STATISTICS IN MEDICINE
*Statist. Med.* 2003; **22**:1335–1353 (DOI: 10.1002/sim.1370)

# Combining multiple outcome measures in a meta-analysis: an application

Lidia R. Arends[1,*,†], Zoltán Vokó[1,2] and Theo Stijnen[1]

[1]*Department of Epidemiology & Biostatistics, Erasmus University Medical School, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands*
[2]*School of Public Health, Medical and Health Science Center, University of Debrecen, Kassai út 26/b, 4028 Debrecen, Hungary*

SUMMARY

In meta-analysis of clinical trials published in the medical literature it is customary to restrict oneself to standard univariate fixed or random effects models. If multiple endpoints are present, each endpoint is analysed separately. A few articles have been written in the statistical literature on multivariate methods for multiple outcome measures. However, these methods were not easy to apply in practice, because self-written programs had to be used, and the examples were only two-dimensional. In this paper we consider a meta-analysis on the effect on stroke-free survival of surgery compared to conservative treatment in patients with increased risk of stroke. Three summary measures per trial are available: short-term post-operative morbidity/mortality in the surgical group; long-term event rate in the surgical group, and the event rate in the conservative group. We analyse the three outcomes jointly with a general linear MIXED model, compare the results with the standard univariate approaches and discuss the many advantages of multivariate modelling. It turns out that the general linear MIXED model is a very convenient framework for multivariate meta-analysis. All analyses could be carried out in standard general linear MIXED model software. Copyright © 2003 John Wiley & Sons, Ltd.

## 1. INTRODUCTION

Meta-analysis of clinical trials aims to combine estimates of treatment effect across related studies. Usually no individual patient data are available and use is made of summary data extracted from published literature and reports. The data per trial are summarized by one or more outcome measure estimates along with their standard errors. In practice mostly the data are reduced to one outcome measure per study, for instance the treatment effect estimated by means of an odds ratio. The data are then analysed by standard methods, using either a (univariate) fixed effect or, as preferred by most statisticians, a (univariate) random effects model [1]. If the summary data are multi-dimensional, then the data analysis is usually restricted to a number of separate univariate analyses. Raudenbush *et al.* [2] showed how to

* Correspondence to: L. R. Arends, Institute of Epidemiology and Biostatistics, Erasmus University Medical School, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.
† E-mail: arends@epib.fgg.eur.nl

*Received July 2001*
*Accepted June 2002*

analyse two or more outcome measures jointly in a fixed effects multivariate linear model. Dear [3] used essentially the same method for combining survival curves in a meta-analysis, where each curve was characterized by estimated survival probabilities at two or more follow-up times. Van Houwelingen *et al.* [4] were the first to consider multivariate random effects meta-analysis. They introduced a bivariate linear random effects model for the joint analysis of one estimated outcome measure per treatment group. Essentially the same model was used by McIntosh [5] and Arends *et al*. [6] in the context of investigating the underlying risk as a source of heterogeneity in treatment effects across trials. Berkey *et al.* [7] introduced the general linear MIXED model as a general random effects regression method for meta-analysis of multiple outcomes. In a recent tutorial on advanced methods in meta-analysis [8], we adopted the general linear MIXED model as a general framework for multivariate meta-analysis and meta-regression. In fact this approach can be considered as a direct generalization of the standard (univariate) DerSimonian–Laird [1] model to higher dimensions. To apply the model the estimated vector of outcome measures along with the corresponding estimated covariance matrix per trial is needed. The parameters are estimated with (restricted) maximum likelihood, acting as if the within-trial covariance matrices are known. In this paper we follow this approach in a meta-analysis about the effect on stroke-free survival of surgery versus conservative treatment in patients with high risk for stroke. Different from van Houwelingen *et al.* [8] and Berkey *et al.* [7], who had two outcome measures, we have three outcome measures: the event rate in the conservative treatment group, and the short-term and long-term event rate in the surgery group. The complication is that, because of the peri-operative mortality, the short-term stroke-free survival in the surgery group is lower than in the conservative group, while stroke-free survival on the long-term is in favour of the surgical treatment because of a lower event rate once the operation is survived. We use a trivariate random effects model for the analysis of the data, and compare the results with univariate analyses. We show that the multivariate analysis is potentially much more informative than univariate analyses and can be carried out relatively easy in practice in standard software. Almost all models were fitted using Proc MIXED of SAS [9], while a few exact analyses were done with Proc NLMIXED. In this paper we focus on the application. More theoretical details and background can be found in the recent tutorial on advanced methods in meta-analysis [8]. In Section 2 we describe the data. In Section 3 the models are introduced and the advantages of multivariate modelling are discussed. In Section 4 we give the results, and the paper ends with a discussion in Section 5.

## 2. DATA

In this paper we analyse data from a meta-analysis of Vokó *et al.* [10] about the effect of carotid endarterectomy on all-cause mortality and stroke-free survival based on the aggregated data from 19 randomized trials. The vascular surgical procedure called carotid endarterectomy aims to remove the atherosclerotic plaque of the internal carotid artery and to restore the lumen of the vessel. To prevent cerebral infarction or death for people with increased levels of stenosis, one frequently performs a carotid endarterectomy [11, 12]. Although the operation mortality and morbidity is not negligible, the hope is that patients on average are better off because of lower event rates once the operation is survived. Several clinical trials comparing carotid endarterectomy plus best medical care with medical treatment alone have been done

 

or are under way. Some of the studies that are already published are combined in the meta-analysis of Vokó *et al.* [10].

All selected trials were randomized clinical trials in which the indication of carotid endarterectomy was stroke prevention rather than treatment of acute stroke and in which the methodology was judged appropriate (no excessive loss to follow-up, symmetrical outcome assessment, analysed by treatment assignment from the moment of randomization onwards). For further details about the selection of the trials we refer to Vokó *et al.* [10]. Together the 19 randomized clinical trials comprise in total 8991 patients being at increased risk of stroke, 4780 allocated to surgery and 4211 to conservative treatment. In this paper we only look at stroke-free survival, so the event of interest is defined as stroke or death.

The basic data available for the 19 trials ($i = 1, \ldots, 19$) were:

1. number of patients in the surgical group ($k_i$) and number of events in the first month after operation ($x_i$);
2. number of events ($y_i$) and person years of follow-up ($n_i$) in the surgical group from 1 month post-operation onwards;
3. number of events ($z_i$) and person-years of follow-up ($m_i$) in the conservative treatment group.

An implicit assumption was that the hazard rate is constant in the surgical group after the first month, and in the conservative group during the whole follow-up. This assumption could be somewhat relaxed, by splitting the time period in more periods and assuming a piecewise constant hazard rate. Of course, this results in more parameters to be estimated, that is, one for each time interval, but the methods of this paper remain applicable. In our case there were no data on events and person-years on sub time intervals available.

The true event probability in the first month (called 'post-surgical risk' in the following) in the surgical group of trial $i$ is denoted by $\pi_i$, estimated by the observed event probability $\hat{\pi}_i = x_i/k_i$. The true event rate after one month (called 'surgical long-term event rate' in the following) in the surgical group is denoted by $\lambda_i$. It is estimated by $\hat{\lambda}_i = y_i/n_i$. The true event rate (called 'conservative long-term event rate' in the following) in the conservative treatment group is denoted by $\mu_i$, estimated by $\hat{\mu}_i = z_i/m_i$. The data are given in Table I.

