

# Application of meta-analysis in reviewing occupational cohort studies

Otto Wong, Gerhard K Raabe

DEC 18 1996

## Abstract

Meta-analysis has been used increasingly in renewing and summarising epidemiological studies. Reviews incorporating meta-analyses have appeared in medical journals in increasing numbers. Although there are several methodology papers on meta-analysis, most of these papers have been written primarily for discussion among epidemiologists. The present paper considers some of the basic methodological issues, the more practical aspects of meta-analysis, and targets an audience of mainly non-epidemiologists. Thus, the main objective of this paper is to provide some basic guidelines for non-epidemiologists to evaluate meta-analysis in occupational cohort studies. In this methodology paper, the limitations and problems of traditional qualitative reviews are pointed out. Some of these problems can be dealt with by quantitative meta-analysis. The potential limitations and benefits of quantitative meta-analysis are discussed. Rather than replacing traditional qualitative review, quantitative meta-analysis should be made part of the overall assessment. The term "meta-review" is proposed to emphasise the importance of both qualitative and quantitative components in a comprehensive review process. The basic steps in a meta-review are outlined, with a discussion on how to recognise and avoid some of the problems which are likely to occur at each step. A meta-review is useful in selecting studies, and in organising, presenting, and summarising results from individual studies. A meta-review can also be used to detect heterogeneity among studies. Major benefits of conducting a meta-analysis (the quantitative component in a meta-review) include the increase in statistical power and the estimate of a properly weighted summary risk estimate.

(*Occup Environ Med* 1996;53:793-800)

**Keywords:** literature review, epidemiology, occupational cohort studies, meta-analysis, meta-review

Meta-analysis has been used increasingly in epidemiological literature reviews. Its use has

generated both enthusiasm and scepticism. Meta-analysis, when used properly, is a powerful research tool. There are several papers in the epidemiological literature on different aspects of meta-analysis. However, these have been written primarily for discussion among epidemiologists. The present paper considers some of the basic methodological issues and the more practical aspects of meta-analysis, and targets as its audience mainly non-epidemiologists, such as occupational physicians, occupational nurses, industrial hygienists, toxicologists, risk managers, or others who have the need and the interest to read epidemiological literature. Thus, a primary objective of this paper is to provide some basic guidelines for non-epidemiologists to evaluate papers on meta-analysis in occupational epidemiology. As such, highly statistical or technical issues will not be discussed. Furthermore, the discussion will focus on occupational cohort studies.

The term meta-analysis was coined by the British psychologist Glass in 1976 to describe "statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings."<sup>1</sup> Although the term meta-analysis was not used before that time, statistical integration of results from studies began long before then. In particular, several excellent papers describing statistical techniques for combining data from individual experiments had been written by pioneers such as Cochran, Fisher, and Pearson.<sup>2-5</sup> Meta-analysis has been used extensively in psychology, sociology and education, and more recently, in clinical trials. In epidemiology meta-analysis has not been widely used until recently. Although the term meta-analysis was not used, in the 1964 Surgeon General's report *Smoking and health*, data from seven cohort studies available at that time were combined to provide summary estimates of cancer risk for smokers.<sup>6</sup>

Various definitions of meta-analysis have been published. For example, similar to the original definition by Glass, Huque defined meta-analysis as "a statistical analysis which combines or integrates the results of several independent clinical trials, considered by the analyst to be combinable."<sup>7</sup> Without any question, statistical analysis is a major part of meta-analysis. However, as will be discussed below, meta-analysis involves more than the application of statistical techniques. There is defi-

Applied Health Sciences, 181 Second Avenue, Suite 628, PO Box 2078, San Mateo, California 94401, USA  
O Wong

Medical Department, Mobil Business Resources Corporation, PO Box 189, New Hope, Pennsylvania, USA  
G K Raabe

Correspondence to:  
Otto Wong, Applied Health Sciences, 181 Second Avenue, Suite 628, PO Box 2078, San Mateo, California 94401, USA.

Accepted 29 July 1996

nately a qualitative component in the selection of studies to be included as well as in the interpretation of the results.

### Meta-review of occupational epidemiology

In a traditional qualitative literature review, the selection of studies is usually not subject to predetermined criteria. This was particularly true in the days before the availability of computerised databases. Furthermore, weights assigned to individual studies were usually subjective. In traditional reviews, no quantitative summaries of risk estimates are provided. Therefore they are primarily narrative, and tend to be subjective.

A comprehensive literature review should consist of both a qualitative review and a quantitative analysis. To be distinguishable from a traditional literature review, we propose the use of the term "meta-review" to describe a comprehensive review that includes a meta-analysis. This term seems appropriate as the Greek word "meta" means "after".<sup>8</sup> According to Webster's dictionary, "meta" is "used with the name of a discipline to designate a new but related discipline designed to deal critically with the original

Table 1 shows the steps in conducting a meta-review. Some of the steps are self explanatory and will be discussed only briefly. Within the framework of occupational studies, the research question (step 1) should consider and identify the exposure of interest, the exposed population, and the extent of exposure. On the response (health outcomes) side, the specific health end points and the corresponding diagnostic criteria should be clearly stated.

