

Bayesian meta-analysis, with application to studies of ETS and lung cancer

Tweedie RL, Scott DJ, Biggerstaff BJ and Mengersen KL
Department of Statistics
Colorado State University
Fort Collins CO 80523, USA

November 14, 1994

Abstract

Meta-analysis enables researchers to combine the results of several studies to assess the information they provide as a whole. It has been used to give a systematic overview of many areas in which data on a possible association between an exposure and an outcome have been collected in a number of studies but where the overall picture remains obscure, both as to the existence or size of the effect.

This paper outlines some innovations in meta-analysis, based on using Markov chain Monte Carlo (MCMC) techniques for implementing Bayesian hierarchical models, and compares these with a more well-known random effects (RE) model. The new techniques allow different aspects of variation to be incorporated into descriptions of the association, and in particular enable us to better quantify differences between studies.

We apply both the classical and Bayesian methods to the current collection of studies of the association between incidence of lung cancer in female never-smokers and exposure to environmental tobacco smoke (ETS), both in the home through spousal smoking and in the workplace. We demonstrate that, compared with the RE model, the Bayesian methods

- (a) allow more detailed modelling of study heterogeneity to be incorporated;
- (b) are relatively robust against a wide choice of specification of such information;
- (c) allow for more detailed and satisfactory statements to be made, not only about the overall risk but about the individual studies, on the basis of the combined information.

For the workplace exposure data set, the Bayesian methods give a somewhat lower overall estimate of relative risk of lung cancer associated with ETS, indicating the care that needs to be taken in using point estimates based on any one method of analysis. On the larger spousal data set the methods give similar answers.

We also consider some of the other concerns with meta-analysis, such as consistency between different geographic areas (such as Asia and the US), and show that Bayesian methods allow us to take into account the overall picture, thus improving the ability to estimate accurately in the subgroups; and publication bias, which we find with the spousal exposure data may lead to an inflated excess risk.

⁰Keywords: Exact tests, Environmental tobacco smoke, Fixed effects, Lung cancer, Meta-analysis, Passive smoking, Publication bias, Random effects, Bayesian analysis

1 Introduction

In recent years there has been an enormous increase in the use of meta-analysis in many medical areas in order to obtain overall evaluations of association in areas where individual studies are equivocal [61]. With this has come a large number of discussion papers which assess the benefits, drawbacks and problems of these techniques (see for example [57, 19, 11, 58, 71]).

Some of the most well documented concerns are about the way in which data can be combined if the collection of studies is not homogeneous by design but is based on a variety of differently structured epidemiological cohort or case control studies [28, 51]. Some of these concerns are matters of judgement, and relate to such issues as differing aims of studies or differing study quality including control of confounders; others relate to the underlying variability in the information presented, and different statistical approaches have been developed to attempt to quantify this objectively.

In the epidemiological literature a standard method of combining estimates of interest is via a “random effects” model, which attempts to allow for inter-study variation, perhaps due to uncontrolled covariates [58]. This has been argued to be preferable to an earlier “fixed effects” model which essentially assumes that any heterogeneity between studies is purely random (cf. [86, 81]) and hence is not modelled explicitly.

The random effects model can be analysed both in a frequentist or a Bayesian framework [58]. In the latter context it extends logically to hierarchical models such as those recently proposed by DuMouchel [16, 15] or Carlin [10]. In order to differentiate between the models we shall refer to the frequentist random effects model as the “RE model” and the hierarchical Bayes model, which is also formally a random effects model, as the Bayesian model. Details of these are given in the Appendix. Interpretations of the two types of statistical approach are different but the context should make the interpretations clear.

Two advantages of the Bayesian approach are its greater flexibility in utilising other (often prior) information or relationships, and the ability to make useful probability statements on the basis of all information. Moreover, new Markov chain Monte Carlo (MCMC) methods now allow analysis of models based on very general formulations of such prior information, which were previously thought to result in mathematical expressions too complex to be solved. Through their use a wider range of inferences can be made in a straightforward way [2], as we demonstrate here.

In the Bayesian meta-analysis context, we will use MCMC to analyse such hierarchical models, without the need to approximate the solutions. Although we do not pursue them here, we note that there are alternative approaches to combining epidemiological studies, also using MCMC methods: a logistic- t model with additional unknown covariates is proposed in [3], and methods of multiple comparisons, proposed in the NRC Report [58, pp 149-158] for detecting non-equivalence between populations, can also be approached through MCMC [55].

In this paper the Bayesian methods are not used, in the main, to describe “prior information” in any strong sense. Rather, one can view the models as describing in more detail the way in which the studies might be heterogeneous, and this allows one to account more explicitly for greater variability in the underlying collection of studies than is done in the fixed or even the RE models.

After describing the methods of analysis, we illustrate a Bayesian MCMC approach through an assessment of the overall association between incidence of lung cancer in female never-smokers and

exposure to environmental tobacco smoke (ETS), or “passive smoking”, both in the workplace and from spousal smoking.

There have been many meta-analyses of the individual studies of ETS exposure associated with spousal smoking ([81, 17, 18, 74, 47, 56] and others) but there has been limited assessment of the current set of papers addressing general workplace exposure, apart from [47, 48, 4]. The results we give below can be compared with those we have derived using somewhat simpler methods in [56, 4], and those papers contain a more complete discussion of aspects we merely survey here.

The ETS studies seem appropriate for meta-analysis for several reasons. Although the association between lung cancer and ETS is an issue of public and legal concern, there has been a tendency to extrapolate results for spousal smoking to the workplace arena (cf [62]). By utilizing all current information about workplace and spousal exposures explicitly, overall estimates of the relative risk relevant to the particular exposure and their variability may be constructed directly and compared across the two sources of exposure.

Our main focus is on statistical methods in this paper. However, any meta-analysis involves choices, not just of the statistical methods, but of many other things: choice of relevant studies, data within those studies, and the preceding statistical analyses of these individual datasets. We assess each of these in the ETS application, and consider whether such choices can influence the outcome. We specifically comment below on

- (i) methods of accounting for homogeneity of studies, and whether the Bayesian methods give different inferences from RE models;
- (ii) the problem of comparability of data and study design so that the meta-analysis can be meaningfully interpreted: this includes the choice of subgroups of studies to include, and whether, say, to include studies from different countries, or case-control and cohort studies;
- (iii) the use of unadjusted or adjusted published data: the meta-analyst is often faced with the problem that some studies report estimates unadjusted for any covariates (often as a 2×2 table), so any effects of covariates are obscured, whilst others report only estimates of relative risks adjusted for (usually different) covariates;
- (iv) the effect of publication bias, recognising that failure to obtain all relevant studies, both published and unpublished, may result in a quite distorted meta-analysis.

Clearly all of these are of concern in principle, but it is not obvious whether they will cause real problems in specific applications. Here we attempt to quantify the degree to which such issues might affect the meta-analyses of ETS studies in practice under the particular models we consider.

2 Data and Analysis of Studies on Exposure to ETS

2.1 Data Comparability and Bias

Meta-analysis is designed to enable combination of results from studies which are *comparable in outcome and exposure*. The interpretation of comparability is a subjective and often difficult one.

In order to paint an honest picture of the aims and applicability of any meta-analysis, we must first define the relevant measures of outcome and exposure with which we are concerned.

The clinical *outcome* assessed in all of the ETS studies is death from “lung cancer”. Several concentrate on or are dominated by one specific form of this disease (e.g. adenocarcinoma), and although some studies give data for different types of cancer, many others do not make such distinctions. Here we choose to combine RR estimates for all lung cancer types, but we are aware that the overall RR estimate may be based on individual RR’s associated with quite different diseases in different studies.

In order to identify studies with comparable *exposures* we primarily restrict the meta-analysis to the subset of all ETS studies of adults asserted to be never-smokers, with exposure to spousal smoking or workplace smoking the declared type of exposure to ETS. However, the relevant data are unavailable in a few “spousal” studies, and for these the restrictions are relaxed slightly to include other household exposure or long-time nonsmokers; see Lee [47] and the EPA Report [18] for further details.

In choosing which studies to combine, we also need to consider the plausibility of comparing different subpopulations. Two obvious questions are whether there are gender or geographic differences. In accord with the practice in most individual studies and other meta-analyses of these data, we have analysed males and females separately, and it is the latter that we report here. For males exposed to ETS in the workplace there is a comparable analysis in [4]. The geographic question seems more appropriately studied through a sensitivity analysis as in [18] and we do so in Section 3.3.