The main questions were to compare the event-free survival of the two treatments depending on the length of the follow-up period and to investigate how the difference is modified by the level of underlying risk in the population. Secondary questions concerned the mean post-operative risk and the heterogeneity in it between trials, the difference between treatments in long-term event rate, and again how these are affected by the underlying risk.

## 3. METHODS

### *3.1. Parameter transformations*

As usual we transform the parameters such that the transformed parameters range from minus to plus infinity. This is more natural when random effects are employed. Moreover, the

 

Table I. Data of the 19 clinical trials.

| Trial | Surgical group | | | | | | Conservative group | | |
|---|---|---|---|---|---|---|---|---|---|
| | First month | | | After one month | | | | | |
| | Events $x_i$ | Patients $k_i$ | Risk $\hat{\pi}_i$ | Events $y_i$ | Person years $n_i$ | Event rate $\hat{\lambda}_i$ | Events $z_i$ | Person years $m_i$ | Event rate $\hat{\mu}_i$ |
| 1 | 19 | 169 | 0.112 | 26 | 564.56 | 0.046 | 38 | 507.06 | 0.075 |
| 2 | 7 | 20 | 0.350 | 5 | 28.55 | 0.175 | 7 | 50.60 | 0.138 |
| 3 | 5 | 91 | 0.055 | 5 | 79.01 | 0.063 | 9 | 96.33 | 0.093 |
| 4 | 5 | 78 | 0.064 | 23 | 446.71 | 0.051 | 16 | 366.17 | 0.044 |
| 5 | 3 | 162 | 0.019 | 57 | 920.04 | 0.062 | 39 | 686.63 | 0.057 |
| 6 | 14 | 200 | 0.070 | 64 | 1141.92 | 0.056 | 46 | 829.85 | 0.055 |
| 7 | 18 | 190 | 0.095 | 47 | 1103.46 | 0.043 | 32 | 720.13 | 0.044 |
| 8 | 22 | 350 | 0.063 | 104 | 2005.75 | 0.052 | 86 | 1403.83 | 0.061 |
| 9 | 22 | 232 | 0.095 | 60 | 1322.38 | 0.045 | 48 | 790.53 | 0.061 |
| 10 | 21 | 231 | 0.091 | 68 | 1316.46 | 0.052 | 69 | 985.82 | 0.070 |
| 11 | 12 | 251 | 0.048 | 86 | 1446.63 | 0.059 | 71 | 900.71 | 0.079 |
| 12 | 5 | 113 | 0.044 | 38 | 650.83 | 0.058 | 35 | 348.37 | 0.100 |
| 13 | 45 | 678 | 0.066 | 163 | 3065.42 | 0.053 | 209 | 3120.17 | 0.067 |
| 14 | 28 | 430 | 0.066 | 92 | 1974.13 | 0.047 | 156 | 1889.95 | 0.083 |
| 15 | 19 | 328 | 0.058 | 28 | 526.23 | 0.053 | 80 | 489.87 | 0.163 |
| 16 | 7 | 206 | 0.034 | 49 | 529.50 | 0.093 | 57 | 526.50 | 0.108 |
| 17 | 0 | 15 | 0.033 | 3 | 41.37 | 0.073 | 1 | 44.33 | 0.023 |
| 18 | 9 | 211 | 0.043 | 41 | 513.16 | 0.080 | 59 | 597.87 | 0.099 |
| 19 | 22 | 825 | 0.027 | 106 | 2004.36 | 0.053 | 146 | 2076.45 | 0.070 |

transformed parameters have better statistical properties if Wald-type confidence intervals and tests are used.

The post-surgical risk parameter $\pi_i$ is transformed to the log-odds scale: $\omega_i = \ln(\pi_i/(1-\pi_i))$. The estimated log-odds is denoted by $\hat{\omega}_i$. Its variance is estimated by

$$\text{var}(\hat{\omega}_i) = \frac{1}{x_i} + \frac{1}{k_i - x_i} \tag{1}$$

(One trial, number 17, had zero events. As is usually done, we added 1/2, so $x_{17} = 0.5$.) The long-term event rates for the surgical and conservative treatment, respectively, are logarithmically transformed:

$$\beta_i = \ln(\lambda_i), \text{ estimated by } \hat{\beta}_i = \ln(\hat{\lambda}_i) \text{ with estimated variance } \text{var}(\hat{\beta}_i) = 1/y_i \tag{2}$$

$$\alpha_i = \ln(\mu_i), \text{ estimated by } \hat{\alpha}_i = \ln(\hat{\mu}_i) \text{ with estimated variance } \text{var}(\hat{\alpha}_i) = 1/z_i \tag{3}$$

The variances follow under the assumption of an exponential survival time distribution or constant hazard rate. The estimated transformed outcome measures are given in Table II.

 

Table II. Transformed data of the 19 trials.

| Trial | Surgical group | | | | Conservative group | |
|---|---|---|---|---|---|---|
| | First month | | After one month | | | |
| | Log-odds post-surgical risk | | log long-term event rate | | log long-term event rate | |
| | $\hat{\omega}_i$ | $\text{var}(\hat{\omega}_i)$ | $\hat{\beta}_i$ | $\text{var}(\hat{\beta}_i)$ | $\hat{\alpha}_i$ | $\text{var}(\hat{\alpha}_i)$ |
| 1 | −2.066 | 0.059 | −3.078 | 0.038 | −2.591 | 0.026 |
| 2 | −0.619 | 0.220 | −1.742 | 0.200 | −1.978 | 0.143 |
| 3 | −2.845 | 0.212 | −2.760 | 0.200 | −2.371 | 0.111 |
| 4 | −2.681 | 0.214 | −2.966 | 0.043 | −3.130 | 0.063 |
| 5 | −3.970 | 0.340 | −2.781 | 0.018 | −2.868 | 0.026 |
| 6 | −2.587 | 0.077 | −2.882 | 0.016 | −2.893 | 0.022 |
| 7 | −2.257 | 0.061 | −3.156 | 0.021 | −3.114 | 0.031 |
| 8 | −2.702 | 0.049 | −2.959 | 0.010 | −2.793 | 0.012 |
| 9 | −2.256 | 0.050 | −3.093 | 0.017 | −2.802 | 0.021 |
| 10 | −2.303 | 0.052 | −2.963 | 0.015 | −2.659 | 0.014 |
| 11 | −2.992 | 0.088 | −2.823 | 0.012 | −2.541 | 0.014 |
| 12 | −3.073 | 0.209 | −2.841 | 0.026 | −2.298 | 0.029 |
| 13 | −2.644 | 0.024 | −2.934 | 0.006 | −2.703 | 0.005 |
| 14 | −2.664 | 0.038 | −3.066 | 0.011 | −2.494 | 0.006 |
| 15 | −2.789 | 0.056 | −2.934 | 0.036 | −1.812 | 0.013 |
| 16 | −3.347 | 0.148 | −2.380 | 0.020 | −2.223 | 0.018 |
| 17 | −3.367 | 2.069 | −2.624 | 0.333 | −3.792 | 1.000 |
| 18 | −3.111 | 0.116 | −2.527 | 0.024 | −2.316 | 0.017 |
| 19 | −3.597 | 0.047 | −2.940 | 0.009 | −2.655 | 0.007 |

### 3.2. *Univariate analyses*

*3.2.1. Surgical risk.* To describe the post-surgical risk among studies we adopt the standard (univariate) random effects model of DerSimonian and Laird [1]:

$$\begin{aligned} \hat{\omega}_i &\cong \text{N}(\omega_i, \text{var}(\hat{\omega}_i)) \\ \omega_i &\cong \text{N}(\omega, \sigma_\omega^2) \end{aligned} \tag{4}$$

