STATISTICS IN MEDICINE
Statist. Med. **17**, 1923–1942 (1998)

AN EMPIRICAL STUDY OF THE EFFECT OF THE CONTROL RATE AS A PREDICTOR OF TREATMENT EFFICACY IN META-ANALYSIS OF CLINICAL TRIALS

CHRISTOPHER H. SCHMID[1]*, JOSEPH LAU[1], MARTIN W. McINTOSH[2] AND JOSEPH C. CAPPELLERI[3]

[1] *Division of Clinical Care Research, Department of Medicine, New England Medical Center and Tufts University School of Medicine, 750 Washington St., Boston, MA 02111, U.S.A.*
[2] *Department of Biostatistics, University of Washington, Seattle, WA 98195, U.S.A.*
[3] *Department of Clinical Research, Pfizer Central Research, Groton, CT 06340, U.S.A.*

SUMMARY

If the control rate (CR) in a clinical trial represents the incidence or the baseline severity of illness in the study population, the size of treatment effects may tend to vary with the size of control rates. To investigate this hypothesis, we examined 115 meta-analyses covering a wide range of medical applications for evidence of a linear relationship between the CR and three treatment effect (TE) measures: the risk difference (RD); the log relative risk (RR), and the log odds ratio (OR). We used a hierarchical model that estimates the true regression while accounting for the random error in the measurement of and the functional dependence between the observed TE and the CR. Using a two standard error rule of significance, we found the control rate was about two times more likely to be significantly related to the RD (31 per cent) than to the RR (13 per cent) or the OR (14 per cent). Correlations between TE and CR were more likely when the meta-analysis included 10 or more trials and if patient follow-up was less than six months and homogeneous. Use of weighted linear regression (WLR) of the observed TE on the observed CR instead of the hierarchical model underestimated standard errors and overestimated the number of significant results by a factor of two. The significant correlation between the CR and the TE suggests that, rather than merely pooling the TE into a single summary estimate, investigators should search for the causes of heterogeneity related to patient characteristics and treatment protocols to determine when treatment is most beneficial and that they should plan to study this heterogeneity in clinical trials. © 1998 John Wiley & Sons, Ltd.

1. INTRODUCTION

Clinical trials with patients randomly assigned to a treatment group or a control group have become the gold standard for evaluating the effect of new drugs, treatments and delivery systems. When the outcome is dichotomous, the group risk is measured by its event rate, either the

* Correspondence to: Christopher H. Schmid, Biostatistics Research Center, Division of Clinical Care Research, New England Medical Center, Box 63, 750 Washington St, Boston, MA 02111, U.S.A. E-mail: cschmid@es.nemc.org

Contract/grant sponsor: Agency for Health Care Policy and Research of the United States Public Health Service
Contract/grant number: R01-HS07782, R01-HS08532

CCC 0277–6715/98/171923–20$17.50

Received March 1997
Revised September 1997

treatment rate (TR) or the control rate. The treatment rate is the proportion of patients in the treatment group with the event of interest and the control rate is the corresponding proportion in the control group. These two rates are commonly compared to measure treatment effect by the risk ratio, odds ratio and risk difference. The proliferation of clinical trials has led to instances in which a series of trials has given conflicting results. At least four sources for the disagreement are: (i) application of treatment to different study populations; (ii) use of different treatment regimens or protocols; (iii) variation in the quality of technical design or execution; and (iv) random variation.[1–4] The heterogeneity of the results, therefore, complicates generalization of the findings.

When the disagreement arises mainly because smaller studies do not have sufficient power to obtain a significant result, two different approaches to reduce random variation by increasing sample size are taken in order to obtain a more precise estimate of the common treatment effect in the population. The first is to conduct a meta-analysis by combining the studies to obtain a pooled treatment effect. The second approach is to conduct a very large clinical trial – mega-trial – typically involving more than 10,000 patients.[5] Mega-trials can also address variation in study design and execution because they are usually planned with input from many different experts to arrive at a consensus.

When heterogeneity in the outcomes of different studies reflects differences in patient populations or treatment protocols, however, the clinical dispute may be substantial and a pooled result may not be the best way to describe treatment efficacy.[6] Notable disagreements that have arisen recently between mega-trials and meta-analyses of smaller trials[7,8] may reflect an attempt to compare apples and oranges, a large trial testing one treatment in a general population versus a collection of smaller trials testing variations of the treatment in specific subpopulations. Because the heterogeneity may indicate that treatments work differently for different patients, substantial disagreement or contradiction among studies is not unexpected and we should not discount it. Rather, the opportunity arises to analyse the sources of heterogeneity and seek to describe their impact on treatment efficacy.[6,9,10]

Meta-regression is a technique for using regression analysis to assess the relationship between treatment efficacy and characteristics of the studies that may induce between-study variation.[11–14] These characteristics relate to study protocols (for example, dosages or durations of treatment), to patients (for example, demographic factors or severity of illness) or to factors concerning the execution of the study (for example, proper blinding). Without the benefit of individual patient data, meta-regression models must rely on the summary results of published studies. These summary results describe only between-study, not between-patient, variation in the risk factors and are therefore most useful for a study of characteristics that differ across studies.