It is also crucial in meta-analysis to attempt to collect all studies relevant to the relationship in question [27]. This involves collecting at least all *published* studies, if possible, and testing for the potential existence and influence of unpublished or uncollected studies. There is an insufficient number of studies of workplace exposure to decide if there might be missing information due to publication bias. In contrast, for the spousal exposure studies detailed in the next section, it is possible to investigate completeness using funnel plots (see [51]), and in Figure 1 of [56] there is a clear indication of the absence of small studies with negative (perhaps nonsignificant) estimates of effect. It does appear from this that there is indeed bias towards publication of raised relative risks, with perhaps 6-10 or so small but negative studies expected but not absent: this may impact on our overall results, and we comment on this in Section 4.

Overall, our experiences with collating these data strongly reinforce those of Felson [19] and Chalmers [11]: data extraction and location is a non-trivial exercise, there are considerable problems in locating studies and relevant data within them, and there are many subjective decisions about data collection and analysis which need to be explicitly documented. We attempt to do this in the following sections.

2.2 Spousal and workplace exposure to ETS

Table 1 lists all studies known to us, through Medline and Cancerlink searches and reference to published reviews [18, 47, 49], which provide data relevant to a meta-analysis of the association between ETS and lung cancer in nonsmoking adults, using spousal smoking as the primary measure of exposure. This currently comprises 40 studies of which 3 are unpublished theses, and we give details of location and sex studied.

TABLE 1: Studies which provide relative risk estimates associated with lung cancer for female nonsmokers exposed to ETS, as measured by spousal smoking and workplace exposure

Study	Country	Sex	Workplace
Case Control Studies			
1	Akiba et al [1]	JAPAN	M+F
2	Brownson et al [7]	US	M+F
3	Brownson et al [6]	US	F
4	Buffler et al [8]	US	M+F
5	Chan and Fung [12]	HONG KONG	F
6	Correa et al [13]	US	M+F
7	Du et al [14]	CHINA	F
8	Fontham et al [21]	US	F YES
9	Gao et al [22]	CHINA	F
10	Garfinkel et al [24]	US	F YES
11	Geng et al [25]	CHINA	F
12	Ger et al [26]	TAIWAN	M+F
13	Humble et al [33]	US	M+F
14	Inoue and Hirayama [34]	JAPAN	F
15	Janerich et al [35]	US	M+F YES
16	Joeckel [36]	GERMANY	F
17	Kabat [37]	US	M+F YES
18	Kabat and Wynder [38]	US	M+F YES
19	Kalandidi et al [39]	GREECE	F YES
20	Koo et al [40]	HONG KONG	F YES
21	Lam T. et al [42]	HONG KONG	M+F
22	Lam W. [43]	HONG KONG	F
23	Lee et al [50]	ENGLAND	M+F YES
24	Liu et al [53]	CHINA	F
25	Liu et al [52]	CHINA	F
26	Pershagen et al [63]	SWEDEN	F
27	Shimizu et al [65]	JAPAN	F YES
28	Sobue et al [67]	JAPAN	F
29	Stockwell et al [69]	US	F
30	Svensson et al [70]	SWEDEN	F
31	Trichopolous et al [73]	GREECE	F
32	Varela [77]	US	M+F
33	Wang et al [82]	CHINA	F
34	Wu et al [84]	US	F YES
35	Wu-Williams et al [85]	CHINA	F YES
36	Ziegler et al [87]	US	M
Cohort Studies			
37	Butler [9]	US	F
38	Garfinkel [23]	US	F
39	Hirayama [29, 30]	JAPAN	M+F
40	Hole et al [32]	SCOTLAND	M+F

TABLE 2: Unadjusted RR, Bayesian shrinkage estimates and adjusted RR with 95% CI for female nonsmokers exposed to ETS through spousal smoking.

Study No.	RR	Unadjusted		Bayes	Adjusted
		Logit CI	Exact CI	RR (CI)	RR (CI)
1	1.52	(0.87-2.63)	(0.85-2.77)	1.31 (0.96-1.83)	1.50 (0.9-2.8)
2	1.82	(0.45-7.36)	(0.33-8.90)	1.27 (0.86-1.88)	1.68 (0.39-2.97)
3	0.96	(0.77-1.20)	(0.77-1.21)	1.03 (0.84-1.23)	1.0 (0.8-1.2)
4	0.80	(0.34-1.90)	(0.32-2.21)	1.15 (0.79-1.59)	–
5	0.75	(0.43-1.30)	(0.42-1.35)	1.06 (0.73-1.41)	–
6	2.07	(0.81-5.25)	(0.75-6.06)	1.33 (0.93-1.98)	–
7	1.09	(0.64-1.85)	(0.62-1.93)	1.19 (0.86-1.60)	–
8	1.26	(1.04-1.54)	(1.14-1.40)	1.25 (1.06-1.49)	1.29 (1.04-1.60)
9	1.19	(0.82-1.73)	(0.80-1.77)	1.21 (0.92-1.57)	About 1.4
10	1.23	(0.81-1.87)	(0.80-1.92)	1.23 (0.93-1.60)	–
11	2.16	(1.08-4.29)	(1.03-4.56)	1.41 (0.99-2.06)	–
12	N/A	N/A	N/A	N/A	1.18 (0.47-2.99)
13	2.34	(0.81-6.75)	(0.76-8.59)	1.34 (0.92-1.98)	2.20 (0.8-6.6)
14	2.55	(0.74-8.78)	(0.67-11.91)	1.32 (0.91-1.99)	2.25 (0.8-8.8)
15	N/A	N/A	N/A	N/A	0.93 (0.55-1.57)*
16	2.27	(0.75-6.82)	(0.68-8.28)	1.32 (0.91-1.99)	–
17	0.90	(0.46-1.76)	(0.44-1.88)	1.15 (0.80-1.57)	–
18	0.79	(0.25-2.45)	(0.22-2.83)	1.18 (0.79-1.68)	–
19	1.55	(0.87-2.83)	(0.83-2.92)	1.32 (0.96-1.84)	2.11 (1.09-4.08)
20	1.55	(0.90-2.67)	(0.86-2.77)	1.32 (0.96-1.81)	1.64 (0.87-3.09)
21	1.65	(1.16-2.35)	(1.14-2.39)	1.43 (1.09-1.87)	–
22	2.01	(1.09-3.72)	(1.04-3.92)	1.41 (1.01-2.01)	–
23	1.03	(0.41-2.55)	(0.38-2.88)	1.20 (0.82-1.69)	1.00 (0.37-2.71)
24	0.74	(0.32-1.69)	(0.31-1.92)	1.13 (0.75-1.57)	–
25	1.66	(0.73-3.78)	(0.68-4.14)	1.30 (0.91-1.89)	–
26	1.03	(0.61-1.74)	(0.59-1.80)	1.16 (0.83-1.54)	1.20 (0.7-2.1)
27	1.08	(0.64-1.82)	(0.62-1.89)	1.18 (0.86-1.56)	–
28	1.06	(0.74-1.52)	(0.73-1.55)	1.14 (0.88-1.46)	1.13 (0.78-1.63)
29	N/A	N/A	N/A	N/A	1.6 (0.8-3.0)
30	1.26	(0.57-2.81)	(0.54-3.16)	1.24 (0.87-1.73)	About 1.5
31	2.08	(1.20-3.59)	(1.16-3.76)	1.45 (1.05-2.09)	–
32	0.75	(0.47-1.20)	(0.46-1.23)	1.02 (0.72-1.34)	–
33	1.41	(0.54-3.67)	(0.49-4.20)	1.25 (0.86-1.80)	1.20 (0.50-3.30)
34	1.20	(0.48-3.01)	(0.39-3.40)	1.23 (0.84-1.76)	–
35	0.79	(0.62-1.02)	(0.61-1.02)	0.92 (0.72-1.14)	0.7 (0.6-0.9)
36	N/A	N/A	N/A	N/A	–
37	2.44	(0.58-10.22)	(0.38-12.55)	1.30 (0.88-1.96)	2.02 (0.48-8.56)
38	1.17	(0.85-1.61)	(0.84-1.64)	1.20 (0.94-1.50)	1.18
39	1.39	(0.97-1.98)	(0.96-2.04)	1.30 (1.01-1.68)	1.45 (1.02-2.08)
40	1.89	(0.22-16.23)	(0.21-8.96)	1.25 (0.83-1.89)	2.41 (0.45-12.83)*
Overall	1.20	(1.07-1.34)		1.22 (1.08-1.37)	

* Results for sexes combined

TABLE 3: Unadjusted RR, Bayesian shrinkage estimates and adjusted RR with 95% CI for female nonsmokers exposed to ETS through workplace smoking.