We will refer to the two sub-models as the measurement error model and the structural model, respectively. Here $\omega_i$ is the true logit(post-surgical risk) for trial $i$. The $\omega_i$'s may vary over trials, and are assumed to follow a normal distribution with mean $\omega$ and standard deviation $\sigma_\omega$, the latter characterizing the heterogeneity among trials. The $\omega_i$'s are not observed, but estimated by $\hat{\omega}_i$, which are assumed to have a normal distribution with mean $\omega_i$ and variance given by (1). The normality assumption for $\hat{\omega}_i$ is usually justified by large enough sample sizes. In fact, it is only assumed that the likelihood of $\hat{\omega}_i$ is well approximated by a normal distribution likelihood, which is a somewhat weaker assumption [8]. The normality assumption for the true $\omega_i$'s, although standard, is more crucial, but for a larger number of trials the inference on $\omega$ and $\sigma_\omega$ is robust against misspecification of this distribution [13, 14]. The parameters $\omega$ and $\sigma_\omega^2$ are estimated by standard maximum likelihood or restricted maximum likelihood methods [14], as if the study specific variances are known.

The model can be fitted using any standard linear MIXED model program provided that it is possible to fix the residual variances at user specified values. We used the procedure Proc MIXED of the SAS package [9].

 

In the recent advanced meta-analysis tutorial [8] exact approaches were discussed that can be fitted in very special and relatively simple cases. Where feasible we will do that, in order to compare the results with the approximate approach. When in model (4) the approximate measurement error model is replaced by the exact one we get

$$\begin{aligned} x_i &\cong \mathrm{Bin}\left(k_i, \frac{\exp(\omega_i)}{1+\exp(\omega_i)}\right) \\ \omega_i &\cong \mathrm{N}(\omega, \sigma_\omega^2) \end{aligned} \tag{5}$$

This is a logistic-normal random effects model which could be fitted for instance with EGRET [15] or MIXOR [16, 17]. We fitted the model using Proc NLMIXED of SAS [9]. (Since $x_i$ is allowed to be zero in this model, we changed $x_{17} = 0.5$ back to $x_{17} = 0$.)

*3.2.2. Long-term risks.* To compare the long-term risks between the treatments we look at the difference $\delta_i = \alpha_i - \beta_i$ and assume again the standard random effects model [1]:

$$\begin{aligned} &\hat{\delta}_i \cong \mathrm{N}(\delta_i, \mathrm{var}(\hat{\delta}_i)) \quad \text{with } \mathrm{var}(\hat{\delta}_i) = \mathrm{var}(\hat{\alpha}_i) + \mathrm{var}(\hat{\beta}_i) \\ &\delta_i \cong \mathrm{N}(\delta, \sigma_\delta^2) \end{aligned} \tag{6}$$

The estimated variances are computed using formulae (2) and (3). The parameters $\delta$ and $\sigma_\delta^2$ are estimated by maximum likelihood, assuming the study specific variances to be known. In this case it is also possible to fit the exact measurement error model. Exploiting the fact that the conditional distribution of $z_i$ given $z_i + y_i$ is binomial with parameters $z_i + y_i$ and $\mu_i m_i/(\mu_i m_i + \lambda_i n_i)$, respectively, the model can be written as

$$\begin{aligned} z_i &\cong \mathrm{Bin}\left(z_i + y_i, \frac{\exp(\log(m_i/n_i) + \delta_i)}{1+\exp(\log(m_i/n_i) + \delta_i)}\right) \\ \delta_i &\cong \mathrm{N}(\delta, \sigma_\delta^2) \end{aligned} \tag{7}$$

This is a logistic-normal random effects model with $\log(m_i/n_i)$ as an offset variable. Again we fitted this model with Proc NLMIXED.

*3.2.3. Cumulative survival ratio.* To compare event-free survival probabilities between the treatments over a fixed follow-up interval $(0, t)$ we look at the ratio of the cumulative $t$-year survival probabilities:

$$\mathrm{CSR}_i(t) = \frac{(1-\pi_i)\exp(-t'\lambda_i)}{\exp(-t\mu_i)}$$

where $t' = t - 1/12$ and $i$ the number of the trial. The choice of $t$ is arbitrary, but in the original meta-analysis of Vokó *et al.* [10] focus was on $t = 3$ years since most of the trials had a mean follow-up duration of about 3 years.

For the analysis it is natural to work with the logarithmically transformed parameter:

$$\rho_i(t) = -\log(1+\exp(\omega_i)) - t'\exp(\beta_i) + t\exp(\alpha_i)$$

 

The $\rho_i$'s are estimated by plugging in the estimates of $\omega_i$, $\alpha_i$ and $\beta_i$, while the variances are estimated using the delta-method by

$$\mathrm{var}(\hat{\rho}_i) = \hat{\pi}_i^2 \,\mathrm{var}(\hat{\omega}_i) + t'^2 \hat{\lambda}_i^2 \,\mathrm{var}(\hat{\beta}_i) + t^2 \hat{\mu}_i^2 \,\mathrm{var}(\hat{\alpha}_i) \tag{8}$$

Again we adopt the standard random effects model for the log(cumulative $t$-years survival ratio):

$$\begin{aligned} \hat{\rho}_i(t) &\cong \mathrm{N}(\rho_i(t), \mathrm{var}(\hat{\rho}_i(t))) \\ \rho_i(t) &\cong \mathrm{N}(\rho(t), \sigma_\rho^2(t)) \end{aligned} \tag{9}$$

For fixed value of $t$, the parameters $\rho(t)$ and $\sigma_\rho^2(t)$ are again estimated by maximum likelihood, assuming the study specific variances (8) to be known.

### 3.3. *Multivariate analyses*

In this subsection we introduce a multivariate model in which all three outcome measures are analysed simultaneously. The model is a direct generalization of the above univariate random effects models. Again we work with the transformed parameters. Given the true trial specific outcome measures we assume that the estimates follow a multivariate normal distribution:

$$\begin{pmatrix} \hat{\omega}_i \\ \hat{\beta}_i \\ \hat{\alpha}_i \end{pmatrix} \cong \mathrm{N}\left[ \begin{pmatrix} \omega_i \\ \beta_i \\ \alpha_i \end{pmatrix}, \begin{pmatrix} \mathrm{var}(\hat{\omega}_i) & 0 & 0 \\ 0 & \mathrm{var}(\hat{\beta}_i) & 0 \\ 0 & 0 & \mathrm{var}(\hat{\alpha}_i) \end{pmatrix} \right] \tag{10}$$

In general the covariances might be non-zero, but in our application the correlations are zero because the likelihood factorizes in three parts each involving only one parameter.

For the true outcome measures we assume a multivariate distribution as well:

$$\begin{pmatrix} \omega_i \\ \beta_i \\ \alpha_i \end{pmatrix} \cong \mathrm{N}\left[ \begin{pmatrix} \omega \\ \beta \\ \alpha \end{pmatrix}, \begin{pmatrix} \sigma_{\omega\omega} & \sigma_{\omega\beta} & \sigma_{\omega\alpha} \\ \sigma_{\omega\beta} & \sigma_{\beta\beta} & \sigma_{\alpha\beta} \\ \sigma_{\omega\alpha} & \sigma_{\alpha\beta} & \sigma_{\alpha\alpha} \end{pmatrix} \right] \tag{11}$$

Marginally this model assumes just a standard DerSimonian–Laird model for each outcome measure. The parameters can be estimated by standard likelihood, as if the variances of the trial specific outcome measures are known. We again used SAS Proc MIXED to fit the model.