Step 2, a literature search or identification of studies, calls for a comprehensive search strategy; and step 3, criteria for inclusion, is vital to the overall search strategy. One question which has been debated to some extent is whether unpublished studies should be included. Negative studies—that is, studies showing no increased risk—are less likely to be published for various reasons, the most common one being that the investigators or journal editors do not think that the results are exciting or informative. This is particularly true for negative studies with a small sample size, because such studies individually are not conclusive. Negative studies are also less likely to be published in the absence of positive studies or any controversy of the subject matter. This is usually referred to as the "file drawer" problem or "publication bias." The general consensus is that a study's quality should not be judged according to whether it is unpublished or not.<sup>10</sup> A more troublesome issue is the identification

of *all* unpublished studies. There is no good solution to this problem. If the reviewer is also a primary investigator in that particular field, identifying unpublished studies may not be difficult. Otherwise, the best approach may be inquiring of researchers (identified through published reports) in that field.

There is also the reverse file drawer problem, the repeated publications or overrepresentation of positive studies. There is a tendency for positive studies to be published or presented at conferences more than once. In a meta-review on asbestos and gasmc cancer, it was found that some data sets had been published repeatedly.<sup>1</sup> Papers reporting on different subcohons of the same population may also complicate the process. Furthermore, studies may have been updated periodically, and results from different versions of the same study may have been published. The general guideline for inclusion in a study is that the most recent and the most complete version of the study should be included.<sup>11</sup> Also, any serious discrepancies between past and current versions should be identified and explained. Finally, studies based on individual companies or individual centres may also be part of an industrywide or multicentre study, and caution must be exercised not to enter the same data more than once in the meta-analysis.<sup>12</sup>

A similar issue is the presentation bias, which refers to selective presentation of study results. For example, the results for rare cancers in a cohort study consisting of workers exposed to certain chemicals of interest may not be presented due to small numbers; yet this is exactly the situation where a meta-analysis is most beneficial. Again, contacting the original authors may be the only solution to this problem. This was done in a recent investigation, in which original investigators were contacted to provide unpublished data for analyses of cell type specific leukaemia in a combined cohort of more than 208 000 workers in the petroleum industry.<sup>13</sup> In most of the original studies, cell type specific leukaemia results were not analysed or reported due to small numbers.

Step 4, the traditional qualitative review, is a main component of the overall process. It summarises the basic information from individual studies, such as design, subject selection, exposure patterns, diagnostic criteria, observation period, potential confounding factors, statistical techniques, etc. The quantitative review should also point out the strengths and weaknesses of individual studies. Table 2 provides a suggested checklist for evaluating studies. Most of the items in table 2 are self explanatory and common to any qualitative literature review, and a

Table 1 Steps in conducting a meta-review

- 1 Defining the research question
- 2 literature search
- 3 Criteria for inclusion in the study
- 4 Traditional qualitative review
- 5 Quantitative meta-analysis
- 6 Integrating traditional qualitative review with quantitative meta-analysis
- 7 Applying criteria for causation in interpretation

Table 2 Checklist for evaluating studies

- |                                           |
|-------------------------------------------|
| Study design                              |
| Selection of study subjects               |
| Sample size and statistical power         |
| Statistical analysis                      |
| Comparison population                     |
| Duration and completeness of follow up    |
| Latency                                   |
| Diagnostic information                    |
| Adequacy and specificity of exposure data |
| Concomitant exposures                     |
| Confounding factors                       |

such will be discussed only briefly when they have special relevance in a meta-analysis. The principal objective of this review step is to identify studies of acceptable quality for inclusion in subsequent meta-analyses and to identify factors contributing to heterogeneity between studies, which can then be used to guide further stratification of data (stratified meta-analyses).

Study design is used to group individual studies for meta-analysis. Generally, separate meta-analyses should be performed on studies of different designs. It is not usually advisable to combine studies of different designs in a single meta-analysis unless it can be determined that study design has little or no influence on study characteristics such as quality of data, specificity of exposure, and uniformity of diagnoses. In reality, study design is usually one of the most important determinants of data quality, exposure specificity, and diagnostic criteria. Similarly, studies with very different statistical techniques, different comparison populations, or different diagnostic categories should generally not be lumped into a single analysis. For example, in the literature on occupational epidemiology, some studies may report proportional mortality ratios (PMRs) and others standardised mortality ratios (SMRs). It is generally agreed that SMR studies are far better than PMR studies, and that PMRs usually overestimate cancer risks in industrial populations due to the healthy worker effect. Therefore, SMR and PMR studies should not be combined in a single meta-analysis. Similarly, if one study examines total digestive cancer in general and another stomach cancer in particular, it would be inappropriate to combine these studies as different diagnostic categories are involved.