Study No.	RR	Unadjusted		Bayes	Adjusted
		Logit CI	Exact CI	RR (CI)	RR (CI)
Case Control					
8	1.12	(0.91-1.36)	(1.01-1.24)	1.11 (0.94-1.32)	1.34 (1.11-1.74)
10	0.93	(0.55-1.55)	(0.53-1.60)	1.07 (0.79-1.38)	0.93 (0.73-1.18)
15	N/A	N/A	N/A	N/A	0.91 (0.80-1.04)*
17	1.00	(0.49-2.06)	(0.46-2.21)	1.10 (0.78-1.45)	1.00 (0.49-2.06)
18	0.68	(0.32-1.47)	(0.30-1.58)	1.05 (0.72-1.37)	0.68
19	1.39	(0.76-2.54)	(0.73-2.67)	1.15 (0.86-1.54)	1.39 (0.76-2.54)**
20	0.91	(0.15-5.37)	(0.08-6.95)	1.10 (0.77-1.51)	0.91
23	0.63	(0.17-2.33)	(0.11-2.49)	1.07 (0.74-1.45)	0.63 (0.17-2.33)
27	1.18	(0.68-2.03)	(0.66-2.09)	1.12 (0.84-1.46)	1.2
34	N/A	N/A	N/A	N/A	1.3 (0.5-3.3)
35	1.22	(0.95-1.57)	(0.94-1.58)	1.16 (0.95-1.41)	1.1 (0.9-1.6)
Overall	1.12	(0.93-1.28)		1.10 (0.90-1.32)	

* Combined Males and Females

** Calculated from [39, Table 2] compared to housewives as unexposed group

The 40 studies in Table 1 comprise 30 studies as described by Lee [47, pp. 101-105], and 10 more recent studies, most of which are also reviewed in the EPA Report [18]. Nine other related studies which do not provide data usable in our meta-analysis are not considered here; Lee [47] provides details of these.

Having chosen the studies for inclusion, there is then a question of choice of *data* to be settled. Different values can be extracted from different parts of some studies: one could for example use different criteria for inclusion of subjects, such as inclusion of ex-smokers or cigar smokers, single or widowed subjects, surrogate respondents, or disease rather than death. In previous analyses [56] we considered these choices and found that they made little difference to this meta-analysis. We chose therefore to adopt the data tabulated in [47, Tables 3.13F and 3.13M], which is well documented [47, pp. 102-103]; fortunately, for the 10 more recent studies [6, 20, 35, 69, 53, 26, 36, 14, 52, 82] the abstraction of comparable data is straightforward.

It is harder to assemble a coherent set of studies with sufficient data reported to enable meta-analysis of the relative risk of lung cancer associated with “general” exposure to workplace ETS.

Table 1 also indicates that 11 of these papers contain statistics or data concerning exposure to ETS in the general workplace. These studies comprise, to our knowledge, all those currently published on this association, based on a Medline search and various reviews. Discussion and review of these studies may be found in Lee [47, p. 117-118] and [48, p. 37-41] (except the results based on [40, Table 2] which is not used there and [20] which is more recent) and we do not repeat details here.

We have omitted from our workplace meta-analyses published studies concerning *occupation-specific* environments such as passenger cabins in commercial airlines [60, 31] or the food service industry [66]; and also the study of Brownson *et al* [7] which relates to smokers and non-smokers and only considers specific high lung-cancer risk occupations, since these are sufficiently different in design to violate the applicability of our models [58]. We have also only included studies for which the exposure is solely in the workplace, excluding those (Lam and Cheng [41] and Svensson *et al* [70])

which give relative risks for lung cancer when ETS is measured through exposure “at home or at work” or “at home and at work”.

2.3 Exact and Approximate Analyses of Individual Studies

Typically, studies report results either in “crude” or “unadjusted” form, as 2×2 tables, or as “reported” results, which may be adjusted for covariates as described by the individual authors.

Ideally one would wish to construct a model with complete control of such covariates (eg [83]). Most often, however, the required information is not available in published epidemiological papers. Instead meta-analysis must be performed only on the basis of summary statistics. These statistical quantities of interest in the individual studies, which are later combined in our meta-analyses, are the point estimates and associated confidence intervals (CIs) of the relative risk (RR) of outcome in a population with some defined exposure (either spousal or workplace ETS in our examples), compared with outcome for an unexposed population.

In Tables 2 and 3 we first provide analyses of the unadjusted data for the spousal and workplace studies respectively. A more detailed description of the methodology we use is relegated to the Appendix, and here we describe the notation and quantities needed to interpret these tables.

We use the following notation throughout: we suppose that we have k studies, and that

$$\begin{aligned} \text{RR}_i &= \text{observed estimate of relative risk in study } i, \\ Y_i &= \log \text{RR}_i, \\ \theta_i &= \text{true log relative risk in study } i, \\ W_i &= \text{an appropriate estimate of } (\text{Var}[Y_i])^{-1}. \end{aligned}$$

In the traditional setting for epidemiological studies, the empirical odds ratio provides a point estimate of the true relative risk for each study, and we use this throughout in this paper.

Tables 2 and 3 also contain estimates of the individual parameters θ_i and corresponding confidence intervals based on logit approximations to the variance, with an assumption of normality of the Y_i which is known to be reasonable, at least for large individual sample sizes. As seen in these Tables we find that, compared with an exact method (also discussed in the Appendix) the logit method gives CI’s that are perhaps 5-10% too short; but for our purposes we will accept this level of accuracy here. (In [4] these methods are also compared with the results generated by Mantel-Haenszel methods [5, p.141], [64], which are found to be typically less accurate again.)

Even without 2×2 tabulations, reported results may be combined provided all the confidence intervals are also reported, through deriving a Normal-based variance estimate for the log relative risk estimate. This is the case for many of the studies in Tables 2 and 3. Note, however, that the different factors for which adjustment was made in each of the studies render it more difficult to be sure that like is being compared with like in such an analysis.

In Table 2 we see that 28 of the 36 relevant studies with female respondents reported an increase in the unadjusted relative risk of lung cancer associated with spousal ETS exposure, with just 5 of these significantly different from 1.0 at the 95% level. (Because we use both frequentist and Bayesian methods, it will be convenient to define the phrase “significantly different from 1.0” to

cover either the situation in which there is a constructed 95% confidence interval which does not cover 1.0, or a Bayesian 95% credible interval which does not cover 1.0: the context should make it clear which is meant.)

In Table 3 we see that only 4 of the 9 relevant studies with female respondents reported an increase in the unadjusted relative risk of lung cancer associated with spousal ETS exposure, with just one of these significantly different from 1.0 (as indicated from the exact CI).

Thus, as stated above, both of these collections of studies are certainly such that a simple interpretation is difficult and in which heterogeneity may well be a problem that both the RE and the Bayesian analyses can help to overcome.

2.4 Random Effects and Bayesian Approaches to Meta-analysis

The RE model for meta-analysis is a natural starting point to describe a Bayesian methodology for meta-analysis. As described more formally in the Appendix, in the RE method we assume that there is a true underlying log RR over all studies, denoted μ , and that the observed log relative risks Y_i for each study are from a distribution governed by quantities θ_i and σ_i^2 which represent the true RR and within-study variability of study i , and a quantity τ^2 which provides a measure of the between- or across-study variability. In the special case in which $\tau^2 = 0$, indicating homogeneity between studies, this RE model reduces to the well-known fixed effects model (see [86, 81] and others).

In this non-Bayesian paradigm, μ, σ and τ are presumed fixed, and the θ 's are random variables with mean μ . In a general hierarchical Bayesian scheme [16], σ_i^2 and τ^2 are also random variables with (in our case) a χ^2 distribution, and these χ^2 distributions are in turn governed by parameters (degrees of freedom) df_σ and df_τ which indicate how well the variance structures are assumed to be known.

The distributions of these quantities are specified *a priori* according to the application. It is standard practice to assume a “flat” or “uninformative” prior for μ as we do below, as even with a small number of studies, “the combined data become relatively informative about the location of the effect-size prior distribution” [10, p. 146]. The imposition of distributions on θ, σ^2 and τ^2 enables us to describe much more explicitly any underlying variability in the way the study outcomes are distributed.

In this formulation, the posterior distributions become quite complicated, leading DuMouchel [16] to make some (reasonable) approximations to normality for computational convenience. In contrast, in this paper we use simulation methods (specifically the Gibbs sampler through the software package BUGS [68]) to carry out the analysis. As previously mentioned, these algorithms provide powerful computational tools for Bayesian analysis and release the user from restrictive assumptions about the distribution of the data and of prior information [2].

