The effectiveness of workplace smoking cessation programmes: a meta-analysis of recent studies

G Smedslund, K J Fisher, S M Boles, E Lichtenstein

Objective: Using meta-analytic procedures, we compare the effectiveness of recent controlled trials of worksite smoking cessation during the 1990s with a previous meta-analysis of programmes conducted in the 1980s.


Study selection: Controlled smoking cessation interventions at the workplace with at least six months follow up published from 1989 to 2001 and reporting quit rates (QRs).

Data extraction: Two reviewers independently scanned titles/abstracts of relevant reports, and we reached consensus regarding inclusion/exclusion of the full text reports by negotiation. A third reviewer resolved disagreements. Two reviewers extracted data according to a coding manual. Consensus was again reached through negotiation and the use of a third reviewer.

Data synthesis: 19 journal articles were found reporting studies conforming to the study’s inclusion criteria. Interventions included self help manuals, physician advice, health education, cessation groups, incentives, and competitions. A total of 4960 control subjects were compared with 4618 intervention subjects. The adjusted random effects odds ratio was 2.03 (95% confidence interval 1.42 to 2.90) at six months follow up, 1.56 (95% CI 1.17 to 2.07) at 12 months, and 1.33 (95% CI 0.95 to 1.87) at more than 12 months follow up. Funnel plots were consistent with strong publication bias at the first two follow ups but not the third. In Fisher et al’s 1990 study, the corresponding ORs were 1.18, 1.66, and 1.18.

Conclusions: Smoking cessation interventions at the worksite showed initial effectiveness, but the effect seemed to decrease over time and was not present beyond 12 months. Compared to the Fisher (1990) analysis, the effectiveness was higher for the six month follow up. Disappointingly, we found methodological inadequacies and insufficient reporting of key variables that were similar to those found in the earlier meta-analysis. This prevented us from determining much about the most effective components of interventions. It is advisable for researchers conducting studies in the future to report on attrition and retention rates of participants who quit, because these variables can affect QRs.

Meta-analysis as a data integration strategy

A meta-analysis is a quantitative synthesis of the results of many different studies of the same genre. Core advantages of meta-analysis are its ability to detect trends in a set of studies that may not be obvious to the naked eye, and to provide a quantitative estimate of the magnitude of effect. The outcome of a meta-analysis is known as an “effect magnitude”. The three most common types of effect magnitude are: (1) the standardised mean difference, called “effect size” or “d”; (2) the correlation coefficient r; and (3) the odds ratio (OR). In the present meta-analysis, we used the ORs in the experimental and control conditions to calculate ORs.
results across studies; and (5) failure to examine the influence of moderating variables on study outcomes. Additionally, meta-analysis itself is subject to several criticisms, which fall into four broad categories:

1. The “apples and oranges” problem. Diverse measuring instruments, different subjects, various treatments and outcomes, and other differences make meaningful comparisons illogical.
2. The use of data from “poor” studies. Meta-analysis procedures often uncritically accept poorly designed or low quality studies into the analysis, rendering aggregated solutions mediocre or uninterpretable.
3. Selection bias in favour of reported research. Published research is biased in favour of significant findings; thus, meta-analytic results are also biased.
4. Non-independent data. Multiple comparisons from large data sets that contribute to meta-analyses make the results seem more reliable than they really are.

In this study, these four issues were addressed in the following ways. The “apples and oranges” problem was dealt with empirically by coding possible moderator variables in each study according to predetermined criteria. This made it possible to investigate the degree to which these variables accounted for variance in the meta-analytic results. For smoking cessation studies, the dependent measure of QR is less problematic in terms of measurement criteria than in many other disciplines because it is a widely used criterion measure of outcome. Potential moderator variables that influence QR can be generated from published reviews of worksite smoking cessation and by examination of recent high quality studies.

The second criticism was addressed by adopting Slavin’s16 20 “best evidence” synthesis, which gives priority to studies highest in internal and external validity based upon defended, well specified inclusion criteria developed after a thorough examination of the literature. In the present study, we judged that the best evidence came from studies with a control or comparison group and having a follow up time of at least six months. Although these two criteria do not guard against inclusion of low quality studies, they lower the probability of bias due to self selection and transitory effects of cessation programmes. Coding procedures were developed to accommodate possible variations in methodology across these best evidence studies.

The problem of potential bias in published studies was addressed by producing funnel plots. This is a graphical method for detecting the relative absence of small, imprecise studies with negative results.