The selection of study subjects, adequacy and specificity of exposure data, concomitant exposures, and confounding factors should be examined at the individual study level; including studies which are significantly deficient in any of these areas would increase the heterogeneity in the resultant meta-analysis. For example, if different concomitant exposures are present in individual studies, heterogeneity in the combined analysis will result, as this would reflect the influence of different concomitant exposures. Finally, although sample size or statistical power are important Criteria for evaluating a study's quality on an individual basis, it should not be part of the basis for including or excluding a study in a meta-analysis. As will be discussed later, one of the main benefits of a meta-analysis is the pooling of small studies to improve statistical power.

Some authors have suggested the use of "quality scoring" to determine whether a study should be included. This suggestion has met with strong opposition. For example, in a recent commentary quality scoring was characterised as "the most insidious form of subjectivity masquerading as objectivity."<sup>14</sup> We do not recommend the use of quality scoring for the simple reason that it would be impossible to treat different study characteristics (size, exposure information, completeness of follow up, etc) that are related to quality as if they are

of equal importance or that they can be measured by a single score. In any event, quality scoring does not improve objectivity, as any scoring system can be arbitrary and subjective.

Whether a particular study should be included or excluded in a meta-analysis should be based on a set of criteria such as those suggested in the checklist in table 2. Only those studies which meet certain minimum requirements should be considered further for inclusion. An additional assessment has to be made to find whether all such eligible studies can be included.

The first consideration in a quantitative meta-analysis (step 5) is the ability of individual studies to be combined. To some extent, the definition of a study can sometimes be artificial or arbitrary. For example, a study of petroleum refinery workers from an oil company can consist of two or more refineries. An industrywide study obviously consists of more than one company. Data from individual refineries or companies could have been published separately. Thus, the issue of whether or not studies can be combined may already exist at the level of individual studies.

The primary question to ask is: do studies to be included in the meta-analysis have sufficient common ground to consider the scientific issue in question? Table 3 provides a list of items to be considered in evaluating whether studies can be combined. In occupational studies, the question often asked is: is there a relation between exposure to a certain chemical and a specific disease? Therefore, both specificity and patterns of exposures are obviously important determinants in deciding whether it would be meaningful to combine certain studies. As discussed already, separate meta-analyses should be performed for studies of different design due to the heterogeneity in data collection and data quality.

Other study characteristics should also be examined. The fundamental question is "Are studies so different that they should not be combined?" Studies should be carefully examined for similarity or differences of characteristics such as eligibility of study subjects, cohort definition, exposure definition and classification, disease diagnostic criteria, data collection, etc. However, because many statistical tests are available to evaluate heterogeneity of study results or effect estimates, in many meta-analysis papers the emphasis on assessment of heterogeneity has been incorrectly placed on the results of studies rather than on the basic characteristics of studies. In other

Table 3 Factors determining whether studies can be combined in a meta-analysis

|                                                                 |
|-----------------------------------------------------------------|
| Study design                                                    |
| Specificity of exposures                                        |
| Patterns of exposures                                           |
| Concomitant exposures                                           |
| Specificity of health end points                                |
| Uniformity of diagnostic criteria                               |
| Statistical procedures                                          |
| Information available from individual studies for meta-analysis |
| Comparison populations                                          |
| Heterogeneity of effect, risk estimates                         |

words, heterogeneity of studies is more a scientific than a statistical issue. Nevertheless, statistical tests for heterogeneity of study results can sometimes provide clues to sources of disagreement among studies." If the test for heterogeneity is significant, the underlying studies must be scrutinised closely. Once the source for heterogeneity is identified, the decision whether to combine the studies must be carefully considered.

From the practical point of view, statistical procedures used in individual studies play an important part in the ability of studies to be combined, as do the type and amount of information presented in individual studies. Only studies presenting adequate information can be included in the meta-analysis. In occupational cohort studies, as will be discussed later, statistical procedures are usually similar and information needed for a meta-analysis is usually presented.

Although several study designs are appropriate for investigating occupational health, for chronic diseases such as cancer, most studies use the historical cohort design. Most occupational cohort studies are based on documented employment or exposure histories, and those included in meta-analyses usually consist of workers at similar places of work or from the same industry. As such, information on exposure is relatively reliable and exposure pattern is homogeneous. When linked to industrial hygiene data, such information can be relatively specific. Another epidemiological study design commonly used is the case-control study. Typically in a case-control study occupational information is based on recalls of the subjects themselves or family members, which may not be reliable or specific. Other population based case-control studies may use population registries to identify the industries and jobs of cases and controls. These different approaches to exposure classification between cohort and case-control studies can result in significant heterogeneity. Diagnostic criteria can also be different between cohort and case-control studies. Therefore, in most situations we do not recommend combining cohort and case-control studies in a single meta-analysis. The meta-analysis should at least be stratified by study design.

As the title suggests, this discussion will focus on the statistical procedure used in

occupational cohort studies—the standardised mortality ratio (SMR). In a meta-analysis, once it has been decided which studies will be combined, one of the goals is to calculate a quantitative summary estimate of the effect or risk of exposure, as well as the corresponding precision, usually in the form of a 95% confidence interval (95% CI). The summary estimate should make maximum use of the information available and preserve the properties of the original estimates in individual studies (such as adjustment for confounding).