In the ETS case, for example, although the model (2) in the Appendix was considered appropriate, approximations to the posterior distribution were not needed, although comparison of our results here to those in [4] show that the Normal approximations of DuMouchel [16] are in fact very effective in this case.

The Bayesian method, as implemented through MCMC software, also enables us to make inferences about the posterior probability that the overall relative risk is above 1.0, enabling more exact inferences to be made and thus more effectively enabling the meta-analysis to achieve one of its overall goals. It is equally possible to quantify statements such as $P\{\text{overall US mean} > 1\}$ using this method, which is not a simple task in the RE models.

In this paper we will show that the use of this more flexible description of the way in which relative risks are spread across studies can lead to small but possibly important differences in the overall conclusions made, and that these conclusions are essentially independent of how the prior distributions are chosen, so that in fact it is the data that are driving the conclusions.

3 Results

3.1 Analysis based on unadjusted relative risks

The results of meta-analyses under both the RE and Bayesian paradigms are given as the “Overall” values at the bottoms of Tables 2 and 3. In the second-last column of Tables 2 and 3 we also give the estimates for the individual studies after “shrinkage” towards the overall mean through “borrowing strength” from the totality of studies. Note that these estimates have much tighter credible intervals than the original study estimates, since they are based on a combination of individual and overall study information.

In Table 2, the overall Bayesian posterior mean estimate for spousal studies (1.22) is slightly higher than that of the logit-based RE model (1.20), although they are very much within each other’s CI. For the spousal exposure studies we find $P\{\mu > 1\} = 0.9996$, significant at well above the 5% level with this data.

The values appear robust to some change in model choice. Under the Bayesian model, there were negligible changes to posterior distributions when input values for prior distributions were changed. Only when the degrees of freedom associated with the distributions of σ^2 and τ^2 or the entries in the matrix controlling the between studies variance were set at extreme and unreasonable values were there any real changes to posterior estimates. Other changes in prior specifications produced no effect at all.

The effect of the different methods is perhaps more noticable in the analysis of workplace exposure than spousal exposure data.

In using the RE approach to obtain an overall estimate of relative risk for females in workplace studies, the estimate of between-study heterogeneity is $\hat{\tau}^2 = 0$ so the RE and fixed effects models coincide, indicating that all studies may have a common true relative risk. (For males, in contrast, as shown in [4], there is indeed detectable between-study heterogeneity.)

The RE estimate of the overall mean μ is then 1.12, although this is not significantly raised above unity. The overall (posterior) mean, based on the Bayesian model which does allow for between-study heterogeneity, is estimated to be 1.10. Both of these values are again well within each other’s CIs: in this sense this is an insignificant difference. However, the *excess risk* which is often fed into calculations of attributable risk is some 20% higher for the RE model.

The value of 1.10 is larger than the value of 1.07 calculated in [4], where it is also shown using a simplified formulation due to Carlin [10] that the posterior mean is estimated by 1.04. However, here we have added the recent result reported by Fontham *et al* [20]: using exactly the same data as in [4] we get a value exactly in accord with the DuMouchel approximation.

These are slight discrepancies. Nonetheless they illustrate that the choice of model can play a serious role, for in using estimates of the overall effect, especially for values estimated as near 1.0 such as this one, a 10%-20% discrepancy in the excess risk makes a considerable difference in interpretation.

Finally, we again quantify the posterior probability that $\mu > 1.0$: for the workplace studies we find $P\{\mu > 1\} = 0.83$, so that this is not significant even at the 10% level with this data.

3.2 Analysis using adjusted relative risks

In any meta-analysis of published studies the role of covariates, either in the design or by adjustment in the analysis, raises questions of comparability. Unadjusted RR's may be quite appropriate for case-control studies since adjustments can often be assumed to have been made by matching and similar techniques in the design stage. Also, restricting the meta-analysis to data which have been adjusted seems extreme for case-control studies, since many studies do not report such adjusted values and would have to be ignored. For cohort studies, conversely, adjusted RR's are probably more appropriate; but of course, as with the case control studies, adjustments for the same covariates are not made on a common basis across studies.

Our goal is fortunately rather more limited than making a final choice between the two types of estimates. We merely wish to see whether this problem of principle actually makes any practical difference in these particular meta-analyses.

From each of Tables 2 and 3, we provide meta-analyses of the adjusted RR estimates where provided by individual studies, and unadjusted estimates for other studies.

For workplace exposure the following two groups of datasets were considered separately:

- (a) only those studies of females which provided unadjusted RR estimates: these results can thus be directly compared with those of Tables 2 and 3;
- (b) all studies, combining both males and females, giving a result directly comparable with that of Lee [48, Table 5].

Results are presented in Table 4. For dataset (a), under either model the combined point estimate for the females exposed to workplace smoking is 1.10-1.11. The dataset (b) includes both genders and indicates that the male studies are somewhat different in the sense that now $\hat{\tau}^2 = 0.017$. The overall estimate of 1.07-1.08 is only marginally higher than that of Lee [48, Table 5] as we should expect since they differ only by two studies.

The Bayesian methodology also enables us to assert that the posterior probability that the overall underlying relative risk is greater than 1.0 is 0.83-0.84 in both these cases.

TABLE 4: Meta-analyses of Results (Adjusted where Available) of Workplace Exposure Studies

Study Type	Bayesian Model	RE Model	
	RR (95% CI)	RR (95% CI)	$\hat{\tau}^2$
(a) Females (9 Studies)	1.10 (0.89-1.32)	1.11 (0.96-1.29)	0.005
(b) Combined (14 Studies)	1.08 (0.92-1.26)	1.07 (0.93-1.24)	0.017

For exposure to spousal smoking, we consider a different approach to the adjusted results, and indicate the effect of combining the case-control and cohort studies. Under a fixed effects model this is not advisable due to the inherent differences in the methodology. Here we are able to take that into account.

We analyse again the totality of studies, consisting of adjusted RR's where given as in Table 2, and unadjusted RR's for other studies, thus using the maximum number of 35 case control and 4 cohort studies for females. Table 5 gives the results of analysing this dataset. Again we note that the Bayesian and the RE models give very similar answers. The inhomogeneity in the studies is supported by a value of $\hat{\tau}^2 = 0.052$, although the inclusion of the cohort studies in this case actually decreases $\hat{\tau}^2$ slightly.

TABLE 5: Meta-Analysis of Results (Adjusted where Available) of Spousal Exposure Studies

Study Type	Bayesian Model	RE Model	
	RR (95% CI)	RR (95% CI)	$\hat{\tau}^2$
Case Control	1.22 (1.07-1.39)	1.22 (1.06-1.41)	0.061
Cohort	1.33 (1.03-1.78)	1.29 (1.02-1.64)	0
All	1.23 (1.09-1.39)	1.23 (1.08-1.39)	0.052

3.3 Choosing subgroups of studies

Ensuring comparability can entail close examination of the data to identify appropriate subsets of studies for combination. As noted in the EPA Report [18], for this particular meta-analysis it may be sensible to consider the effect of grouping studies by geographic region, especially given the rather inexplicit nature of "exposure to ETS" and the way in which it might vary in different cultures. To illustrate the effect of geographic location we provide in Table 6 meta-analyses of two subgroups of studies of spousal smoking which may *a priori* be considered internally more homogeneous: Asian populations only (China, Japan, Hong Kong, Taiwan), comprising 15 case control and one cohort study; and US studies only, comprising 11 case control and two cohort studies.

Resulting overall unadjusted relative risks with logit variance estimates are again compared under both frequentist and Bayesian approaches. It can be seen that there are considerable differences between the two country groups with respect to both overall estimate and between-study heterogeneity (and that again there is 10%-15% difference between using RE and Bayesian methods).

The relative risk estimate is significantly increased above 1.0 at the 5% level for the Asian studies, but for the U.S. studies we calculate the probability of the relative risk being above 1.0 as 0.92.

TABLE 6: Meta-analysis of Asian and U.S. subgroups of studies of spousal exposure

Study Type	Bayesian Model	RE Model	
	RR (95% CI)	RR (95% CI)	$\hat{\tau}^2$
Asian Studies	1.25 (1.03-1.50)	1.25 (1.02-1.52)	0.067
US Studies	1.13 (0.95-1.34)	1.11 (0.98-1.26)	0.003

One implication of the different RR's is that extrapolation of overall results to individual studies and from one country group to another may not be appropriate. It certainly highlights a need for further investigation of these differences, perhaps through a closer exploration of covariates or possible biases. Some recognised covariates in the association between lung cancer and exposure to ETS include diet and socioeconomic status. Possible biases include different underlying rates of lung cancer and misclassification of active smoking.