In addressing the fourth criticism relating to the use of a large number of results from the same study, we selected only one comparison for each study. In the studies with more than one treatment group, we consistently compared the simplest treatment (or a no treatment control group) versus the next simplest treatment. Typically, the other groups received nested interventions with which there were problems with disentangling the active ingredients of the treatment. We ran separate analyses for six month follow up, 12 month follow up, and more than 12 month follow up. Each study could contribute a maximum of three effect magnitudes (three follow up time points), but only one effect size per study was allowed into any single analysis.

METHODS
Inclusion criteria
A thorough search of all published and unpublished studies on smoking cessation from January 1989 through December 2000 was conducted. For potential inclusion in our meta-analysis, studies had to conform to the following criteria: (1) a study of worksite smoking cessation that (2) reported QRs for a follow up of at least six months post-treatment, (3) included a control or comparison group, and (4) published between January 1989 and December 2000.

Search strategies
Electronic database searches were conducted, and review articles were consulted. A manual search of reference lists from retrieved publications was also conducted.

Search criteria—“Smoking cessation” was used as a subject descriptor with “worksite” or “work site” or “workplace.” Studies appearing in publications from the years 1989 to 2000 were included.

Databases searched—We searched the following databases for studies that met the prescribed criteria: ABI/Inform, BRS, Combined Health Information Database (CHID), Dissertation Abstracts International Database, ERI, Medline, Occupational Health and Safety Database, PsycInfo, Smoking and Health Database, Social Sciences Citation Index, and Sociological Abstracts Database.

Selecting papers for coding
Titles and abstracts provided the basis for initial decisions and selection of documents. Two raters independently scanned the set of documents for inclusion/exclusion criteria met by the titles and/or abstracts. The two raters then met and on the basis of relevance contained in the title or abstracts, agreed on a group of 109 full text reports that would be considered for coding. Each rater then made an independent decision on whether a paper met the inclusion criteria. Initially, 25 published articles were judged to meet the coding criteria, but after further consideration, six papers were excluded. One paper was excluded because it reported data that were already reported in one of the included papers. Two other papers employed workplace randomised designs and were excluded because they reported outcomes in the form of workplace prevalence rates and not individual level QRs. We elected to omit these from the meta-analysis but report their outcomes. Hence, a final set of 19 papers was coded using a coding manual adapted from Lipsey and Wilson. Inter-rater reliability was computed on the set of 25 studies.

Coding variables
Design variables—The coding variables based on study design included random or non-random assignment, unit of randomisation (subject or worksite), unit of analysis (subject or worksite), attrition, lost at follow up, biochemical verification, length of follow up, and pre-test group comparison.

Sample descriptors—The coding descriptors included age, race, sex, education, smoking rate (cigarettes/day), smoking prevalence, and previous quit attempts.

Organisation variables—These variables included company size, type of industry, private or public workplace setting, and stringency of non-smoking policy. Whenever there were multiple sites in a study, we calculated an average company size or “typical” size (using the mode statistic).

Intervention descriptors—These descriptors included intervention type (self help, physician advice, incentives, cessation group, steering committee, and other), treatment duration (fixed time or intermittent), whether the intervention took place on company time, and whether the intervention was part of a larger programme.

Effect size data—These variables included control group n; treatment group n; QR for six months, 12 months, and/or more than 12 months; QR type (seven day point prevalence, 30 day point prevalence, or continuous abstinence).
Excluded randomised trials
Two large trials were highly relevant to the present analysis but were not included in the meta-analysis because they estimated QRs indirectly from prevalence changes. Moskowitz et al\textsuperscript{27} estimated the effects of local workplace smoking laws in California. They found that smokers who worked in localities with a strong workplace ordinance (compared with no workplace ordinance) were more likely to report quitting smoking in the prior six months (OR 1.5, 95% confidence interval (CI) 1.1 to 1.7).

Sørensen et al\textsuperscript{27} presented results of the Working Well Trial, the largest worksite cancer prevention and control trial to date in the USA. The Working Well Trial used a randomised, matched pair evaluation design, with the worksite as the unit of assignment and analysis. The study was conducted in 111 worksites (n = 28,000 workers). The effects of the intervention were evaluated by comparing changes in intervention and control worksites, as measured in cross sectional surveys at baseline and follow up. The two year intervention targeted both individuals and the worksite environment. Changes in tobacco use were in the desired direction but were not significant.