The multivariate modelling has several advantages. First, instead of doing a number of univariate analyses each tailored to one specific question, the multivariate approach gives a complete and concise description of all data at one stroke. Once the model is fitted, it immediately gives the estimated mean post-surgical risk, the long-term risks under both treatments and the between-trial variances of these parameters. Inference on derived parameters can readily be carried out. For instance, the estimate of the difference in log(long-term risk) (see Section 3.2.2) can easily be computed, and the associated $P$-value and confidence interval follow directly from the covariance matrix of the estimates. In the univariate approach, a separate analysis had to be done for estimating the log(cumulative survival ratio) over $(0, t)$ for each value of $t$ that was of interest. The multivariate approach yields an estimate and confidence interval for the typical log(cumulative survival ratio) over $(0, t)$ as a relatively simple

 

function of $t$. Second, the multivariate approach yields the estimated correlations between the outcome measures. This can lead to more insight. For instance it might be interesting to know whether high post-surgical risks are associated with higher or rather with lower long-term risks, either under the conservative or under the surgical treatment. One would probably want to adjust the latter association for the long-term risk under the conservative treatment, then one looks at the partial correlation between $\omega_i$ and $\alpha_i$ given $\beta_i$. The third advantage that we mention is related to the previous point. Often one is interested in whether a measure of treatment benefit is modified by some measure of baseline risk. For instance, is the difference in log(long-term risk)'s between treatments associated with larger baseline risks as measured by the long-term risk under the conservative treatment? Or, is the log(cumulative survival ratio) over $(0,t)$ modified by the baseline risks as measured by the long-term risk under the conservative treatment? For these types of questions the univariate analyses of the previous section fall short and multivariate modelling is necessary. We elaborate on this in Section 4.2. The fourth advantage of the multivariate approach that we mention is that it is capable of dealing adequately with incomplete trials, that is when one or two outcome measures are missing, and therefore makes more efficient use of the data. If the missing outcomes are missing at random but not completely at random, the multivariate approach might also be more valid than the univariate analyses that necessarily leave out the incomplete trials. In our application the event of interest was defined as stroke or death. Most trial reports, but not all, also give the outcome event death alone. Probably both outcomes will be highly correlated. Therefore, if one is interested in the effect of treatment on the endpoint death, it would be advantageous to carry out a multivariate analysis with both the outcome death and outcome stroke or death. One can also think of a situation where one has a surrogate endpoint for all trials, and a smaller number of trials reporting both the 'true' endpoint as well as the surrogate. Then one is specifically interested in the trial-level correlation between both endpoints [18, 19].

## 4. RESULTS

### *4.1. Univariate analyses*

The univariate models (4), (6) and (9) were fitted using Proc MIXED of SAS [9]. Since this is not completely trivial, we refer the reader to the recent tutorial on advanced meta-analysis methods [8] for an explanation of how to do that.

The estimated mean log-odds of a post-surgical event was $\hat{\omega} = -2.681$, with standard error 0.133. Thus on the original scale the estimated mean post-surgical risk is 0.064 with approximate 95 per cent confidence interval (0.050, 0.082). The between-trials variance on the log-odds scale was estimated as 0.224, giving an approximate 95 per cent coverage interval of the true post-surgical risks of (0.026, 0.148). This indicates quite a large between-trial variation in post-surgical risks. The likelihood ratio test on $H_0$: $\sigma_\omega^2 = 0$ was borderline statistically significant ($P = 0.08$). Proc MIXED also gives a Satterthwaite approximation based 95 per cent confidence interval for the between-trials variance, (0.097, 0.950).

The exact measurement error model fitted by SAS Proc NLMIXED gives an estimated mean log-odds of a post-surgical event $\hat{\omega} = -2.739$, with standard error 0.130. This corresponds on the original scale to an estimated mean post-surgical risk of 0.065 with a 95 per cent confidence interval (0.049, 0.085). The estimated between-trials variance on the log-odds

 

Table III. Results of the univariate random effects model for the cumulative $t$-years survival probability ratio for some selected values of $t$.

| | $t$ (years) | | |
|---|---|---|---|
| | 1 | 3 | 8 |
| Estimated mean log CSR ($\hat{\rho}(t)$) | −0.0379 | −0.00756 | 0.0747 |
| Standard error of $\hat{\rho}(t)$ | 0.00844 | 0.0155 | 0.0356 |
| $P$-value for $H_0$: $\rho(t)=0$ | <0.0001 | 0.63 | 0.036 |
| Between-trials variance $\hat{\sigma}_\rho^2(t)$ | 0.000744 | 0.00232 | 0.0111 |
| LR text $p$-value for $H_0$: $\sigma_\rho^2=0$ | 0.0002 | 0.008 | 0.026 |
| Estimated cumulative survival ratio CSR($t$) | 0.963 | 0.9925 | 1.078 |
| 95 per cent confidence interval for CSR($t$) | 0.947, 0.979 | 0.963, 1.023 | 1.005, 1.155 |

scale was 0.211, resulting in a 95 per cent coverage interval of the true post-surgical risks of (0.026, 0.159). All of these estimates are quite similar to the approximate likelihood estimates.

The estimate of the mean difference in log(long-term event rate)'s was $\hat{\delta}=0.277$ with standard error 0.061, so on average the long-term event rate of the surgical treatment was highly significantly better than the conservative treatment. The estimated hazard ratio is 1.32 with approximate 95 per cent confidence interval (1.17, 1.49). The estimated between-trials variance in true log(long-term event rate difference)'s is 0.0268 (95 per cent confidence interval (0.0078, 0.6425)), significantly different from zero at the 5 per cent level ($P=0.04$). The approximate 95 per cent coverage interval of the true long-term event rate ratios is (0.96, 1.82), again indicating quite a large between-trials variation.

The exact measurement error model fitted by SAS Proc NLMIXED gives an estimated mean difference in log(long-term event rate)'s equal to $\hat{\delta}=0.278$, with standard error 0.064. The estimated hazard ratio is 1.32 with 95 per cent confidence interval (1.15, 1.51). This is very similar to the approximate likelihood estimates. The estimated between-trials variance in the true log(long-term event rate difference)'s is 0.032, significantly different from zero ($P=0.02$ likelihood ratio test). The 95 per cent coverage interval of the true long-term event rate ratios is (0.93, 1.58). This is all very similar to the results based on the approximate likelihood. Model (9) was fitted for a number of different values of $t$. The results for some selected values of $t$ are given in Table III.