There are many other breakups of the data that could be accomplished by the methods used here. For example, there has been considerable recent interest [54, 76, 18, 57] in accounting for the differing quality of studies in a meta-analysis. The EPA Report [18] groups studies into four tiers based on a qualitative score. This is intended to take into account various design aspects and susceptibilities to common sources of bias and misclassification which may impact on the observed results. In Table 7 of [56] we show the effect of using only the studies in Tiers 1-2 and 1-3 in a classical model in order to avoid bias which may be inherent in the lower quality studies. For each gender, using only the "good quality" studies appears to give a more homogeneous collection, as indicated by the decreasing $\hat{\tau}^2$ values with increasing "quality" although as also commented in [18], the "good quality" Tier 1-2 studies exhibited exactly the same point estimate of combined RR for females as does the complete set of studies.

4 Discussion

4.1 Effect of different statistical models

In this paper we have compared a frequentist (random-effects) and a Bayesian approach to meta-analysis of epidemiological studies, and implemented the approaches for studies of the association of lung cancer with both workplace and spousal exposure to ETS. In general one would expect that the Bayesian methods would give a more explicit overall picture of the effect of variability in a collection of studies. The use of Bayesian methods was facilitated by the use of MCMC algorithms, which allow for more flexibility in the formulation of prior information and models, and for a wider range of inferences and comparisons through simulation.

The workplace studies were analysed using a similar Bayesian model in [4], but there approximations were used to enable direct analytical solutions; our results confirm the approximate analysis there. The spousal exposure studies have not been analysed previously by Bayesian methods, although

the results are very similar to those for RE models of the same data: the overall relative risk is 1.20-1.22 with a CI of (1.07,1.35) on either method.

For workplace exposure, there is also a small but proportionally more noticeable difference between the two estimates: the overall mean estimate under the RE model is [1.12 (0.93,1.28)] (essentially as noted by Lee [48]), higher than that obtained under the Bayesian formulation [1.10 (0.90,1.32)]. Use of the RE model thus gives a point estimate of excess risk which is 20% higher than for the Bayesian method. This effect of choice of approach is rather more marked before adding the recent Fontham study [20]: it was shown in [4] that a variation of almost 70% in excess risk (from 0.07 to 0.12) resulted by changing the model used.

Such differences are almost certainly due to the different estimates of between-study variation. In this case we have an estimate of $\tau^2 = 0$ in the classical model but the *assumed* non-zero across-study variability in the Bayes methods automatically down-weights individual study estimates with relatively small variances, so that they do not dominate overall estimates as much.

Using MCMC methods, we are able to estimate the probability that the overall relative risk is greater than unity: for the workplace studies this is $P\{\mu > 1.0\} = 0.83$.

Taking into account variation between studies through the adoption of a Bayesian model is of considerable importance. Most previous meta-analyses of ETS spousal exposure studies have used simple fixed effects models (see [81, 59, 74, 18, 17]) and substantial conclusions have been based on them. We have shown that there is considerable indication of heterogeneity between studies and that moving to a Bayesian approach can give a more detailed approach to analysis of such data.

4.2 Use of appropriate studies

One of the outcomes of this analysis is the ability to determine whether one can use spousal data to evaluate relative risks in the workplace. It is clear that this is a dangerous practice: the overall spousal relative risk estimated here is around 1.20. The overall workplace relative risk is estimated to be around 1.11. Thus the observed excess risk in one environment is only some 50% of the excess estimated in the other environment.

In these analyses above we have deliberately ignored any need to adjust for covariates in the population. If the reported (and sometimes adjusted) data for spousal exposure are used, then there is little indication of any change in the estimate of relative risk. Similarly, the difficult choice between unadjusted and adjusted data seems immaterial in the case of workplace exposure, where adjustments up and down also seem to cancel.

In the adjusted data there continues to be indication of substantial between-study variation. This warrants special attention in a meta-analysis, especially in applying overall results to subpopulations and individuals.

In this example, the choice of subpopulations to be combined had very obvious effects. For example, the overall estimates vary considerably within different geographic areas: it is difficult not to conclude from Table 5 that in combining over the studies conducted in Asia and the US, one may well be ignoring meaningful differences. Again the excess risk in the current Asian studies is almost double that of the excess risk in the U.S.

It is clear from these observations that quotation of point estimates alone, and their use as a basis for decisions, seems very unwise. A much more acceptable practice is to report the corresponding confidence intervals, taking explicit account of the between-study variation and, if necessary, reducing this variation through the adoption of more homogeneous subpopulations or expansion of the model to include relevant covariates.

Detection of heterogeneity in a meta-analysis context becomes, of course, only the first step: it is then revealing for the analyst to investigate further the source of such variation, as we have tried to do in isolating study groups.

4.3 Publication and other biases

Although we have not concentrated on it in this paper, the sensitivity of the combined point estimate and confidence interval to possible biases should be acknowledged and, where possible, explicitly taken into account in inferences based on the size and significance of the overall relative risk. In [56] we showed that there is substantial indication of publication bias in the set of spousal exposure studies. It is calculated in that paper that the possible impact of this bias is to reduce the combined RR from 1.20 to 1.12 (95% CI of (1.01, 1.24)) using the RE model. Following the same approach of excluding the “high outliers” for which there appear to be no matching “low outliers”, due perhaps to publication bias, we also get a combined RR of 1.11 (95% CI of (0.99, 1.24)) using the Bayesian analysis.

Thus as much as 45% of the observed excess risk could be due to publication bias.

We must also note that there is a further well-known bias in the spousal studies, due to misclassification of smokers as non-smokers, which spuriously elevates the observed relative risk in this context. In using the results calculated here we should adjust both the overall point and interval estimates to allow for such a bias. Relevant methods were first developed in [81] and have been adopted widely [59, 17, 18, 44, 74]. The true extent of misclassification bias has been debated vigorously [44, 46, 45, 80, 79, 78, 18, 75]; we will not repeat the arguments here .

This systematic bias must be accounted for in a meta-analysis as in the analysis of a single study. Applying, say, the EPA Draft Review [17] estimate of an overall “spurious excess risk” of 0.12 due to this misclassification of smokers as nonsmokers to the estimates based on the RE model in Table 4, we would derive an estimated combined risk for all females reduced from 1.20 to 1.08, and the associated CI would also fall to around (0.96, 1.24). Based on the final EPA Report [18] in which adjustment for misclassification is made to individual studies prior to meta-analysis, correction of the combined estimate for this bias would not be as severe.

It is interesting to note that an adjustment of around 0.10 would bring the spousal exposure relative risk to around the same level as that of the relative risk currently estimated for workplace exposure.

4.4 Overall comments

Meta-analysis is often used simply to increase statistical power: that is, in effect to narrow the confidence limits around an estimate of effect, even if results are fairly consistent and clearcut. It can be used to greater advantage, however, in situations for which individual outcomes are difficult

to interpret—and it has become increasingly popular for this purpose [61]—or when excess risks are small or not significant in each study alone. It is important to realise that the impact of choice of method, selection of studies to be combined, and evaluation of bias, can be substantial, as we have seen in this one example.

There are many problems with meta-analysis as a tool. We try to combine studies with different designs, of different quality, and from different areas. There may be consistent biases, either upward or downward, and these will flow from individual studies to an overall assessment. We have noted that in this data-set (and even more in the earlier analysis before the Fontham study [20] was released) the method of analysis may inflate the excess risk estimate by 20%; publication bias may account for almost 50% of the observed excess risk; addition of studies from other geographic areas may raise the U.S. excess risk by more than 100%, and so their appropriateness needs careful consideration; misclassification of subjects, such as the use of active instead of non-smoking subjects in studies of exposure to spousal ETS (which is well-documented [75]) may account for 50% of the observed excess risk.

But with all of these problems, meta-analysis is an increasingly common and useful practice and one that needs to be improved where possible.

In this paper we approach one particular issue (study inhomogeneity) by trying to capture study differences in an expanded hierarchical model, and summarised a number of others from the earlier analyses in [56, 4].

These are all shown to make small but important differences in point estimates: on this data-set the overall excess risk for workplace exposure is shown to be increased slightly if the RE model is used rather than the Bayesian model, and in previous analyses of smaller data sets the increase has been considerably higher. Since none of these values are significantly different from each other, or from a null effect, this may still, of course, be a product of random variation in the data: but it does indicate that there can be considerable danger in ignoring the heterogeneity between studies.

