14.6%; interquartile range 5.05-22.1%). In the fluoridated areas there was a significant increase in the proportion of children without caries in 19 of 30 analyses. Only one analysis found a significant decrease in the proportion of children without caries in the fluoridated area. We estimate that that a median of six people would need to receive fluoridated water for one extra person to be free from caries (interquartile range of the distribution of number needed to treat was 4 to 9 people).

Fifteen of 16 analyses found a significantly greater mean change in decayed, missing, and filled primary/permanent teeth in the fluoridated areas than the non-fluoridated areas (fig 2). The range (median) of mean change in decayed, missing, and filled primary/permanent teeth was 0.5-4.4 (2.25) teeth (interquartile range 1.28-3.63 teeth).

Meta-regression showed that the proportion of children without caries at baseline, the setting, and the validity score show a significant association with the difference in risk in the proportion of children without caries. A table of the results of the meta-regression can be found on the *BMJ's* website. Baseline decayed, missing, and filled primary/permanent teeth, age, setting, and duration of study show a significant association with the mean difference in decayed, missing, and filled primary/permanent teeth.

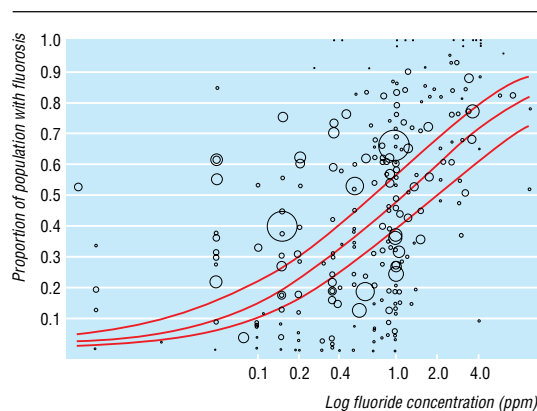
### Negative effects

A total of 175 included studies examined possible negative effects of water fluoridation.

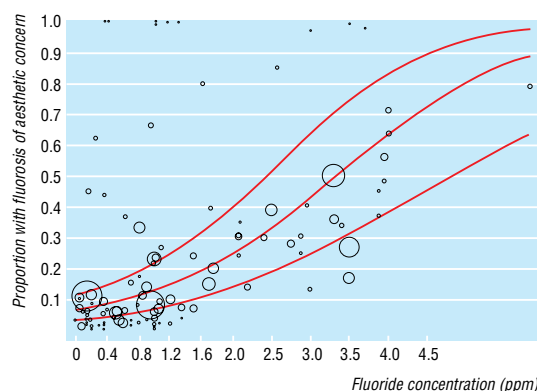
#### Dental fluorosis

We included 88 studies of dental fluorosis. These were largely cross sectional designs, with only four controlled before-after designs. The mean (range) validity score for fluorosis was only 2.8 (1.3-5.8) out of 8. All of the studies were of evidence level C (lowest quality), except one level B study. A full list of citations is available elsewhere.<sup>2</sup>

Regression analysis showed a significant dose-response relation for both methods of measuring the prevalence of fluorosis (figs 3 and 4). From these models, the pooled estimate of the prevalence of fluorosis at a water fluoride concentration of 1.0 ppm was 48% (95% confidence interval 40% to 57%) and for fluorosis of aesthetic concern 12.5% (7.0% to 21.5%). There was, however, considerable heterogeneity between results of individual studies.