In general, the summary estimate is some type of weighted average of the effect measures or risks from the original studies, with weights inversely proportional to the variance of the individual estimates—that is, a combination of stratum specific effect measures, with individual studies as the strata. For cohort mortality studies, the summary risk estimate is the summary SMR or meta-SMR, which is defined as the ratio of the total number of observed deaths (summed over all studies) to the total number of expected deaths. That is,

$$\text{meta-SMR} = \Sigma_i O_i / \Sigma_i E_i,$$

where the subscript  $i$  denotes the  $i$ th study,  $O$  the observed deaths, and  $E$  the expected deaths. Under the null hypothesis ( $\text{SMR} = 1$ ), the approximate variance of SMR in the  $i$ th study is  $1/E_i$ . Furthermore:

$$\text{meta-SMR} = \Sigma_i \{W_i \times \text{SMR}_i\}$$

where  $W_i = E_i / \Sigma E_i$  is the weight. That is, the weights for the meta-SMR are proportional to the inverse of the variance of individual studies. In other words, more weight is given to larger studies, and vice versa.

In calculating a meta-SMR, each study is treated as a separate stratum, and the data pooling is therefore similar to the indirect standardisation in individual studies. The total data from individual cohorts can be considered as a "megacohort." Figure 1 shows the concept graphically. It should be evident from both the statistical expression  $\text{meta-SMR} = \Sigma_i \{W_i \times \text{SMR}_i\}$  and fig 1 that a meta-SMR is based on a series of comparisons within each study, although the computational formula  $\text{meta-SMR} = \Sigma_i O_i / \Sigma_i E_i$  shows that it is the ratio of the total observed deaths to the total expected deaths. A meta-SMR is simply a summary risk index based on all the data, with

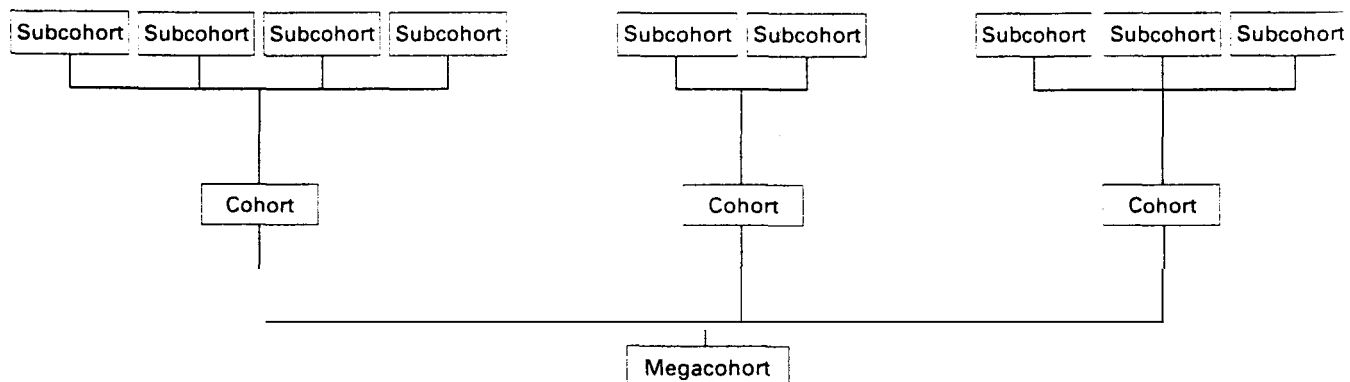


Figure 1 Relations between subcohorts, cohorts, and megacohort.

individual studies as a stratifying variable. The 95% CI of the meta-SMR can be calculated with the same procedure as SMRs for individual studies (for example, treating it as a Poisson variable<sup>16</sup>). Data needed to calculate a meta-SMR are the observed and expected deaths for the causes of interest from individual studies, which are usually reported.

Once the meta-analysis has been completed, the result should be integrated with the qualitative component of the meta-review (step 6), especially the discussion on the strengths and weaknesses of the individual studies. It should be emphasised that a meta-analysis will not compensate for any deficiencies that the individual studies may have, other than low statistical power. As such, it is important to discuss the impact of suspect individual studies on the meta-analysis. Analysis of sensitivity can be helpful for this. In such an analysis, studies of marginal quality, with slightly dissimilar exposures or slightly different from other studies in some aspects, are removed, and the result is then compared with the one that includes them.

The results of a meta-analysis are directly applicable to two important criteria of causation (step 7): strength and significance of association and consistency. The strength of association is given directly by the magnitude of the meta-SMR, and the statistical significance is measured by the 95% CI. Risk estimates from individual studies, especially for rare diseases, can be unstable and inadequate in statistical power. If the results from individual studies are in general agreement, however, a summary risk estimate such as the meta-SMR takes the consistency of results into consideration. The meta-SMR, based on the combined database, is more stable and is less vulnerable to fluctuations due to small numbers than the original studies. Thus, a meta-analysis takes into consideration not only the size of risk estimates from individual studies but also their consistency.

So far our discussion of statistical procedures in meta-analysis of cohort studies has been based on the summation of observed and expected deaths. This method is commonly used because the information needed is usually available. Depending on the information presented in the underlying studies, other statistical techniques may be used.<sup>17,18</sup> In general, the weights for the summary risk estimate are inversely proportional to the variance of the individual study.