ANALYSIS
The 19 controlled trials in this review may be considered a “best evidence”\textsuperscript{16} subset of the literature on the topic and, as such, are well suited to meta-analytic procedures. Each of the studies was cast into a short term (six month QR), intermediate term (12 month QR), or long term (more than 12 months QR) frequency table based on the count of smokers and quitters in treatment and control conditions at each assessment point. The number of quitters and the total number of participants was recorded in Review Manager 4.2\textsuperscript{,}\textsuperscript{27} which is the Cochrane Collaboration’s software for meta-analysis. A random effects model was used, assuming that the effect magnitudes come from a distribution of true effect magnitudes and not a single true magnitude. The effect magnitudes were weighted by the standard error of their natural logarithms, and a 95% CI was computed.

RESULTS
The studies in the present analysis are mutually exclusive of the studies in Fisher et al’s\textsuperscript{25} analysis. Table 1 offers an overview of the setting, design, and intervention of the included studies.\textsuperscript{27} 28 30 32–44 46–48 The sample size in the 19 included studies ranged from 56 to 2,317, with a median of 426.5. Nine studies were randomised trials. All but three\textsuperscript{27} 30 45 reported follow up times of at least 12 months. As table 1 shows, the papers report a high degree of diversity regarding intervention types, settings, and participants. Types of interventions coded were bibliotherapy (that is, self help written materials) (16 studies), cessation group (13 studies), “other type” (12 studies), incentives (nine studies), steering committee (seven studies), non-smoking policy (four studies), physician advice (two studies), and pharmacological treatment (two studies).

Inter-rater reliability
Initially, we coded the 25 studies that made the final cut in terms of satisfying the inclusion criteria as a pilot practice. The mean inter-rater reliability for this pilot coding was 72%. Any disagreements related to the pilot coding procedures and the decision to reduce the final set of studies to 19 was resolved through a third reviewer (EL). For the final set of 19 studies the two raters (GS and KJF) achieved 100% agreement.

Results and reporting of moderator variables
All included studies were published articles in peer reviewed journals. Some dissertations qualified for inclusion initially, but were later excluded because the results appeared in a subsequent published article. Some studies were conducted in a number of different sites. Company size was coded by using the average number of workers at each site. One study had fewer than 100 employees per site, three had between 100–249 employees, three had between 250–750 employees, and 11 had more than 750 employees at the typical site. One study did not report data on number of employees. Six worksites were coded as “manufacturing/construction”, two were coded as “utilities/transport/communication”, three were coded as “service/educational/hospital”, six were coded as “mixed”, one was coded as “other”, and one had missing data on this variable. Seven worksites were private, five were public, five were mixed, and two had missing data. Two worksites had no smoking policy, two allowed smoking in separate areas, none had a total ban, and the remaining 12 did not report stringency of non-smoking policy. Four studies reported that smoking cessation was part of a larger programme to improve health. Small sample sizes prevented us from conducting meaningful sub-analyses of effects of most moderator variables.

Among methodological characteristics, attrition during the intervention was reported in only six of the 19 studies and participants lost at follow up were reported in only seven studies. Twelve studies reported some kind of biochemical validation of smoking abstinence, while seven studies did not. Thirteen studies had interventions that lasted a fixed time. In six studies, the interventions were intermittent, meaning that cessation programmes or activities were implemented at certain times rather than being regularly accessible.

Quit rates
Table 2 shows quit rates for each study, group and follow up. For example, the first study (Bertera 1990) reported only data for more than 12 months, while the Jason 1997 study reported data on all three follow up points. The variation in the quit rates is tremendous in the intervention groups for all follow ups. The variation is also large in the control groups.

Table 3 provides data on characteristics and reporting on some of the other moderator variables. Most studies have been conducted in workplaces with high smoking prevalence and with white, heavy smokers. Possible important moderator variables, such as previous quit attempts, cigarettes per day, and education level, were inconsistently reported.

Odds ratios and quit rates
For each study, we recorded data on up to three independent treatment conditions and at three time points (42 effects sizes), but since only 11 effect sizes came from a second or third treatment condition, we elected to focus mainly on one treatment condition for each study (the two simplest conditions). Hence, most of the reported analyses were restricted to 28 QRs (table 2). The weighted ORs were 2.03 (95% CI 1.42 to 2.90), 1.56 (95% CI 1.17 to 2.07), and 1.33 (95% CI 0.95 to 1.87) for the three follow up points, respectively (figs 1–3). At six months the overall 16.7% in the intervention groups versus 8.5% in the control groups, at 12 months it was 20.8% versus 12.2%, and at more than 12 months it was 17.2% versus 13.9%.