However, key risk factors that vary across patients and that can be measured only as aggregate values, such as age and gender, are difficult to address adequately by meta-regression. One reason for this is that aggregated values tend to exhibit little between-study variation, thus providing minimal information across the potential range of the factor. Use of aggregated values may also introduce bias because of the failure to account for the within-study variation.[15,16] For example, while patients in two studies may have a mean age of 40 years, patients in one study may have much greater variation in their ages than patients in the other study. If the outcome tends to occur more frequently in the very old, the study with the greater within-patient variation and a higher proportion of very old individuals will have higher incidence rates.

Several investigators have proposed that the control rate can serve as a surrogate for these patient risk factors.[17–20] The control rate is a summary measure always reported. Variation in

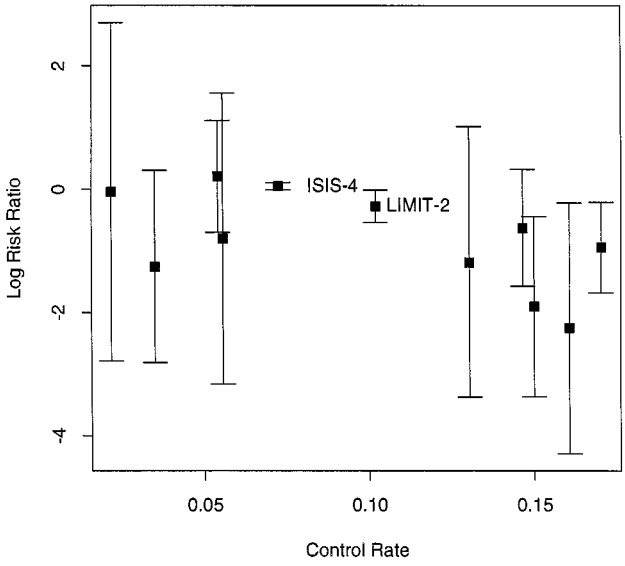

Figure 1. Plot of the natural logarithm of the risk ratio versus the control rate for 11 RCTs of the effect of magnesium therapy for acute myocardial infarction. Error bars indicate a 95 per cent confidence interval on the log risk ratios

control rates across studies may reflect different patient populations, underlying baseline risk of patients, length of study follow-up and treatment delivery. As a readily available, simple measure to explore treatment effect heterogeneity, the control rate may help delineate when one needs further research on risk factors.

Some analyses in the medical literature have already used the control rate as a covariate in a regression analysis. In a meta-analysis of tocolysis by beta-mimetics, Brand and Kragt proposed a correlation between increasing baseline risk (control rate) and increasing benefits (decreasing odds ratio).[18] Thompson[6] and Smith[21] found that the benefits of lowering the serum cholesterol on mortality depended on the initial level of risk. Boissel[22] reanalysed the data on the apparent lethal effects of anti-arrhythmic drugs in post-myocardial infarct patients[23,24] and suggested that high-risk patients may still derive benefits from the therapy. In the case of magnesium therapy for patients with acute myocardial infarction, one of the causes of the discrepancy between the results of ISIS-4[25] and prior meta-analyses of small studies,[26,27] as well as between the results of ISIS-4 and a moderately sized study (LIMIT-2)[28] may have been the low mortality rate in the control group of ISIS-4[29] (Figure 1).

Though there is increasing appreciation for the importance of meta-regression, the proper statistical methods to perform the analysis remain at issue. Senn[30] and Sharp *et al.*[31] have pointed out pitfalls in the fitting of regression models in which the dependent variable is an explicit function of the independent variable. McIntosh[32] has developed a hierarchical model that adjusts for this functional relationship in addition to the measurement error introduced by the uncertainty in the estimation of the control rate. This paper is an empirical study of meta-regression with the control rate across a wide range of many meta-analyses. After presenting the statistical models in Section 2, we describe in Section 3 the frequency with which we found control rate effects under the observed and hierarchical models; we also evaluate the adequacy of

commonly employed measures of between-study heterogeneity. Section 4 discusses the implications of our findings for meta-analysis and clinical research.