In the previous analysis it was seen that the long-term event rate was better for the surgical treatment. However, for relative short follow-up times, the event-free cumulative survival probability is in favour of the conservative treatment because of the post-surgical risk. For example, from Table III it is seen that for one-year follow-up duration survival is very significantly worse for the surgical treatment. For longer follow-up duration the balance is in favour of the surgical treatment. From Table III it is seen that for $t=3$ years follow-up the estimated event-free survival probability is about equal for both treatments. From about $t=8$ years, cumulative survival for the surgical treatment is significantly better than for the conservative treatment. In Figure 1 the estimated cumulative survival ratio and its 95 per cent confidence interval is given as a function of $t$. Moreover the approximate 95 per cent coverage interval is given, that is, the interval in which the true cumulative survival probability ratio of a new trial will lie with about 95 per cent probability. It is seen that the estimated length of follow-up for which the two treatments are equivalent is 3.5 years with a 95 per cent

 

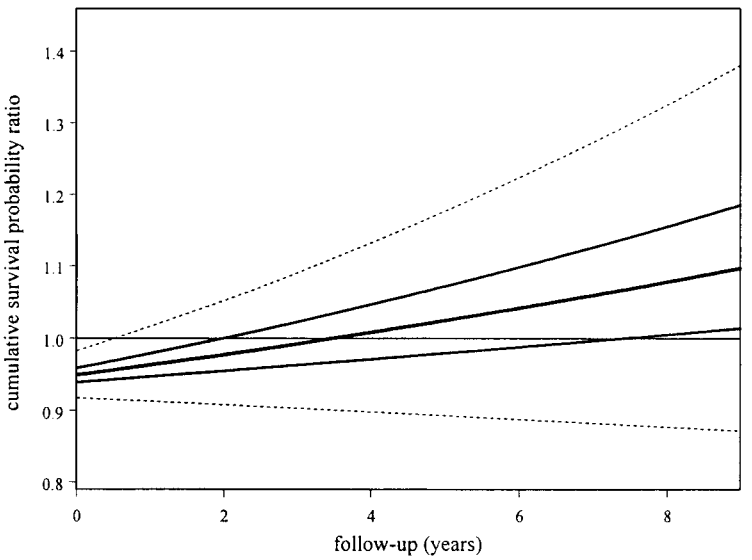


Figure 1. Estimated mean cumulative survival probability ratio, based on *univariate* analyses for different lengths of follow-up (bold curve). The inner two curves give the 95 per cent confidence interval, and the outer curves approximate 95 per cent coverage intervals.

Table IV. Some results of the trivariate models (10) and (11).

| Estimate | Outcome measure | | |
|---|---|---|---|
| | Logit of post-surgical risk ($\omega$) | Log long-term event rate of surgical treatment ($\beta$) | Log long-term event rate of conservative treatment ($\alpha$) |
| Mean | −2.707 | −2.891 | −2.573 |
| Standard error | 0.1337 | 0.0440 | 0.0777 |
| Between-trials variance | 0.2299 | 0.0167 | 0.0852 |

confidence interval running from 2.0 to 7.4 years. These values were determined by using a fine grid of values of $t$ and running the analysis for each value of $t$.

### 4.2. *Multivariate analysis*

The multivariate model could be fitted using SAS Proc MIXED as well. The main results of the multivariate meta-analysis are given in Table IV. The full estimated covariance matrix of the estimated mean outcome measures is

$$\text{covar}\begin{pmatrix}\hat{\omega}\\ \hat{\beta}\\ \hat{\alpha}\end{pmatrix} = \begin{pmatrix} 0.017888 & -0.002154 & -0.0005938\\ -0.002154 & 0.001939 & 0.001925\\ -0.0005938 & 0.001925 & 0.006036\end{pmatrix} \tag{11}$$

 

The estimated covariance matrix of the random effects is

$$\operatorname{covar}\begin{pmatrix}\omega_i\\ \beta_i\\ \alpha_i\end{pmatrix}=\begin{pmatrix}0.2299 & -0.03666 & -0.01144\\ -0.03666 & 0.01675 & 0.03220\\ -0.01144 & 0.03220 & 0.08519\end{pmatrix} \tag{12}$$

This covariance matrix turned out to be positive semi-definite. The estimated correlation matrix of the random effects is

$$\operatorname{corr}\begin{pmatrix}\omega_i\\ \beta_i\\ \alpha_i\end{pmatrix}=\begin{pmatrix}1 & -0.59 & -0.08\\ -0.59 & 1 & 0.85\\ -0.08 & 0.85 & 1\end{pmatrix}$$

SAS Proc MIXED also gives the estimated covariance matrix of the estimates of the random effects parameters (not shown).

Notice that for the logit of the post-surgical risk the result is almost identical to the above given univariate analysis. This is also true for the other two outcomes (univariate results not shown).

A number of questions could be answered using the results of the multivariate analysis. Let us start with comparing the two long-term event rates. The estimated mean difference in log(long-term event rate)'s is 0.318 with standard error $\sqrt{(0.001939 + 0.006036 - 2 \times 0.001925)} = 0.064$, not much different from the univariate analysis. The associated between-trials variance is estimated as $0.01675 + 0.08519 - 2 \cdot 0.0322 = 0.0375$, slightly larger than from the univariate analysis.

We now look at the $t$-years cumulative survival ratio. In the univariate approach we had to repeat the analysis for each value of $t$ of interest. An advantage of the multivariate model is that the estimated $t$-years cumulative survival ratio can be given as a simple function of $t$:

$$\widehat{\mathrm{CSR}}(t) = \exp(-0.06 + 0.021t) = 0.94\mathrm{e}^{0.021t}$$

Note that the interpretation of this is a little bit different from the CSR($t$) from the univariate analysis. In the univariate approach the mean log(CSR($t$)) was estimated. After exponentiating it can be interpreted as the estimated CSR($t$) for a trial with average log(CSR($t$)), or as the estimated median CSR($t$). Now we have estimated the CSR($t$) for the 'typical' clinical trial having average post-surgical log-odds and average log(long-term risk) under both treatments. Of course, based on the multivariate analysis it would be possible to compute the analogue of the univariate parameter estimate, by integrating the estimated log(CSR($t$)) over the estimated trivariate normal distribution of the random effects, but that is not very simple and there is no need to do that since the present parameter estimate is perfectly interpretable. The value of $t$ for which the cumulative survival probability over $(0,t)$ is equal for both treatments is estimated as 2.88, somewhat smaller than found in the univariate analyses.

The estimated variance of its logarithm is computed from the analogue of (8) and the covariance matrix (11), and is a simple quadratic function of $t$:

$$\operatorname{var}(\ln(\mathrm{CSR}(t))) = (0.712 - 0.0893t + 0.248t^2)/10^4$$

 

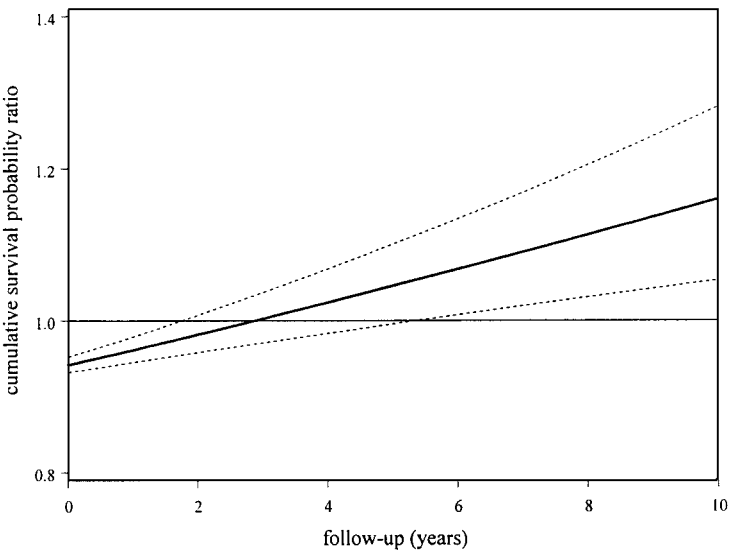


Figure 2. Estimated cumulative survival probability ratio, based on the *multivariate* analysis as a function of length of follow-up. The dotted curves give the 95 per cent confidence band.