The choice of a particular statistical technique depends on information presented in individual studies. Any valid statistical procedure of meta-analysis should be based on an objective weighting system. This implies that exposed and non-exposed groups are compared within each study. In other words, each study is treated as a stratum in the meta-analysis. The separate identity of each study is preserved, and it is the estimates of the effect from each study that are combined according to appropriate weights.<sup>19</sup> As already discussed, the method commonly used in combining cohort studies clearly satisfies this criterion. On

the other hand, simply pooling results from individual studies, ignoring study as a stratifying variable, is not a valid procedure.<sup>20</sup>

For completeness, we also mention another statistical issue of meta-analysis: fixed effects versus random effects models. Fixed effects models assume a single effect estimate for all studies, whereas random effects models take variability between studies into consideration. Most methods of meta-analysis commonly used (including the one discussed here) are fixed effects. There is an ongoing debate among epidemiologists about the choice of models. In this paper it suffices to say that the choice of models does not change the risk estimate, but may affect the width of the 95% CI. It should also be noted that, as already discussed, the definition of a study can be arbitrary and the concept of variability between studies can be ambiguous. Random effects models have been criticised for assigning more even weight to all studies, thus incorrectly overemphasising the results of small studies.<sup>21</sup> Interested readers are referred to some recent publications on the issue of fixed and random effects models.<sup>14,19,22</sup>

## Discussion and an example

Meta-analysis is a powerful tool in reviewing, summarising, and integrating scientific data from multiple studies. As stated by Glass *et al* "the findings of multiple studies should be regarded as a complex data set, no more comprehensible without statistical analysis than would hundreds of data points in one study."<sup>23</sup> A meta-analysis not only provides a summary risk estimate which is properly weighted, but also takes consistency of results of individual studies into consideration. An additional benefit of meta-analysis is that it comprehends the multiple comparison problem exhibited in individual cohort studies. Typically in an occupational cohort study more than 40 causes of deaths are examined. Statistically significant results for as many as two causes could have occurred by chance alone given a type 1 error of 5% (or a 95% confidence level). By combining data across studies, chance excesses or deficits for a particular cause in individual studies are less likely to influence the meta-analysis.

The meta-SMR discussed has several desirable attributes. It is the natural extension of the indirect standardisation of summarising results from age, sex, and race specific strata in the original studies. Being an extension of the indirect standardisation, the meta-analysis preserves the statistical properties of the original estimates as well as the original interpretation of the effect measure. In particular, by treating each study as a stratum, the summary effect measure preserves the adjustment for confounding made in the original studies. The number of expected deaths in each study is proportional to the size and age of the individual cohort as well as the duration of observation. From the practical point of view, the data required for a calculation of a meta-SMR are almost always available in the individual reports.

Table 4 Summary of mortality from brain cancer in petroleum workers<sup>a</sup>:

| Author, year                    | Place                         | Cohort size | Follow up | Observed deaths | Expected deaths | SMR (95% CI)         |
|---------------------------------|-------------------------------|-------------|-----------|-----------------|-----------------|----------------------|
| Wong 1980*                      | 17 US refineries              | 20163       | 1962-71   | 8               | 11.29           | 0.71 (0.31 to 1.40)  |
| Kaplan 1986*                    | 17 US refineries              | 20163       | 1962-80   | 22              | 24.63           | 0.89 (0.56 to 1.35)  |
| Schottenfeld <i>et al</i> 1981* | US refineries                 | 55007       | 1977-79   | 8               | 4.90            | 1.63 (0.70 to 3.21)  |
| Nelson 1985                     | 10 US refineries              | 10763       | 1970-82   | 8               | 9.60            | 0.83 (0.36 to 1.63)  |
| Wong <i>et al</i> 1986          | 2 California refineries       | 14179       | 1950-80   | 22              | 17.48           | 1.25 (0.78 to 1.89)  |
| Hanis <i>et al</i> 1982*        | 1 Louisiana refinery          | 8666        | 1970-77   | 5               | 4.90            | 1.02 (0.33 to 2.38)  |
| Hanis <i>et al</i> 1985         | 3 US refineries               | 21698       | 1970-77   | 15              | 13.00           | 1.15 (0.64 to 1.90)  |
| Wen <i>et al</i> 1983           | 1 Texas refinery              | 15095       | 1937-78   | 23              | 22.85           | 1.01 (0.64 to 2.37)  |
| Morgan and Wong 1984            | 1 Texas refinery              | 6139        | 1945-79   | 9               | 8.26            | 1.09 (0.50 to 2.07)  |
| Morgan and Wong 1985            | 1 Pennsylvania refinery       | 4263        | 1946-79   | 7               | 6.64            | 1.05 (0.42 to 2.16)  |
| Enterline and Henderson 1985    | 1 California refinery         | 1621        | 1959-78   | 3               | 2.09            | 1.43 (0.30 to 4.18)  |
| Divine <i>et al</i> 1985        | 13 US refineries              | 19077       | 1947-77   | 31              | 28.00           | 1.11 (0.76 to 1.42)  |
| Divine and Brown 1987           | Production and pipeline in US | 11098       | 1946-80   | 11              | 15.90           | 0.69 (0.35 to 1.58)  |
| Rushton and Alderson 1981       | 8 UK refineries               | 34781       | 1950-75   | 36              | 44.77           | 0.80 (0.56 to 1.11)  |
| Rushton and Alderson 1981       | Distribution centers in UK    | 23306       | 1950-75   | 39              | 36.54           | 1.07 (0.76 to 1.46)  |
| Theriault and Goulet 1979*      | 1 Canada refinery             | 1205        | 1928-75   | 3               | 0.77            | 3.89 (0.80 to 11.37) |
| Theriault and Provencher 1987   | 1 Canada refinery             | 1207        | 1928-81   | 4               | 1.87            | 2.13 (0.58 to 5.45)  |
| Christie 1987                   | Australia refineries          | 10489       | 1981-86   | 2               | 2.70            | 0.74 (0.09 to 2.67)  |
| Total (meta-analysis)           |                               |             |           | 210             | 209.70          | 1.00 (0.87 to 1.15)  |