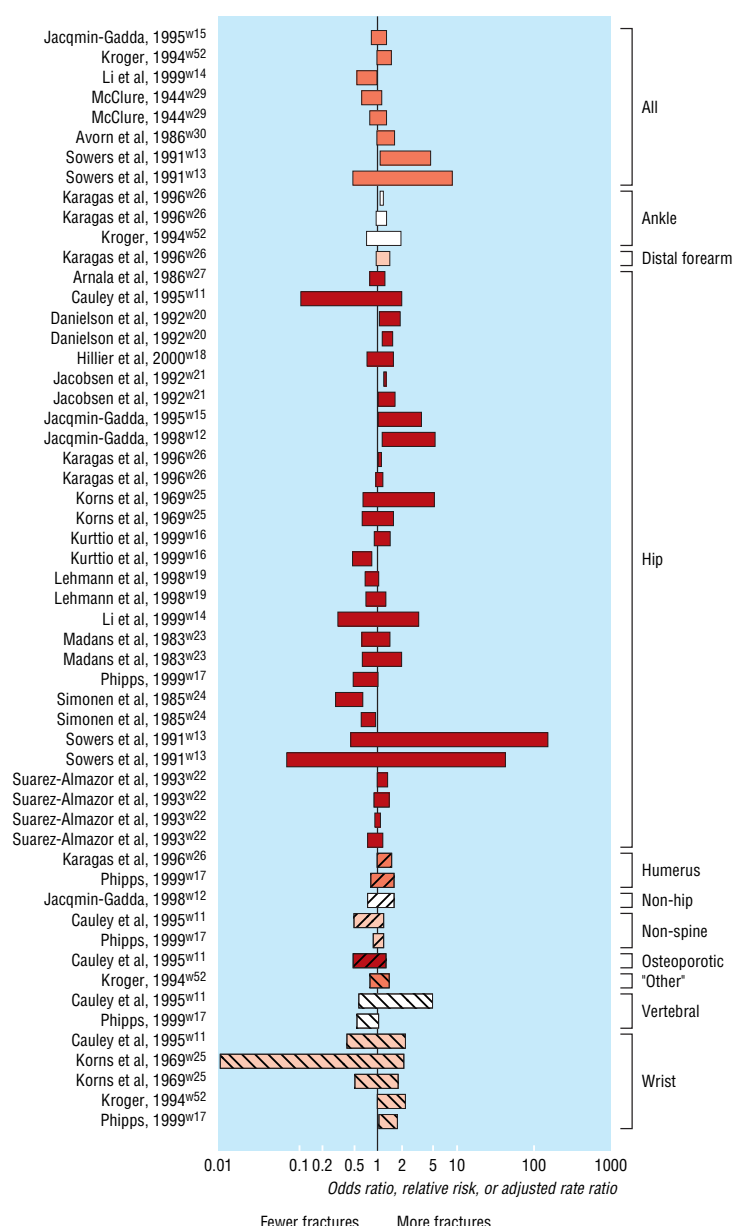


**Fig 3** Proportion of population with dental fluorosis by water fluoride concentration with 95% confidence interval for proportion. Fluoride concentration is plotted on log scale because of linear association between this and log (odds) of fluorosis. Each circle represents a study area in which the proportion of people with fluorosis is estimated—the larger the circle, the higher the precision of the estimate



**Fig 4** Proportion of population with fluorosis of aesthetic concern by water fluoride concentration (plotted on untransformed scale because of linear association between this and log (odds) of "aesthetic fluorosis"). Each circle represents a study area in which the proportion of people with fluorosis is estimated—the larger the circle, the higher the precision of the estimate

These results show a strong association between water fluoride concentration and the proportion of the population with dental fluorosis. We estimate that six people (95% confidence interval 4 to 21) would have to be exposed to water fluoride concentrations of 1.0 ppm for one additional person to develop fluorosis of any degree, compared with a theoretical low fluoride concentration of 0.4 ppm. Of these, about one quarter will have fluorosis of aesthetic concern (number needed to treat 22, 95% confidence interval 13.6 to  $\infty$ ). These estimates apply only to the comparison of 1.0 ppm with 0.4 ppm. The model may not fit data at the extreme ends (low or high concentrations) well because of the small numbers of data points at these concentrations. Though many areas in Britain may have water fluoride concentrations lower than 0.4 ppm, this concentration was chosen as the comparator (low fluoride) to ensure that the results were as reliable as possible.



**Fig 5** Incidence of bone fracture (estimate of effect and 95% confidence interval). See the *BMJ*'s website for further details

#### Bone fracture and problems with bone development

Twenty nine studies were included on the association with bone fracture or problems with bone development and water fluoride. These studies had a mean (range) validity score of 3.4 (1.5-6.0) out of 8. All but one study was evidence level C (the other being level B).