The approximate 95 per cent confidence interval for CRS($t$) is thus given by

$$0.94\exp(0.021t \pm 1.96\sqrt{\{(0.712 - 0.0893t + 0.248t^2)\}/100}) \tag{13}$$

The confidence interval for the value of $t$ for which the event-free survival probability is equal for both treatments is conveniently computed by converting (13) and turns out to be (1.83, 5.50). As an illustration we give in Figure 2 the estimated CSR($t$) and its corresponding 95 per cent confidence interval.

An advantage of the above multivariate analysis is that the correlations between the different outcome measures are estimated. Notice that it would not be adequate to look at the simple correlations between the observed outcome measure estimates, since we are interested in the correlations between the underlying true trial specific outcomes. Moreover, the observed outcome measures have different precisions between trials and the errors might in general be correlated too (although this was not the case in our application), so that within- and between-trial correlation would be mixed up. An interesting finding in our example is that there was almost no correlation between the post-surgical risk and the long-term event rate under conservative treatment. This is an indication that the post-operative risk is not higher in high risk populations, in contrast to what was expected beforehand. Another finding is that there is a moderately high negative correlation of $-0.59$ between the post-surgical risk and the long-term risk under the surgical treatment. This is an indication that the most vulnerable patients tend to have an event in the first month after surgery leading to a selected group of patients with good long-term prognosis, a kind of 'survival of the fittest' phenomenon. Probably one would want to adjust this correlation for the event rate in the conservative treatment. The partial correlation turns out to be equal to $-1$, which is an even stronger

 

indication of this selection phenomenon that the patients with a post-operative event are probably the ones that otherwise would have had an event later on.

Above it appeared that there is quite some variation among trials in treatment effect measures such as the difference in log(long-term events rate)'s and the $t$-years cumulative survival probability ratios. One possibility of exploring this heterogeneity would be to make use of trial level covariates. The above multivariate model is straightforwardly extended with covariates, which might be different for different outcomes. In the application of this paper there were no covariates available. In the absence of covariates, although not only then, it is quite common to consider whether there is any association between patients' underlying risk of the event in question and the treatment effect measure. The underlying risk is a convenient and clinically relevant trial-level measure which can be interpreted as a summary of a number of unmeasured patient characteristics. In our application the log(long-term event rate) $\alpha_i$ is the straightforward choice for the baseline risk measure. Simply regressing the estimated treatment on the observed baseline risk measure would be a mistake for several reasons (see for instance Sharp [20]), and a more sophisticated approach is needed. A number of articles has been written on how to estimate the relation between treatment effects and underlying risk in meta-analyses [5, 6, 21–25]. The approach of this paper is in the spirit of Arends *et al.* [6] and McIntosh [5], and is easily carried out using the results of the multivariate analysis.

As a first example, suppose that one is interested in whether the long-term treatment effect is different between low and high risk populations, or, in other words, whether the long-term treatment effect depends on the long-term event rate in the conservative treatment group. Then it is natural to look at the regression line of $\delta_i = \alpha_i - \beta_i$ on $\alpha_i$, which is given by

$$\delta_i = \delta + \left(1 - \frac{\sigma_{\alpha\beta}}{\sigma_{\alpha\alpha}}\right)(\alpha_i - \alpha)$$

All ingredients that we need are available from the multivariate results above. The estimated regression line is $\delta_i = 1.918 + 0.622\alpha_i$. The estimated standard error of the slope is 0.2924, computed with the delta method using the covariance matrix of the estimated covariance matrix (not shown). We conclude that the slope differs significantly from zero, so the long-term event risk ratio increases with increasing baseline event rate in favour of the surgical treatment. This is illustrated in Figure 3. A confidence band for the regression line might be computed with the delta method, using the estimated covariance matrices of the fixed effects and covariance parameters. The residual standard deviation is 0.069 and the percentage explained variance is quite high, 87.5 per cent. The fit of the regression line to the observed long-term event ratios appears to be quite good, except maybe for the very small trial number 17 at the right hand side below.

Another relationship of interest is between the $t$-years cumulative survival ratio and the underlying risk. Therefore we look at the conditional distribution of $(\omega_i, \beta_i)$ given $\alpha_i$, which is bivariate normal with mean

$$\begin{pmatrix} \omega + \dfrac{\sigma_{\omega\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i - \alpha) \\ \beta + \dfrac{\sigma_{\beta\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i - \alpha) \end{pmatrix}$$

 

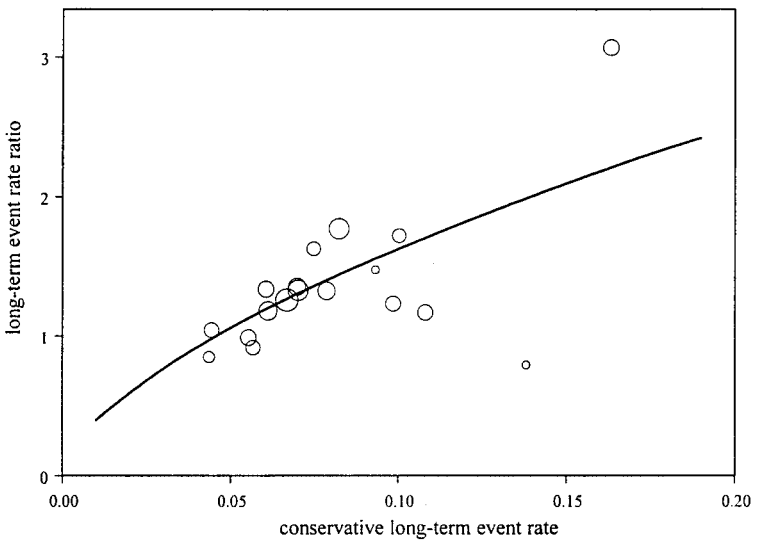


Figure 3. Observed long-term event rate ratios (conservative relative to surgical treatment) plotted against observed long-term event rate in the conservative treatment group and estimated regression line of true long-term event rate ratio on true conservative long-term event rate. Area of circles is proportional to the number of long-term events.

The typical log(CRS($t$)) therefore is

$$\rho(t;\alpha_i) = -\ln\left(1+\exp\left(\omega+\frac{\sigma_{\omega\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i-\alpha)\right)\right) - \left(t-\frac{1}{12}\right)\exp\left(\beta+\frac{\sigma_{\beta\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i-\alpha)\right)+t\exp(\alpha_i)$$

Using the results of the multivariate analysis this is estimated by

$$\hat{\rho}_i(t;\alpha_i) = -\ln(1+\exp(-3.053-0.134\alpha_i)) - \left(t-\frac{1}{12}\right)\exp(-1.918+0.378\alpha_i)+t\exp(\alpha_i)$$

Given $\alpha_i$, the estimated break-even value of $t$, that is the value of $t$ for which the survival probability over $(0,t)$ is equal for both treatments, is given by

$$t_{\text{break-even}} = \frac{\ln(1+\exp(\omega+\frac{\sigma_{\omega\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i-\alpha)))-\exp(\beta+\frac{\sigma_{\beta\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i-\alpha))/12}{\exp(\alpha_i)-\exp(\beta+\frac{\sigma_{\beta\alpha}}{\sigma_{\alpha\alpha}}(\alpha_i-\alpha))}$$

 