\*Excluded from the calculation of the meta-SMR because of overlap of data.

Table 4 provides an example of meta-analysis of data on brain cancer in petroleum workers originally reported by Wong and Raabe in 1989.<sup>12</sup> As discussed in the original report, both published and unpublished (but otherwise meeting the eligibility criteria) cohort studies of petroleum workers were eligible for inclusion in the meta-analysis. The PMR studies were not included because of their inherent limitations. Case-control studies were also not included because of the inadequacy of exposure information (based primarily on recalls), but they were included in the discussion. An examination was made of all cohort studies for evidence for heterogeneity of cohort identification, data collection, and statistical analysis.

Based on this examination, two cohort studies were excluded because of methodological or data deficiencies. For the remaining cohort studies, it was decided that no significant heterogeneity was present and the data could be combined in a meta-analysis. All these studies were historical cohort studies based primarily on employment records. Exposures in these studies were relatively homogeneous in the sense that all study subjects were petroleum workers. Methods of follow up and comparison populations were also similar. In the meta-analysis, duplicate data (such as industrywide studies or data based on earlier follow ups) were excluded.

Table 4 and figure 2 show that the brain

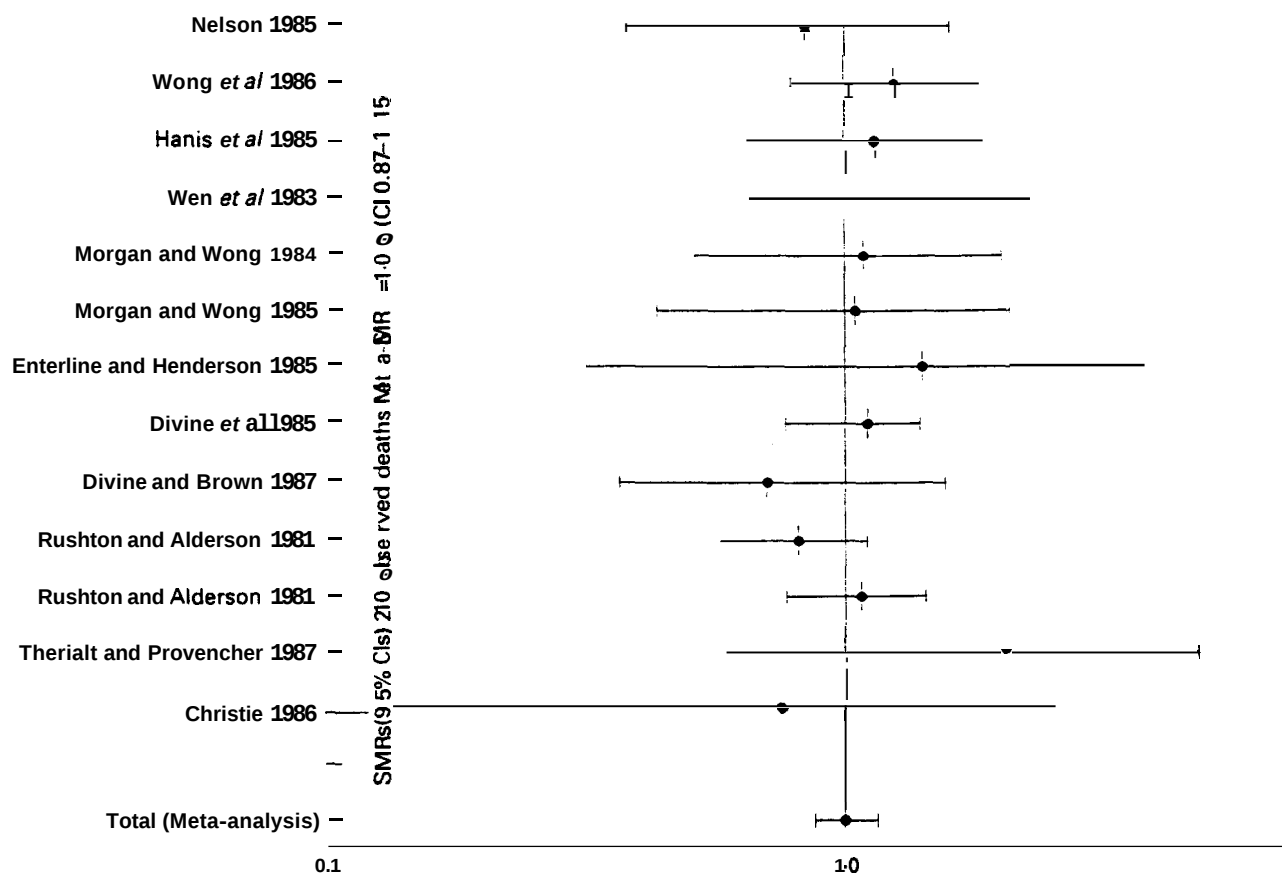


Figure 2 Mortality from brain cancer in petroleum workers.

cancer meta-analysis consisted of listing the SMRs and the corresponding 95% CIs from individual studies as well as the meta-SMR and its 95% CI. The presentation of risk estimates as well as other key study variables from all individual studies in a single table is obviously very informative. Similarly, the graphical Presentation is a helpful visual aid. Unfortunately, in some recent meta-analyses only graphical presentations are available. The lack of a presentation of numerical data often makes it difficult for the readers to determine what data from the original studies were actually used in the meta-analyses. Thus, we recommend presenting data both numerically and graphically.

In our example of brain cancer in petroleum workers, the meta-SMR was 1.00 (95% CI 0.87 to 1.15). Therefore, the combined data had reasonable power to rule out an increase in risk of brain cancer as small as 15%. The improvement in statistical power through the use of meta-analysis is evident from the narrow 95% CI of the meta-SMR and the small increase in risk that the data could detect. A meta-review should not stop at the quantitative meta-analysis. In the example of brain cancer in petroleum workers, the result of the meta-analysis based on cohort studies was discussed in conjunction with results from case-control studies. It was concluded that petroleum workers did not experience any increased risk of brain cancer."