Randomised and non-randomised studies
Figure 1 shows a forest plot of the six month follow up. The studies are divided into randomised and non-randomised. The non-randomised show a much stronger effect (OR 4.65) than the randomised ones (OR 1.74). The precision of the
estimates is greater for the randomized studies (95% CI 1.26 to 2.40) than for the non-randomised ones (95% CI 1.92 to 11.28). The χ² test for homogeneity confirmed that both types of studies show homogeneity at the six month follow up. The funnel plot shows strong evidence of publication bias.

Figure 2 shows similar results for the 12 month follow up. The effects have decreased for both types of study and the precision is greater for the randomised studies. Both subgroups are homogenous and there is strong evidence for publication bias.

Figure 3 shows that the effects do not seem to last beyond 12 months as evidenced by confidence intervals for the total group and both subgroups surrounding 1. The set of randomised studies is homogenous, but the non-randomised ones are heterogeneous. There is no evidence for publication bias, except for the small Bertera 1990 study.

**Comparison of the 1990 and 2001 surveys**
The overall ORs found in the present study (2.03, 1.56, 1.33 at 6, 12, and >12 months) is quite similar to the corresponding ORs of 1.18, 1.66, and 1.18 found in the Fisher et al study. In 1990, 34 comparisons produced an overall QR of 13% in the intervention groups. In the 2001 analysis, 28 comparisons produced a QR of around 18%.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Setting and comparison tested</th>
<th>Company size</th>
<th>Random allocation</th>
<th>Follow up (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bertera 1990</td>
<td>70</td>
<td>Stop smoking clinic versus self-help kit. ALA “Freedom From Smoking”</td>
<td>&gt;750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Burling 1989</td>
<td>58</td>
<td>VA Medical Center employees. Computerised nicotine fading versus contest only</td>
<td>NR</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Erfurt 1991</td>
<td>671</td>
<td>Manufacturing plants. Health education, counselling, social organisation.</td>
<td>&gt;750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Glasgow 1993</td>
<td>512</td>
<td>Stop smoking contest. Incentive versus no incentive</td>
<td>250–750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Glasgow 1995</td>
<td>431</td>
<td>Take Heart Study. 26 worksites in Oregon. Steering committee and menu approach. Early intervention versus delayed intervention</td>
<td>100–249</td>
<td>Yes</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Gomel 1997</td>
<td>431</td>
<td>28 ambulance stations in Sydney, Australia. Counselling. Risk factor education versus health risk assessment</td>
<td>&lt;100</td>
<td>Yes</td>
<td>6, 12</td>
</tr>
<tr>
<td>Helyer 1998</td>
<td>104</td>
<td>Air Force smoking cessation groups versus no treatment</td>
<td>&lt;100</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Hynowitz 1991</td>
<td>252</td>
<td>Smoking cessation, health education, worksite non-smoking policy. Group + enriched milieu versus group only</td>
<td>&gt;750</td>
<td>Yes</td>
<td>12</td>
</tr>
<tr>
<td>Jason 1997</td>
<td>844</td>
<td>Worksite television program, newspaper, support groups. Self help manual + incentives versus manual only</td>
<td>100-249</td>
<td>Yes</td>
<td>6, 12, &gt;12</td>
</tr>
<tr>
<td>Koffman 1998</td>
<td>177</td>
<td>Aerospace industry. Multicomponent programme (self help, telephone, incentives) versus standard smoking cessation</td>
<td>&gt;750</td>
<td>No</td>
<td>6, 12</td>
</tr>
<tr>
<td>Lang 2000</td>
<td>1269</td>
<td>Electrical and gas company. Contract, quit date + follow up support by worksite physician versus simple advice to quit</td>
<td>100-249</td>
<td>Yes</td>
<td>6, 12</td>
</tr>
<tr>
<td>Muto 1998</td>
<td>56</td>
<td>Nicotine gum. Group treatment + competition versus group only</td>
<td>&gt;750</td>
<td>No</td>
<td>6</td>
</tr>
<tr>
<td>Olsen 1991</td>
<td>2317</td>
<td>Dow Chemical Texas Operations. Smoking Cessation Incentive Program (SCIP) versus control</td>
<td>&gt;750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Salina 1994</td>
<td>419</td>
<td>38 Chicago companies. Self help manuals, television, quitting techniques, social support versus manuals and television only</td>
<td>100-249</td>
<td>Yes</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Sorensen 1993</td>
<td>681</td>
<td>Consultation. Cessation classes versus control</td>
<td>&gt;750</td>
<td>Yes</td>
<td>6, 12, &gt;12</td>
</tr>
<tr>
<td>Willemsen 1995</td>
<td>394</td>
<td>Dutch telecommunications company. Tailored letter of advice, self help guide, group cessation, telephone quit line versus control</td>
<td>&gt;750</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Willemsen 1999</td>
<td>498</td>
<td>Enriched environment versus control</td>
<td>250-750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Willemsen 1998</td>
<td>438</td>
<td>Self help manual, group, mass media, smoking policies versus self help manual only</td>
<td>&gt;750</td>
<td>No</td>
<td>&gt;12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Quit rates at different follow up points for the selected conditions (n = 19 studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>6 month quit rate (%)</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
</tr>
<tr>
<td>Bertera 1990</td>
<td>21.4</td>
</tr>
<tr>
<td>Burling 1989</td>
<td>10.0</td>
</tr>
<tr>
<td>Erfurt 1991</td>
<td>30.8</td>
</tr>
<tr>
<td>Glasgow 1993</td>
<td>10.3</td>
</tr>
<tr>
<td>Glasgow 1995</td>
<td>23.0</td>
</tr>
<tr>
<td>Gomel 1997</td>
<td>38.9</td>
</tr>
<tr>
<td>Helyer 1998</td>
<td>12.0</td>
</tr>
<tr>
<td>Hynowitz 1991</td>
<td>7.5</td>
</tr>
<tr>
<td>Jason 1997</td>
<td>6.1</td>
</tr>
<tr>
<td>Koffman 1998</td>
<td>3.0</td>
</tr>
<tr>
<td>Lang 2000</td>
<td>25.0</td>
</tr>
<tr>
<td>Muto 1998</td>
<td>22.9</td>
</tr>
<tr>
<td>Olsen 1991</td>
<td>23.0</td>
</tr>
<tr>
<td>Salina 1994</td>
<td>12.0</td>
</tr>
<tr>
<td>Sorensen 1993</td>
<td>7.5</td>
</tr>
<tr>
<td>Willemsen 1995</td>
<td>6.1</td>
</tr>
<tr>
<td>Willemsen 1998</td>
<td>38.9</td>
</tr>
<tr>
<td>Unweighted mean</td>
<td>16.7</td>
</tr>
<tr>
<td>Range</td>
<td>6.1–30.8</td>
</tr>
</tbody>
</table>
DISCUSSION