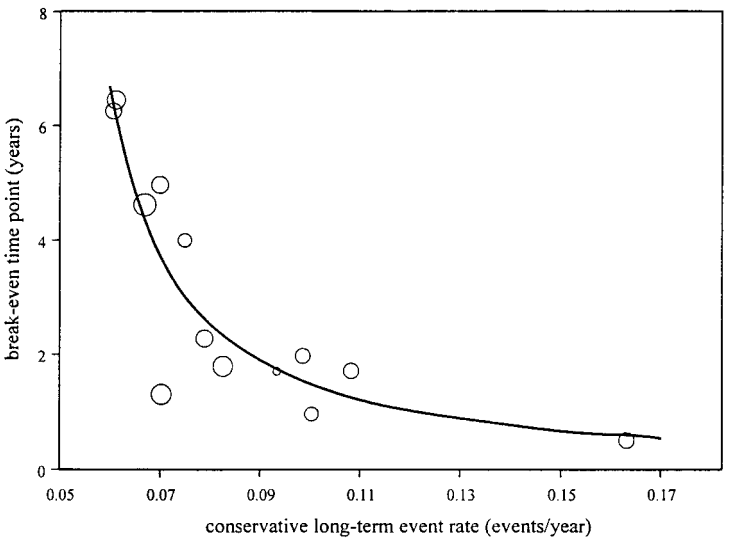


Figure 4. Observed and predicted break-even time against long-term event rate for the conservative treatment. Area of circles is proportional to the total number of events in a trial.

This is estimated by

$$\hat{t}_{\text{break-even}} = \frac{\ln(1+\exp(-3.053-0.134\alpha_i)) - \exp(-1.919+0.378\alpha_i)/12}{\exp(\alpha_i) - \exp(-1.919+0.378\alpha_i)}$$

The estimated break-even point is positive as long as the long-term event rate for the surgical treatment is lower than the predicted long-term event rate for the conservative treatment. As an illustration we plot in Figure 4 the observed and predicted break-even point $t$ against the observed event rate in the conservative treatment, for the trials with positive observed and predicted break-even times. Confidence intervals might be computed with the delta method, using the estimated covariance matrices of the fixed effects and covariance parameters. The fit of the observed break-even times to the predicted break-even times appears to be quite good.

In Figure 5 the estimated CSR($t$) is given as a function of the true conservative long-term event rate $\mu_i = \exp(\alpha_i)$ for selected values of $t$. From the picture it can be seen for instance that in a population with baseline incidence over about 8 events per 100 person-years the typical 3-years survival probability under surgical treatment is better than under conservative treatment. Again confidence intervals for $\rho(t;\alpha_i)$ can be constructed via the delta method. Notice that the predicted 3-years survival probability ratio curve fits the observed 3-years survival probability ratios very well, except for one outlier, the very small trial number 17.

## 5. DISCUSSION

To our knowledge this is the first example of a multivariate random effects meta-analysis combining more than two outcomes. The model that we used is quite generally applicable.

 

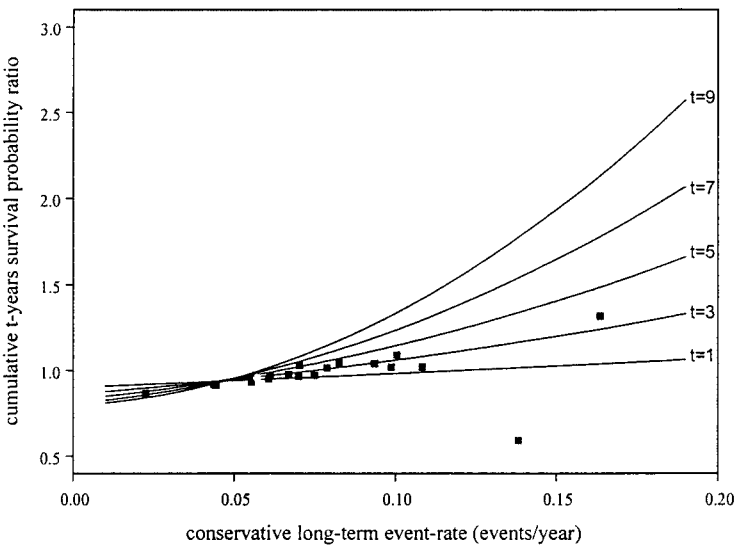


Figure 5. Estimated cumulative survival ratio probability ratio (surgical relative to conservative treatment) for different follow-up periods $(0,t)$ ($t$ in years). The points are the observed 3-years cumulative survival probability ratios plotted against the observed long-term event rate in the conservative treatment group.

In our application we had no covariates available that could explain heterogeneity between trials. If they are available, they can be used without any further difficulties. In our example, the different outcome measures were independent within trials. Also this is no limitation of the method and the presence of correlations can easily be accommodated. For a bivariate meta-analysis example where this is done we refer to van Houwelingen *et al.* [8]. We demonstrated the advantages of the multivariate analysis upon the univariate analyses. One multivariate analysis yielded much more information than a number of separate univariate analyses. The multivariate analysis revealed the relations between the different outcomes and gave simple expressions for estimation of derived treatment effect parameters such as, for instance, the cumulative survival probability ratio as a function of follow-up duration. Furthermore, the results of the multivariate modelling enabled us to easily estimate the relation of different treatment effect parameters and the underlying risk. We did not have missing outcome measures in our example, but our method allows them. In other applications this can increase efficiency compared with the analysis restricted to only the trials with a complete set of outcome measures. This also makes the model very useful in modelling the relationship between surrogate and true endpoints in a meta-analysis with a mix of trials, some of them reporting both the surrogate and the true outcome and the others only the surrogate outcomes. Fortunately, the multivariate model can easily be fitted in standard general linear MIXED model programs, although not every program will have the appropriate options. We used SAS Proc MIXED, but we guess that other packages such as S-plus or MLWin might also be used, although we do not have extensive experience with these programs. The essential requirement is that the residual variances can be fixed at arbitrary values per individual trial [8].

 

Table V. Some results of the trivariate model fitted by BUGS.

| Estimate | Outcome measure | | |
|---|---|---|---|
| | Logit of post-surgical risk ($\omega$) | Log long-term event rate of surgical treatment ($\beta$) | Log long-term event rate of conservative treatment ($\alpha$) |
| Mean | −2.760 | −2.907 | −2.604 |
| Standard error | 0.1322 | 0.0464 | 0.0805 |
| Between-trials covariance matrix of the random effects | $\text{covar}\begin{pmatrix}\omega_i\\ \beta_i\\ \alpha_i\end{pmatrix} = \begin{pmatrix}0.2226 & -0.0356 & -0.0266\\ -0.0356 & 0.0175 & 0.0224\\ -0.0266 & 0.0224 & 0.0974\end{pmatrix}$ | | |