Even low observed excess risks, if used to generate attributable risk figures as in [62], can appear important: if they are inaccurate by 10%, or 20%, or 50%, as we have shown may be the case from choice of models, or of data, or because of publication bias, then very much more caution needs to be shown in using them than appears commonly to be the case.

Appendix: Bayesian and Classical Hierarchical Models for Meta-Analysis in Epidemiology

In order to implement the meta-analysis methods, we need to estimate the individual parameters θ_i introduced in Section 2.4 and to calculate corresponding variances for these estimates. In our analysis we have used

- (a) Fisher’s exact method [5, p.124], which gives a point estimate and a non-parametric confidence interval (CI) but no variance estimate for RR_i ;
- (b) The logit method [5, pp. 129–130], which gives a point estimate Y_i and approximate variance W_i^{-1} , with a corresponding confidence interval based on an assumption of normality which is known to be reasonable, at least for large sample sizes.

A frequentist or classical random effects (RE) model (of which a fixed effects model is a particular case) for meta-analysis is an appropriate starting point to describe the classical and Bayesian methods of meta-analysis. We consider the simple formulation

$$\begin{aligned} Y &= \theta + e \\ \theta &= X\mu + \varepsilon, \end{aligned} \tag{1}$$

in which $Y = (Y_1, \dots, Y_k)'$ are the observed log relative risks for each study, $\theta = (\theta_1, \dots, \theta_k)'$ are the corresponding true log relative risks, $e = (e_1, \dots, e_k)'$ and $\varepsilon = (\varepsilon_1, \dots, \varepsilon_k)'$ are random errors, X is a $k \times p$ design matrix, and μ is a $p \times 1$ -vector of parameters. We shall take X to be the $k \times 1$ -vector of 1's, and so μ to be a scalar parameter, representing the true underlying log RR over all studies.

In (1) Y_i are used since they can be shown to be asymptotically Normal $Y \sim N(\theta, W^{-1})$, with $W^{-1} = \text{diag}(W_1^{-1}, \dots, W_k^{-1})$; accordingly we assume that $\theta \sim N(X\mu, \tau^2 I)$, and hence that the e_i are independent $N(0, \sigma_i^2)$ random variables, the ε_i are independent $N(0, \tau^2)$ random variables, and the e_i and ε_i are mutually independent.

A classical approach to meta-analysis using this model is widely used [58, 72]. Here μ, σ^2 and τ^2 are considered fixed parameters and τ^2 is estimated most commonly through an approximation proposed by DerSimonian and Laird (cf [58]).

DuMouchel [16, 15] describes a general hierarchical Bayesian scheme, and this is the model we use here. This allows considerable flexibility in application, and can be viewed as a Bayesian generalization of the frequentist RE model. DuMouchel makes the following further distributional assumptions:

$$\begin{aligned} Y|\theta, \sigma &\sim N(\theta, \sigma^2 \mathbf{C}), \\ \sigma^{-2} &\sim \chi^2(\text{df}_\sigma)/\text{df}_\sigma, \end{aligned} \tag{2}$$

and

$$\begin{aligned} \theta|\mu, \tau &\sim N(X\mu, \tau^2 \mathbf{V}), \\ \mu|\tau &\sim N(0, D \rightarrow \infty), \\ \tau^{-2} &\sim \chi^2(\text{df}_\tau)/\text{df}_\tau, \end{aligned} \tag{3}$$

where \mathbf{C} and \mathbf{V} are $k \times k$ observed and prior variance-covariance matrices respectively, and the degrees of freedom df_σ and df_τ indicate how well \mathbf{C} and \mathbf{V} , respectively, are known. (Note that for consistency with notation in (1) we have interchanged σ^2 and τ^2 relative to [16], in accordance with notation in [15].)

Implementation of this Bayesian model requires several initial specifications. We need to specify the matrix \mathbf{C} describing within-study variability; here, since the studies are assumed to be independent, we take \mathbf{C} as a diagonal matrix with individual logit estimates of variance of Y_i on the diagonal. The specification of df_σ reflects our faith in these estimates in \mathbf{C} . The average number of exposed cases from each study was used as a conservative estimate of the degrees of freedom df_σ , with the average taken over those studies for which 2×2 tables are given.

Specification of the matrix \mathbf{V} and corresponding df_τ similarly represents our prior beliefs. For most of the analyses the diagonal elements of \mathbf{V} were taken to be the interstudy variability found in the corresponding frequentist analysis. In some cases this choice did not seem to allow the data to be modelled appropriately. In particular when the frequentist model reduced to a fixed effects

model the initial choice of the diagonal elements of \mathbf{V} were given a very small value (0.001), and then the individual study estimates were found to be almost identical, the posterior mean of τ^2 was large and the posterior distribution of τ^2 was very severely skewed to the right. To deal with this problem, different values were tried for the diagonal elements of \mathbf{V} until it was found that the posterior mean of τ^2 was nearly one. This problem appears to arise because the DuMouchel model does not permit the prior distribution of $\tau^2 V$ to be sufficiently uninformative, so the data can not drive the posterior distribution to satisfactorily match the data. Future work is planned to consider a more appropriate way to model the variability of $\theta|\mu, \tau$.

As noted by DuMouchel [16, p. 515], the particular prior distributions given here are chosen for convenience, so that the posterior distribution of θ given Y is a mixture of multivariate Student- t distributions. For computational convenience, however, he suggests using a multivariate normal approximation to the posterior, which can then be described through the posterior mean and covariance matrices. One of the advances in the present paper is the evaluation of this model using MCMC methods, which avoid the need for this approximation and also allow us to assess how much the results actually depend on some of these assumptions. In particular, the Gibbs algorithm was used in these analyses through the software package BUGS.

A more restricted version of this hierarchical model is explored by Carlin [10], who more closely follows the RE version by taking \mathbf{V} to be the the $k \times k$ identity matrix, \mathbf{C} to be a diagonal matrix with the corresponding diagonal entries the (assumed known) variances of the individual observations Y_i and $\sigma^2 = 1$, thus omitting the non-degenerate prior on σ^{-2} . Both μ and τ^2 are still unknown hyperparameters representing, as before, overall mean and between-study variance, respectively.

In [4], both the DuMouchel method and the Carlin method are applied to the workplace exposure studies of Table 3. Both Carlin and DuMouchel suggest ways to examine the sensitivity of their respective methods to the assumptions made, and it is desirable to investigate the dependence of the posterior estimates of μ and θ on the specifications of df_τ and df_σ . Because of computational restrictions this was not done in [4], but the MCMC methods allow us to do this. Although sensitivity was not addressed for all of the analyses of this paper, the sensitivity analyses we conducted indicate that the only initial specifications which have any effect on the estimates are those for df_τ and df_σ , and for meta-analyses such as these where the data-set is large, the changes were essentially negligible.

Acknowledgements

The authors thank Professor P. Mielke for programs to do Fisher's exact methods, and the MRC Biostatistics Unit at the University of Cambridge for supplying a prototype of the BUGS software for the Bayesian analysis. This work was carried out with partial support from three tobacco companies: the conclusions here are however entirely those of the authors and should not be otherwise ascribed.