The retrieval of 19 controlled studies of worksite smoking cessation interventions from the decade covering the 1990s was a similar yield to a previous meta-analysis of 20 studies covering the decade from the 1980s, and similar results were obtained (19 studies with 28 observations in 2001 versus 20 studies with 34 observations in 1990). The studies in the 1990s were almost twice as effective at the six months follow up as the studies in the 1980s, but this difference did not last at longer follow up points. Because we used the conservative strategy of consistently comparing the control group with the least comprehensive intervention arm, the effects may be somewhat underestimated—but at least this made the results more interpretable.

Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean/%</th>
<th>Number missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>39.8</td>
<td>2</td>
</tr>
<tr>
<td>Race</td>
<td>12 white/1 mixed/1 other</td>
<td>5</td>
</tr>
<tr>
<td>Males</td>
<td>63.8%</td>
<td>1</td>
</tr>
<tr>
<td>Education</td>
<td>8 with some college/2 with no college</td>
<td>9</td>
</tr>
<tr>
<td>Cigarettes per day</td>
<td>22.4</td>
<td>6</td>
</tr>
<tr>
<td>Smoking prevalence</td>
<td>29.3%</td>
<td>5</td>
</tr>
<tr>
<td>Previous quit attempts</td>
<td>3.0</td>
<td>18</td>
</tr>
<tr>
<td>Randomisation</td>
<td>10 random/9 not random</td>
<td>0</td>
</tr>
</tbody>
</table>