It should be pointed out that the application of meta-analysis is not without criticisms. Some of the common criticisms are: the inclusion of both good and poor studies; publication bias resulting from the tendency of positive studies being more likely to be published; the reliance on published studies only; overrepresentation of data from studies published repeatedly; and most seriously, the oversimplification of a complicated issue by focusing on a single summary risk estimate. Most of these criticisms arise from improper application of meta-analysis. The general guidelines provided earlier should help to minimise these problems. It should also be noted that all of these criticisms, except for oversimplification, are equally applicable to a traditional qualitative review.

The last criticism of oversimplification of study results is equally applicable to summaries of data specific to strata—for example, age groups—in an individual study. Several decades ago, the pros and cons of summarising data in an individual study were vigorously debated.<sup>25-29</sup> The issue at that time was whether risk estimates from different age groups in a study should be summarised with a single index such as SMR; yet today it would be inconceivable that SMRs or some other summary indices are not calculated in a cohort study. As already discussed, a reasonable and practical approach is to examine data for evidence of heterogeneity, and if necessary provide stratified meta-SMRs or sensitivity analyses. The issue of heterogeneity is more scientific than statistical. Although a statistical test for heterogeneity of study results is help-

ful, other aspects of the original studies must be investigated before studies can be combined (table 3).

The usefulness of a meta-analysis is obviously dictated by the quality of the input data of the original studies and the amount of information available to the meta-analyst. With regard to the amount of information available, hopefully in the future authors will report their investigations in a uniform manner and provide sufficient details to facilitate any future meta-analysis, and journal editors will include as part of their decisions the potential contribution of a manuscript to a future meta-analysis of similar studies.

## Summary

Meta-analysis is useful in selecting studies, and in organising, presenting, and summarising results from individual studies. The benefits of conducting a meta-analysis in a review include enhancing statistical power (particularly when original studies are small), providing a summary risk estimate, taking consistency of study results into consideration, minimising the problem of multiple comparisons, and examining data for heterogeneity. Compared with a traditional qualitative review, a review incorporating a meta-analysis tends to be less subjective. Rather than replacing the traditional qualitative review, a meta-analysis should be made part of the overall assessment. We propose the use of the term meta-review to emphasise the importance of both qualitative review and quantitative meta-analysis.