References

- [1] S. Akiba, H. Kato, and Blot W. J. Exposure to ETS and lung cancer among Japanese women. *Cancer Research*, 46:4808–4807, 1986.
- [2] J. Besag, P. Green, D. Higdon, and K. Mengersen. Bayesian computation and stochastic systems. *Statistical Science*. (To appear).
- [3] J. Besag, D. Higdon, and K Mengersen. Meta-analysis via MCMC. (In preparation).
- [4] B. Biggerstaff, K.L. Mengersen, and R.L. Tweedie. Passive smoking in the workplace: Classical and Bayesian meta-analyses. *Int. Arch. Occupational and Environmental Health*, page (to appear), 1994.
- [5] N.E. Breslow and N.E. Day. *Statistical Methods in Cancer Research*, volume II. IARC, Lyon, 1986. IARC Scientific Publication 82.
- [6] R.C. Brownson, M.C.R. Alavanja, E.T. Hock, and T.S. Loy. Passive smoking and lung cancer in non-smoking women. *Am. J. Public Health*, 82:1525–1530, 1992.
- [7] R.C. Brownson, J.S. Reif, T.J. Keefe, S.W. Ferguson, and J.A. Pritzl. Risk factors for adenocarcinoma of the lung. *Am. J. Epidemiol*, 125:25–34, 1987.
- [8] P.A. Buffler, L.W. Pickle, C.J. Mason, and C. Constant. The causes of lung cancer in Texas. In *Lung Cancer: Causes and Prevention*, pages 83–89, New York, 1984. Verlag Chemie International. eds M. Mizell and P. Correa.
- [9] T.L. Butler. *The relationship of passive smoking to various health outcomes among Seventh-Day Adventists in California*. PhD thesis, UCLA, Los Angeles, USA, 1988.
- [10] John B. Carlin. Meta-analysis for 2×2 tables: a Bayesian approach. *Statistics in Medicine*, 11:141–158, 1992.
- [11] T.C. Chalmers. Problems induced by meta-analysis. *Statistics in Medicine*, 10:971–980, 1991.
- [12] W.C. Chan and S.C. Fung. Lung cancer in non-smokers in Hong Kong. In E. Grundmann, editor, *Cancer Campaign, Vol.6 Cancer Epidemiology*, pages 199–202. Stuttgart, Fischer, 1982.
- [13] P. Correa, L.W. Pickle, E.L. Fontham, Y. Lin, and W. Haenszel. Exposure to ETS and lung cancer. *The Lancet*, pages 595–597, 1983.
- [14] Y.X. Du, Q. Cha, Y.Z. Chen, and J.M. Wu. Exposure to environmental tobacco smoke and female lung cancer in Guangzhou, China. In *Proceedings of Indoor Air 1993*, volume 1, pages 511–516, 1993.
- [15] W. DuMouchel. Predictive cross-validation for hierarchical Bayesian meta-analysis. In *Bayesian Statistics 5*. Oxford University Press, To appear. Eds Berger, J.O., Bernardo, J.M., Dawid, A.P., and Smith, A.F.M.
- [16] William DuMouchel. Bayesian meta-analysis. In *Statistical Methods for Pharmacology*, pages 509–529. Marcel Dekker, New York, 1990. ed. D. Berry.

- [17] EPA Draft Review. *Health Effects of Passive Smoking: Assessment of Lung Cancer in Adults and Respiratory Disorders in Children*. National Academy Press, United States EPA, Washington, 1990.
- [18] EPA Review. *Health Effects of Passive Smoking: Assessment of Lung Cancer in Adults and Respiratory Disorders in Children*. National Academy Press, United States EPA, Washington, 1992.
- [19] D.T. Felson. Bias in meta-analytic research. *J. Clin. Epidemiol.*, 45:885–892, 1992.
- [20] E.T.H. Fontham, P. Correa, P. Reynolds, A. Wu-Williams, P.A. Buffler, R.S. Greenberg, V.W. Chen, T. Alterman, P. Boyd, D.F. Austin, and J. Liff. Environmental tobacco smoke and lung cancer in non-smoking women. *JAMA*, 271:1752–1759, 1994.
- [21] E.T.H. Fontham, P. Correa, A. Wu-Williams, P. Reynolds, A.S. Greenberg, P.A. Buffler, V.W. Chen, P. Boyd, T. Alterman, D.F. Austin, J. Liff, and S.D. Greenberg. Lung-cancer in non-smoking women: A multicenter case-control study. *Cancer Epidemiology, Biomarkers and Prevention*, 1:35–43, 1991.
- [22] Y-T. Gao, W.J. Blot, W. Zheng, A.G. Ershow, C.W. Hsu, L.I. Levin, R. Zhang, and J.F. Fraumeni. Lung cancer among Chinese women. *Int. J. Cancer*, 40:604–609, 1987.
- [23] L. Garfinkel. Time trends in lung cancer mortality among nonsmokers and a note on exposure to ETS. *JNCI*, 66:1061–1066, 1981.
- [24] L. Garfinkel, O. Auerbach, and L. Joubert. Involuntary smoking and lung cancer: A case study. *JNCI*, 75:463–469, 1985.
- [25] G-Y. Geng, Z.H. Liang, A.Y. Zhang, and G.L. Wu. On the relationship between smoking and female lung cancer. In M. Aoki, S. Hisamichi, and S. Tominaga, editors, *Smoking and Health 1987, Proceedings of the 6th World Conference on Smoking and Health*, Tokyo, 1988.
- [26] L.P. Ger, W.-L. Hsu, K.-T. Chen, and C.-J. Chen. Factors of lung cancer by histological category in Taiwan. *Anticancer Research*, 13:1491–1500, 1993.
- [27] L.V. Hedges. Modeling publication selection effects in meta-analysis. *Statistical Science*, 7:227–236, 1992.
- [28] L.V. Hedges and I. Olkin. *Statistical Methods for Meta-analysis*. Academic Press, 1985.
- [29] T. Hirayama. Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. *British Medical Journal*, 282:183–185, 1981.
- [30] T. Hirayama. Lung cancer in Japan: Effects of nutrition and exposure to ETS. In *Lung Cancer: Causes and Preventions*, pages 175–195. Verlag Chemie Weinheim, 1984.
- [31] Larry C. Holcomb. Impact of environmental tobacco smoke on airline cabin air quality. *Environ. Technol. Letters*, 9:509–514, 1988.
- [32] D.J. Hole, C.R. Gillis, C. Chopra, and V.M. Hawthorne. Passive smoking and cardiorespiratory health in a general population in the west of Scotland. *Br. Med. J.*, 299:423–427, 1989.

- [33] C.G. Humble, J.M. Samet, and D.R. Pathak. Marriage to a smoker and lung cancer risk. *Am. J. Public Health*, 77:598–602, 1987.
- [34] R. Inoue and T. Hirayama. Passive smoking and lung cancer in women. In M. Aoki, S. Hisamichi, and S. Tominaga, editors, *Smoking and Health 1987, Proceedings of the 6th World Conference on Smoking and Health*, Tokyo, 1988.
- [35] D.T. Janerich, W.D. Thompson, L.R. Varela, P. Greenwald, S. Chorost, C. Tucci, M.B. Zaman, M.R. Melamed, M. Kiely, and M.F. McKneally. Lung cancer and exposure to tobacco smoke in the household. *New England J. Med.*, 323:632–636, 1990.
- [36] K.-H. Joeckel. Passive smoking—evaluation of the epidemiological findings (in German). VDI Reports 888, Association of German Engineers, 1991. Commission on Air Pollution of the VDI and DIN, Carcinogenic Substances in the Environment.
- [37] G.C. Kabat. Epidemiologic studies of the relationship between passive smoking and lung cancer. In *Proc 1990 Toxicology Forum*, pages 187–199, 1990.
- [38] G.C. Kabat and E.L. Wynder. Lung cancer in nonsmokers. *Cancer*, 53(5):1214–1221, 1984.
- [39] A. Kalandidi, K. Katsouyanni, N. Voropoulou, G. Bastas, R. Saracci, and D. Trichopoulos. Passive smoking and diet in the aetiology of lung cancer among non-smokers. *Cancer Causes and Control*, 1:15–21, 1990.
- [40] L.C. Koo, J.H. Ho, D. Saw, and C. Ho. Measurements of passive smoking and estimates of lung cancer risk among non-smoking Chinese females. *Int. J. Cancer*, 39:162–169, 1987.
- [41] T. H. Lam and K.K. Cheng. Passive smoking is a risk factor for lung cancer in never smoking women in Hong Kong. In *Smoking and Health 1987, Proceedings of the 6th World Conference on Smoking and Health*, Tokyo, 1988. Eds. Aoki, M. and Hisamichi, S. and Tominaga, S.
- [42] T.H. Lam, I.T.M. Kung, C.M. Wong, W.K. Lam, J.W.L. Kleevens, D. Saw, C. Hsu, S. Seniveratne, S.Y. Lam, K.K. Lo, and W.C. Chan. Smoking, passive smoking and histological types in lung cancer in Hong Kong Chinese women. *Br. J. Cancer*, 6:673–678, 1987.
- [43] W.K. Lam. *A Clinical and Epidemiological Study of Carcinoma of Lung in Hong Kong*. PhD thesis, M.D. Thesis submitted to University of Hong Kong, 1985.
- [44] P.N. Lee. *Misclassification of Smoking Habits and Passive Smoking: a Review of the Evidence*. Springer-Verlag, Berlin, 1988.
- [45] P.N. Lee. Lung cancer and passive smoking (continued): Letter to the Editor. *British J. Cancer*, 64:200, 1991.
- [46] P.N. Lee. Lung cancer and passive smoking: Letter to the Editor. *British J. Cancer*, 63:161–162, 1991.
- [47] P.N. Lee. *Environmental Tobacco Smoke and Mortality*. Karger, Basel, 1992.