Review: Smoking cessation in the workplace
Comparison: 6 month follow up
Outcome: Randomised versus non randomised studies

Figure 1 Upper panel: six months follow up data stratified into randomised and non-randomised studies. Lower panel: funnel plot of the odds ratios versus the standard error of the logged odds ratio.
Because we included both randomised and non-randomised studies, we were able to assess the effect of randomisation on the dependent variable. At all three follow-up points, the non-randomised studies showed larger effects. This is consistent with the earlier meta-analysis, where quasi-designed studies had the largest effect sizes.10 The randomised results are probably closer to the truth, as the non-randomised studies are probably overestimating the effects. Because of the probable publication bias, we believe that the randomised studies are also overestimating the effects to some degree.

We collected a number of possible moderator variables for this study. However, because of the low number of observations, we did not find it feasible to do any meaningful subgroup analyses. The effects based on moderator variables were examined on three time points and on two methodological subgroups (randomised/non-randomised).

In the period covered by the present analysis, nicotine replacement therapies and other pharmacological treatments became increasingly popular. It is surprising that this trend did not show up more often as a treatment condition in studies. One reason might be that most pharmacological treatments for smoking cessation (for example, patches, gum) did not become over-the-counter medications until the late 1990s. Another reason might be that researchers excluded employers who confessed to using such medications.

Why did the results not improve even more over the decade of the 1990s? Not only did many new medications become more widely available and affordable, but there were also more restrictions on worksite smoking, and smokers were generally more marginalised in the work place as well as in society. These factors possibly contribute to the proportion of committed, “hard core” smokers who may be less motivated to quit and more likely to be nicotine dependent.
Smoking cessation is influenced not only by the cessation programmes but also by the workplace settings and organisational context. These broader features and influences are also important to consider in relation to interpreting the results.

A possible source of bias in worksite intervention rises when the unit of randomisation differs from the unit of analysis. Eight studies31–33 35 36 38 42 43 used the worksite as the unit of randomisation. In all but one31 of these studies, the analysis took into account the amount of intracluster correlation. Failing to do so may inflate the observed differences between groups.

We acknowledge some limitations to our study. First, this analysis included only studies published in English. From our electronic searches, however, there were few non-English research publications on worksite smoking cessation. This may have been because they were either unpublished or were published in journals not indexed in the databases we searched.

Although many studies of workplace smoking cessation were published during the 1990s (our initial pool of references cited 114 reports), a sizeable proportion had similar methodological or design inadequacies that were identified in Fisher et al’s previous meta-analysis. Although the autonomy, resource limitations, and independence of researchers are respected, unless there is some conformity and consistency in reporting key variables, including design, type of industry, and key employee characteristics, generalisability of findings will be severely limited. Since the publication of the 1990 meta-analysis, the CONSORT statement (consolidated standards of reporting trials) has become available. The CONSORT contains a 21 item checklist and a flow diagram. The format provides readers with consistency from report to report as to where they can expect to find relevant information. It would be highly valuable if future authors use the CONSORT checklist when preparing primary reports.

**Figure 3**

Upper panel: More than 12 months follow up data stratified into randomised and non-randomised studies. Lower panel: funnel plot of the odds ratios versus the standard error of the logged odds ratio.

Smoking cessation is influenced not only by the cessation programmes but also by the workplace settings and organisational context. These broader features and influences are also important to consider in relation to interpreting the results.

A possible source of bias in worksite intervention rises when the unit of randomisation differs from the unit of analysis. Eight studies29–33 35 36 38 used the worksite as the unit of randomisation. In all but one of these studies, the analysis took into account the amount of intracluster correlation. Failing to do so may inflate the observed differences between groups.

We acknowledge some limitations to our study. First, this analysis included only studies published in English. From our electronic searches, however, there were few non-English research publications on worksite smoking cessation. This may have been because they were either unpublished or were published in journals not indexed in the databases we searched.
ACKNOWLEDGEMENTS

We thank Elizabeth Mondulick and Ron Renchler from Oregon Research Institute for helping in the preparation of this manuscript, and the anonymous reviewers who made valuable suggestions during the review process.

Authors’ affiliations

G Smedslund, Directorate for Health and Social Affairs, Oslo, Norway
K J Fisher, S M Boles, E Lichtenstein, Oregon Research Institute, Eugene, Oregon, USA

REFERENCES

Studies included in the meta-analysis marked by *. Excluded studies marked by †.