- 1 Glass GV. Primary, secondary and meta-analysis of research. *Educational Research* 1976;5:3-8.
- 2 Cochran WG. Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society Supplement B* 1937;4:102-18.
- 3 Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101-29.
- 4 Fisher RA. *Statistical methods for research workers*. 4th ed. London: Oliver and Boyd, 1932.
- 5 Pearson ES. The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika* 1938;30:134-48.
- 6 Surgeon General. *Smoking and health*. Washington: US Government Printing Office, 1964.
- 7 Huque MF. Experience with meta-analysis in NDA submissions. *Proceedings of the Biopharmaceutical Section of the American Statistical Association* 1988;2:28-33.
- 8 Webster's new world dictionary of American English. New York: Prentice Hall, 1994.
- 9 Webster's new collegiate dictionary. Springfield, Massachusetts: Merriam, 1990.
- 10 Cook DJ, Guyatt GH, Ryan G, Clifton J, Buckingham L, Willan A, et al. Should unpublished data be included in a meta-analysis? *JAMA* 1993;269:2749-53.
- 11 Morgan RW, Foliant DE, Wung O. Asbestos and gastric cancer, a review of the literature. *West J Med* 1985;143:60-5.
- 12 Wong O, Raabe GK. Critical review of cancer epidemiology in petroleum industry employees, with a quantitative meta-analysis by cancer site. *Am J Ind Med* 1989;15:283-310.
- 13 Wung O, Raabe GK. Cell-type-specific leukemia analyses in a combined cohort of more than 208 000 petroleum workers in the United States and the United Kingdom, 1937-89. *Regul Toxicol Pharmacol* 1995;21:307-21.
- 14 Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994;140:290-h.
- 15 Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987;9:1-30.
- 16 Bailar JC, Ederer F. Significance factors for the ratio of a Poisson variable to its expectation. *Biometrics* 1964;20:639-43.
- 17 Armitage P, Berry G. *Statistical methods in medical research*. 3rd ed. Oxford: Blackwell, 1994.

- 18 Morgan RW, Wong O. A review of epidemiological studies on artificial sweeteners and bladder cancer. *Fundam Appl Toxicol* 1985;23:529-33.
- 19 Pocock SJ, Thompson SG. Assessing the evidence: the role of meta-analysis. In: Marmot M, Elliott P, eds. *Coronary heart disease epidemiology, from aetiology to public health*. Oxford: Oxford University Press, 1991.
- 20 Dickersin K, Berlin JA. Meta-analysis: state-of-the-science. *Epidemiol Rev* 1992;14:154-75.
- 21 Peto R. Why do we need systematic reviews of randomized trials? *Stat Med* 1987;6:242.
- 22 Meier P. Commentary. *Stat Med* 1987;6:329-31.
- 23 Bailey KR. Inter-study differences: how should they influence the interpretation and analysis of results? *Stat Med* 1987;6:351-8.
- 24 Glass G, McGaw B, Smith ML. *Meta-analysis in social research*. Beverly Hill, California: Sage, 1981.
- 25 Yule GU. On some points relating to vital statistics, more especially statistics of occupational mortality. *Journal of the Royal Statistical Society* 1934;97:1-72.
- 26 Kitagawa EM. Component of a difference between two rates. *Journal of the American Statistical Association* 1955;50:1168-94.
- 27 Kitagawa EM. Standardized comparisons in population research. *Demography* 1964;1:296-315.
- 18 Kitagawa EM. Theoretical considerations in the selection of a mortality index, and some empirical comparisons. *Hum Biol* 1966;38:293-308.
- 29 Elveback LR. Discussion of "Indices of mortality and tests of their statistical significance." *Hum Biol* 1966;38:322-4.

## Vancouver style

All manuscripts submitted to *Occup Environ Med* should conform to the uniform requirements for manuscripts submitted to biomedical journals (known as the Vancouver style.)

*Occup Environ Med*, together with many other international biomedical journals, has agreed to accept articles prepared in accordance with the Vancouver style. The style (described in full in the *BMJ*, 24 February 1979, p 532) is intended to standardise requirements for authors.

References should be numbered consecutively in the order in which they are first mentioned in the text by Arabic numerals above the line on each occasion the reference is cited (Manson<sup>1</sup> confirmed other reports<sup>2,3</sup> . . .). In future references to papers submitted to *Occup Environ Med*

should include: the names of all authors if there are seven or less or, if there are more, the first six followed by *et al*; the title of journal articles or book chapters; the titles of journals abbreviated according to the style of *Index Medicus*; and the first and final page numbers of the article or chapter. Titles not in *Index Medicus* should be given in full.

Examples of common forms of references are:

- 1 International Steering Committee of Medical Editors. Uniform requirements for manuscripts submitted to biomedical journals. *BMJ* 1979;1:532-5.
- 2 Soter NA, Wasserman SI, Austen KF. Cold urticaria: release into the circulation of histamine and eosinophil chemotactic factor of anaphylaxis during cold challenge. *N Engl J Med* 1976;294:687-90.
- 3 Weinstein L, Swartz MN. Pathogenic properties of invading micro-organisms. In: Sodeman WA Jr, Sodeman WA, eds. *Pathologic physiology, mechanism of disease*. Philadelphia: W B Saunders, 1974:457-72.

M CC  
Dep  
Occ:  
Fact:  
3450  
Suite  
Qué:  
A Du  
CM  
PLo  
Sen  
Pne:  
Hosp  
PUn  
Sher  
Sher  
Can:  
R Be  
S Ma  
Insti  
en S:  
Trav  
Labo  
oues  
Mais  
Monr  
Cana  
G Per  
Corre  
Dr A I  
Unive:  
Occup  
Facult  
Univer  
Montr  
H3A 2  
Accept