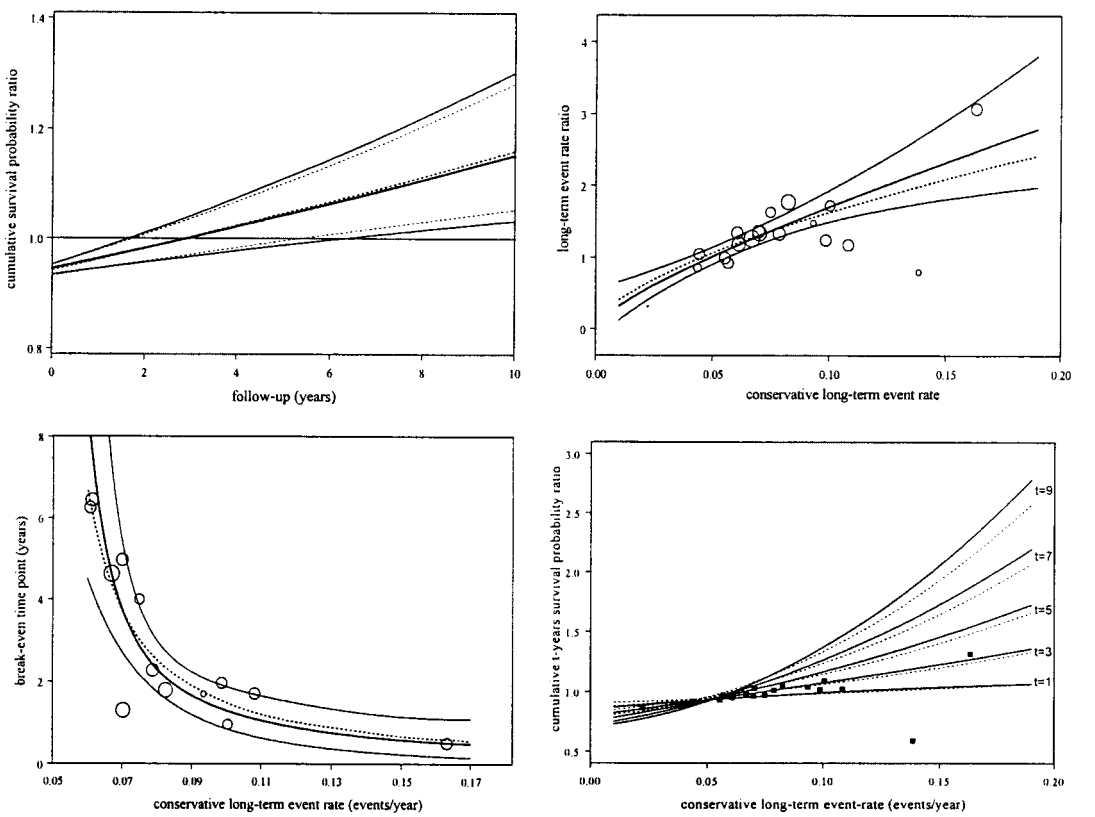


Figure 6. Analogues of Figures 2 to 5, with approximate likelihood as well as the Bayesian approach. The dotted lines represent the approximate likelihood estimates, the solid lines represent the results of the BUGS analysis.

 

We fitted the multivariate model using straightforward likelihood, but approximate because we acted as if the residual variances were estimated without error. In a few special univariate cases an exact likelihood was possible as well. In those cases the results turned out to be very similar. At present, an exact likelihood approach is not feasible in the multivariate case. An alternative approach would be to fit the model using Bayesian methods. This can for instance be done in the free available Bayesian analysis package BUGS [26]. One advantage is that the exact likelihood can be used by specifying the distribution for the outcome measure, in our example a binomial distribution for the number of post-operative events and a Poisson distribution for the events on long-term in both treatment groups. Another advantage is that, since BUGS uses MCMC methods to sample from the posterior distribution of all parameters, the inference based on the results of the fitted model can be easily built in. In Section 4.2 we computed by hand derived results such as the cumulative survival probability ratio and the break-even point, and their regression with the underlying risk. The estimates were quite easily computed, but the standard errors and confidence intervals using the delta method are more cumbersome, especially for the regression relationships. In BUGS this kind of derived inference including the (Bayesian) confidence intervals can be done very conveniently in the program. Some results of the BUGS analysis are presented in Table V.

The results are very similar to those of the approximate likelihood approach (Table III). As a further illustration we reproduce the analogues of Figures 2 to 5 now with the Bayesian approach, see Figure 6. In Figure 6 the dotted lines represent the approximate likelihood estimate, while the solid lines represent the results of the BUGS analysis. Again the results are very comparable.

Of course, a practical disadvantage of this approach is that fitting this kind of model in a program like BUGS can be quite time consuming, and therefore the approach presented in this paper is much more practical.

## REFERENCES

1. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**(3):177–188.
2. Raudenbush SW, Becker BJ, Kalaian H. Modeling multivariate effect sizes. *Psychological Bulletin* 1988; **103**(1):111–120.
3. Dear KB. Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics* 1994; **50**(4):989–1002.
4. Van Houwelingen JC, Zwinderman K, Stijnen T. A bivariate approach to meta-analysis. *Statistics in Medicine* 1993; **12**:2272–2284.
5. McIntosh MW. The population risk as an explanatory variable in research syntheses of clinical trials. *Statistics in Medicine* 1996; **15**:1713–1728.
6. Arends LR, Hoes AW, Lubsen J, Grobbee DE, Stijnen T. Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Statistics in Medicine* 2000; **19**(24):3497–3518.
7. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine* 1998; **17**:2537–2550.
8. Houwelingen HC van, Arends L, Stijnen T. Tutorial in Biostatistics. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 2002; **21**:589–624.
9. SAS [program]. Version 8.0. SAS Institute, Inc.: Cary, N.C., 1999.
10. Vokó Z. Etiology and prevention of stroke. The Rotterdam Study. Thesis, Erasmus University Rotterdam, 2000; 59–75.
11. Dyken ML, Pokras R. The performance of endarterectomy for disease of the extracranial arteries of the head. *Stroke* 1984; **15**:948–950.
12. Gillum RF. Epidemiology of carotid endarterectomy and cerebral arteriography in the United States. *Stroke* 1995; **26**:1724–1728.
13. Verbeke G, Lesaffre E. The effect of misspecifying the random effects distribution in linear models for longitudinal data. *Computational Statistics and Data Analysis* 1997; **23**:541–556.

 

14. Verbeke G. Linear mixed models for longitudinal data. In *Linear Mixed Models in Practice*, Verbeke G, Molenberghs G (eds). Springer-Verlag: New York, 1997; 63–153.
15. CYTEL Software Corporation. EGRET for Windows [program]. Version 2.0.3. CYTEL Software Corporation: Cambridge, 1999.
16. MIXOR [program]. Version 2.0. Hedeker D, Gibbons RD: Chicago, 1996.
17. Hedeker D, Gibbons RD. MIXOR: A computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine* 1996; **49**:157–176.
18. Buyse M, Molenberghs G, Burzykowsky T, Renard D, Geys H. The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics* 2000; **1**(1):49–68.
19. Gail MH, Pfeiffer R, van Houwelingen HC, Carroll RJ. On meta-analytic assessment of surrogate outcomes. *Biostatistics* 2000; **1**(3):231–246.
20. Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *British Medical Journal* 1996; **313**:7059:735–738.
21. Brand R. Importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials (letter). *Statistics in Medicine* 1994; **13**(3):293–296.
22. Senn SJ. Relation between treatment benefit and underlying risk in meta-analysis. *British Medical Journal* 1996; **313**:1550.
23. Sharp SJ, Thompson SG. Analysing the relationship between treatment effect and underlying risk in meta-analysis: comparison and development of approaches. *Statistics in Medicine* 2000; **19**:3251–3274.
24. Thompson SG, Smith TC, Sharp SJ. Investigating underlying risk as a source of heterogeneity in meta-analysis. *Statistics in Medicine* 1997; **16**(23):2741–2758.
25. Van Houwelingen HC, Senn S. Investigating underlying risk as a source of heterogeneity in meta-analysis (letter). *Statistics in Medicine* 1999; **18**:107–113.
26. Gilks W, Thomas A, Spiegelhalter D. A language and program for complex Bayesian modelling. *Statistician* 1994; **43**:169–177.

 