- [48] P.N. Lee. An assessment of the epidemiological evidence relating lung cancer risk in never smokers to environmental tobacco smoke exposure. In *Environmental Tobacco Smoke*, pages 28–70. Springer-Verlag, New York, 1993. ed. H. Kasuga.
- [49] P.N. Lee. Lung cancer and ETS: Is the epidemiologic evidence conclusive? In *Joint Statistical Meetings of the ASA, IBS, SSC and IMS*, August 1994. Toronto, Canada.
- [50] P.N. Lee, J. Chamberlain, and M.R. Alderson. Relationship of passive smoking to risk of lung cancer and other smoking-associated diseases. *Br. J. Cancer*, 54:97–105, 1986.
- [51] R.J. Light and D.B. Pillemer. *Summing Up: the Science of Reviewing Research*. Harvard Univ. Press, 1984.
- [52] Q. Liu, A.J. Sasco, E. Riboli, and M.X. Hu. Indoor air pollution and lung cancer in Guangzhou, People’s Republic of China. *Amer. J. Epidemiol.*, 137:145–154, 1993.
- [53] Z. Liu, X. He, and R.S. Chapman. Smoking and other risk factors for lung cancer in Xuanwei, China. *Int. J. Epidemiology*, 20:26–31, 1991.
- [54] Matthew P. Longnecker, Jesse A. Berlin, Michele J. Orza, and Thomas C. Chalmers. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA*, 260:652–656, 1988.
- [55] K. Mengersen and J. Besag. Ranking and selection using MCMC. In *Proceedings of the 3rd Schwerin Conference in Mathematical Statistics, Germany*, September 1993.
- [56] K.L. Mengersen, R.L. Tweedie, and B.J. Biggerstaff. The impact of method choice in meta-analysis. *Australian J. Statistics*. (Accepted subject to review).
- [57] Frederick Mosteller and Thomas Chalmers. Some progress and problems in meta-analysis of clinical trials. *Statistical Science*, 7:227–236, 1992.
- [58] NRC Committee on Applied and Theoretical Statistics. *Combining Information: Statistical Issues and Opportunities for Research*. National Academy Press, Washington, 1992.
- [59] NRC Committee on Passive Smoking. *Environmental Tobacco Smoke. Measuring Exposures and Assessing Health Effects*. National Academy Press, Washington, 1986.
- [60] Guy B. Oldaker III and Fred C. Conrad, Jr. Estimation of effect of environmental tobacco smoke on air quality within passenger cabins of commercial aircraft. *Environ. Sci. Technol.*, 21, No. 10:994–999, 1987.
- [61] I. Olkin. Meta-analysis: methods for combining independent studies. *Statistical Science*, 7:226, 1992.
- [62] OSHA. Proposed Rule on Indoor Air Quality. *Federal Register*, 59 (65):15968–16039, 1994.
- [63] G. Pershagen, Z. Hrubec, and C. Svensson. Passive smoking and lung cancer in Swedish women. *Am. J. Epidem.*, 125:17–24, 1987.

- [64] Tosiya Sato. Confidence limits for the common odds ratio based on the asymptotic distribution of the Mantel–Haenszel estimator. *Biometrics*, 46:71–80, 1990.
- [65] H. Shimizu, M. Morishita, K. Mizuno, T. Masuda, Y. Ogura, M. Santo, M. Nishimura, K. Kunishima, K. Karasawa, K. Nishiwaki, M. Yamamoto, S. Hismichi, and S. Tominaga. A case-control study of lung cancer in nonsmoking women. *Tohoku J. Exp. Med.*, 154B:389–397, 1988.
- [66] Michael Siegel. Involuntary smoking in the restaurant workplace. *JAMA*, 270, No. 4:490–493, 1993.
- [67] T. Sobue, R. Suzuki, N. Nakayama, and 14 others. Passive smoking among non-smoking women and the relationship between indoor air pollution and lung cancer incidence—results of a multicentre case controlled study. *Gan to Rinsho*, 36:329–332, 1990.
- [68] D. Spiegelhalter, A. Thomas, N. Best, and W. Gilks. BUGS Manual 0.30. Technical report, MRC Biostatistics Unit, Cambridge, 1994.
- [69] H. G. Stockwell, E. C. Candelora, A. W. Armstrong, and P. A. Pinkham. Environmental tobacco smoke and lung cancer in never smoking women. In *Proc. 24th Ann Meeting Soc. Epidemiologic Research*, 1991.
- [70] C. Svensson, G. Pershagen, and J. Klominek. Smoking and passive smoking in relation to lung cancer in women. *Acta Oncologica*, 28:623–629, 1989.
- [71] S.G. Thompson and S.J. Pocock. Can meta-analyses be trusted? *Lancet*, 338:1127–1130, 1991.
- [72] S.G. Thomson. Controversies in meta-analysis: the case of the trials of serum cholesterol reduction. *Statistical Methods in Medical Research*, 2:173–192, 1993.
- [73] D. Trichopoulos, A. Kalandidi, and L. Sparros. Lung cancer and exposure to ETS. Conclusion of Greek study. *The Lancet*, ii:677–678, 1983.
- [74] R.L. Tweedie and K.L. Mengersen. Lung cancer and passive smoking: Reconciling the biochemical and epidemiological approaches. *Br. J. Cancer*, 66:700–705, 1992.
- [75] R.L. Tweedie, K.L. Mengersen, and J.A. Eccleston. Garbage in, garbage out: Can statisticians quantify the effects of poor data? *Chance*, 7(2):20–27, 1994.
- [76] K. Uberla. Lung cancer from passive smoking: Hypothesis or convincing evidence? *Int. Arch. Occup. Environ. Health*, 59:421–437, 1987.
- [77] L.R. Varela. *Assessment of the Association Between Passive Smoking and Lung Cancer*. PhD thesis, Yale University, 1987.
- [78] N. J. Wald, H. S. Cuckle, K. Nanchahal, and S. G. Thompson. Response to letter from Dr Lee: Letter to the Editor. *British J. Cancer*, 64:201, 1991.
- [79] N. J. Wald, K. Nanchahal, H. S. Cuckle, and S. G. Thompson. Response to the letter from Dr P. Lee: Letter to the Editor. *British J. Cancer*, 63:163, 1991.

- [80] N.J. Wald, K. Nanchahal, H.S. Cuckle, and S.G. Thompson. Lung cancer and passive smoking: Letter to the Editor. *British J. Cancer*, 61:337, 1990.
- [81] N.J. Wald, K. Nanchahal, S.G. Thompson, and H.S. Cuckle. Does breathing other people's tobacco smoke cause lung cancer? *Br. Med. J.*, 293:1217–1222, 1986.
- [82] F.-L. Wang, E.J. Love, N. Liu, and X.-D. Dai. Indoor air pollution and lung cancer in Guangzhou, People's Republic of China. *Int. J. Epidemiol.*, 23:223–230, 1994.
- [83] R.L. Wolpert and W.J. Warren-Hicks. Bayesian hierarchical logistic models for combining field and laboratory survival data. In *Bayesian Statistics 4*, pages 525–546. Oxford University Press, 1992. Eds Bernardo, J.M., Berger, J.O., Dawid, A.P., and Smith, A.F.M.
- [84] A.H. Wu, B.E. Henderson, M.C. Pike, , and M.C. Yu. Smoking and other risk factors for lung cancer in women. *JNCI*, 74:747–751, 1985.
- [85] A.H. Wu-Williams, X.D. Dai, W. Blot, Z.Y. Xu, X.W. Sun, H.P. Xiao, B.J. Stone, S.F. Yu, Y.P. Feng, A.G. Ershow, J. Sun, J.F. Jr. Fraumentti, and B.E. Henderson. Lung cancer among women in north-east China. *Br. J. Cancer*, 62:982–987, 1990.
- [86] S. Yusuf, R. Peto, J. Lewis, R. Collins, and P. Sleight. Beta-blockade during and after myocardial infarction: An overview of the randomized trials. *Prog. Cardiovasc Dis.*, 27:335–371, 1985.
- [87] R. G. Ziegler, T. J. Mason, A. Stemhagen, et al. Dietary carotene and vitamin A and risk of lung cancer among white men in New Jersey. *JNCI*, 73:1429–1435, 1984.

R.L. Tweedie, PhD, DSc
 D.J. Scott, PhD
 B.J. Biggerstaff, MS
 K.L. Mengersen, PhD

Department of Statistics,
 Colorado State University,
 Fort Collins, CO 80523, USA.