Figure 5 shows the estimate of effect of water fluoridation compared with control for studies that provided sufficient information.<sup>w11-30</sup> The estimates are distributed evenly around the line of no effect (1.0). There were four analyses that indicated a significant increase in risk of fracture<sup>w12 w13 w20 w21</sup> and five that indicated a significant decrease in risk at the 5% significance level.<sup>w14 w16 w19 w24 w28</sup> Significant heterogeneity was found ( $P < 0.001$ ) among studies. There were no definite patterns of association for fractures of the hip or "other sites" taken as a group.

Meta-regression showed that the only variable associated with the summary measure was duration of study, with studies that were 10 years or longer in duration associated with a protective effect of water fluoridation (fewer fractures).

#### Cancer studies

We included 26 of the association of water fluoridation and cancer. Eighteen of these studies were of evidence level C and eight of level B. The mean (range) validity score was 3.8 (2.8-4.8). Incidence of all cause cancer and mortality was considered as an outcome in 10 studies, and 22 analyses were made.<sup>w31-40</sup> Of these, only two studies found a significant association: one found a negative association (more cancers) in one of eight subgroups,<sup>w32</sup> the other found a significant positive effect (fewer cancers).<sup>w31</sup> Of nine studies comprising 20 analyses of bone cancers,<sup>w41-49</sup> one found a significant negative effect in both men and boys (more cancers).<sup>w41</sup> Because of the varying outcome measures we could not formally pool results.

#### Other possible adverse effects

We included 32 studies of the association of water fluoridation with other possible negative effects. These studies examined various different outcomes, including Down's syndrome, mortality, senile dementia, goitre, and IQ. The quality of these studies was low; all studies were of evidence level C, and the average validity checklist score was 2.7 (range 1.5-4.5) out of 8. None of the studies had a prospective follow up or incorporated any form of blinding. While 22 studies mentioned potential confounding factors, only six used an analysis that controlled for them.

Three of the 33 studies found significant effects. One found a significant negative effect of water fluoride on Alzheimer's disease (increased incidence) and a significant positive effect on impaired mental functioning (decreased incidence).<sup>w49</sup> The other found a significant positive association with congenital malformations in one of two sets of data.<sup>w50</sup> A third study found that the combination of low iodine and high fluoride concentrations was associated with goitre and learning difficulties.<sup>w53</sup> Because of the varying outcome measures we could not formally pool results.

#### Discussion

The most serious defect of the studies of possible beneficial effects of water fluoridation was the lack of appropriate design and analysis. Many studies did not present an analysis at all, while others did not attempt to control for potentially confounding factors. Age, sex, social class, ethnicity, country, tooth type (primary or permanent), mean daily regional temperature, use of fluoride, total fluoride consumption, method of measurement (clinical exam or radiographs, or both), and training of examiners are all possible confounding factors in the assessment of development of dental caries.

While some of these studies were conducted in the 1940s and 50s, before the common use of such analyses, later studies also failed to use methods that were then commonplace. Many studies lacked any measure of variance for the estimates of caries presented. While most of the studies evaluating the proportion of



children without caries contained sufficient data to calculate standard errors, only four of the eight studies that reported decayed, missing, and filled primary/permanent teeth provided any estimate of variance.

### Outcomes measured and bias

The outcome of fluorosis was the most studied of all the adverse effects considered. Observer bias may be of particular importance in studies that assess fluorosis. Because assessment is subjective, unless the observer is blinded to the exposure status of the person being evaluated, bias can be introduced. Efforts to reduce potential observer bias were rarely undertaken in the included studies. The prevalence of fluorosis is overestimated by the indices used in the included studies because enamel opacities not caused by fluoride may be included. The degree to which the estimated 48% prevalence of fluorosis at a water fluoride concentration of 1 ppm overestimates the true prevalence is unknown. Figures 3 and 4 do not originate at 0% fluorosis because all areas included in the studies had at least a small amount of fluoride in the water. In addition, the effects of fluoride from other sources may also be playing a part.

Many studies of other potential negative effects also did not take steps to reduce bias or use analytic techniques to control for potential confounding factors. Interpretation of the results of these studies is difficult because few met inclusion criteria on each specific outcome and studies were generally of poor quality.

Statistical heterogeneity among studies may explain why individual studies report differing estimates of effect. Significant heterogeneity was found among studies of caries, fluorosis, and bone fracture and was also apparent among studies of cancer and other negative effects but could not be tested for. In addition, methodological and clinical diversity was present among these studies.

Publication bias is defined as the failure to publish research on the basis of the nature and directional significance of the results. Because of this, systematic reviews that fail to include unpublished studies may overestimate the true effect of an intervention. Because of the nature of the outcomes and study designs that we examined in this review we considered that the standard methods developed to investigate publication bias were not practical or appropriate. It is thus difficult to estimate whether publication bias is having an effect. As we took such a broad approach in searching for studies, any missed studies would have to be large and different from those that were included to overturn the overall result.

### Conclusions

Given the level of interest surrounding the issue of public water fluoridation, it is surprising to find that little high quality research has been undertaken. As such, this review should provide both researchers and commissioners of research with an overview of the methodological limitations of previous research.

The evidence of a reduction in caries should be considered together with the increased prevalence of dental fluorosis. No clear evidence of other potential negative effects was found. This evidence on positive and negative effects needs to be considered along with the ethical, environmental, ecological, financial, and

### What is already known on this topic

Dental caries cause morbidity and suffering and incur costs

Artificial water fluoridation has been used as a community intervention to reduce the prevalence of dental caries for decades in some communities, but its use remains controversial

### What this study adds

A systematic review of water fluoridation reveals that the quality of the evidence is low

Overall, reductions in the incidence of caries were found, but they were smaller than previously reported

The prevalence of fluorosis (mottled teeth) is highly associated with the concentration of fluoride in drinking water

An association of water fluoride with other adverse effects was not found

legal issues that surround any decisions about water fluoridation. Any future research into the safety and efficacy of water fluoridation should be carried out with appropriate methodology to improve the quality of the existing evidence base.

We thank Dr Keith Abrams, University of Leicester, for contributions to the analysis; Vanda Castle, NHS Centre for Reviews and Dissemination, University of York, for secretarial support; Dr Alan Glanz, Department of Health, for coordination and organisation with the Department of Health; and Marijke van Gestel, University of Maastricht, for technical assistance early in the review process. Details of the members of the advisory panel can be found on the *BMJ's* website.

Contributors: All authors contributed to the design of the protocol, execution of the review and content of the paper. JK led the project and provided methodological skill to the review. MSM was lead reviewer. KM designed and implemented the electronic search strategies and assisted in locating authors. PFW, JC, MM, and MB pilot tested data extraction forms, screened studies, and extracted data. ET and PFW assessed study validity. RT and IC provided clinical interpretation of included dental trials and terminology. PFW and AJS conducted analysis of results. PMW contributed to the interpretation of the results. The advisory panel provided peer review and advice regarding the protocol, analysis, and interpretation. MSM, JK, PFW, and ET are guarantors of the paper.

Funding: This review was commissioned and funded by the Department of Health. The views expressed in this review are those of the authors and not necessarily those of the Department of Health.

Competing interests: None declared.

- 1 Secretary of State for Health. *Saving lives: our healthier nation*. London: Stationery Office, 1999.
- 2 NHS CRD. *A systematic review of public water fluoridation*. York: NHS Centre for Reviews and Dissemination, University of York, 2000. (Report 18.)
- 3 NHS CRD. *Undertaking systematic reviews of research on effectiveness*. York: NHS Centre for Reviews and Dissemination, University of York, 1996. (Report 4.)
- 4 Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* 1999;18:2693-708.
- 5 Dersimonian R, Laird N. Meta-analysis in clinical trials. *Cont Clin Trials* 1986;7:177-88.
- 6 Sharp S. Meta-analysis regression: statistics, biostatistics, and epidemiology 23 (sbe23). *Stata Tech Bull* 1998;42:16-22.

(Accepted 12 September 2000)