2. METHODS

2.1. Potential Relationships Between Treatment Effect and Control Rate

To describe the potential relationship between the treatment effect and the control rate, we must choose a measurement scale for the outcome and for the control rate. The three common measures of treatment effect for comparing two proportions are the risk difference, $RD = TR - CR$, the risk ratio, $RR = TR/CR$, and the odds ratio, $OR = \{TR/(1 - TR)\}/\{CR/(1 - CR)\}$. We use the log risk ratio and log odds ratio because of their nicer statistical properties in normal linear models.

Interpreting the control rate as a proxy for either the prevalence or severity of illness, we might expect that some experimental treatments are more beneficial to sicker populations either because more sick people are available to cure or because those that are sick have more disease to cure. This trend might manifest itself as a roughly linear regression effect with increased baseline risk suggesting increased therapeutic benefit. Non-linear relationships between treatment effects and the control rates might also arise from a variety of clinical and biological rationales. For example, U-shaped curves describe treatments beneficial to patients with moderate risk, but harmful to low and high-risk patients. These could arise if the treatment failed to cure the very sick and had some associated toxicity. If the treatment was not toxic, but was still ineffective in the presence of severe disease, the curve would be J-shaped. Threshold effects introduce another type of non-linearity. If an antibiotic treatment were effective only against bacterial concentrations below a threshold, and these concentrations were related to disease severity, then the treatment would have no effect on very sick patients. High control rate groups would therefore tend to show no benefit and the regression curve might be some sort of bent line or a curve that approached an asymptote.

In this paper, we employ regression models linear in the control rate for two major reasons. First, the number of trials available in each meta-analysis is often small and moving beyond a linear model may require unavailable degrees of freedom. Second, the linear model is the most generally plausible model and will usually be the first model fit in many problems. In a survey such as this, we are not attempting to investigate evey possibility for every data set, but are merely pointing out general trends in the most parsimonious fashion. If no linear trend is indicated, often non-linear trends will not be present either.

2.2. Problems of Fitting by Weighted Least Squares

Brand and Kragt[18] suggested fitting the regression of the treatment effect on the control rate by weighted least squares (WLS) where the weights are the inverse of the variance of the treatment effect. Senn[30] and Sharp *et al.*[31] have shown, however, that such analyses are inappropriate because the presence of the control rate in the definition of the treatment effect (if measured by the risk difference, relative risk or odds ratio) naturally induces an observed correlation between the control rate and the treatment effect, even if there is no true association between them.

Two simple approaches can be used to eliminate this induced correlation. The independence between the observed treatment and control rates suggests that one can remove this correlation by using the treatment rate as the dependent variable regressed against the observed control rate.

The results may then be translated back to a more interpretable metric. If the risk ratio is of interest, for example, the regression based on the treatment rate, $TR = \alpha + \beta \times CR$, may be mapped directly onto the risk ratio scale as $RR = TR/CR = \beta + \alpha/CR$. The second approach regresses the risk difference, $TR - CR$, against the average of the treatment and control rates, $(TR + CR)/2$.[30] This removes the correlation if the variance of the treatment rate and control rate are the same. Again, one can express results on the relative risk scale as the sum of a constant and a coefficient multiplied by the inverse of the control rate.

Neither of these approaches, however, addresses the error in the measurement of the control rate. The control rate is a stochastic quantity so that a regression that includes the control rate in the independent variable must account for the random error in its estimation. Least squares regression assumes that the independent variables are fixed and gives an expected response conditional on the independent variables observed. Thus, least squares is inappropriate to measure the regression of the observed treatment effect on the observed control rate. Instead, we describe a hierarchical model introduced by McIntosh[32] that explicitly models the observation error and estimates the correct regression function.

2.3. The Hierarchical Model

Let Y_{E_i} denote the observed treatment effect and Y_{C_i} the observed control rate in the ith clinical trial and let θ_{E_i} and θ_{C_i} denote the corresponding true values. The regression model that describes the true relationship between θ_{E_i} and θ_{C_i} is given by the following *Structural model*

$$\theta_{E_i}|\theta_{C_i} \sim N(\mu_E + \beta_{true}(\theta_{C_i} - \mu_C), \tau_E^2)$$
$$\theta_{C_i} \sim N(\mu_C, \tau_C^2).$$

The parameter β_{true} is the population (true) slope of the regression of the treatment effect on the control rate, μ_E is the mean treatment effect at the average control rate μ_C, τ_C^2 represents the variance of control rates across studies and τ_E^2 is the residual variance from the regression of θ_{E_i} on θ_{C_i}. If we could observe θ_{E_i} and θ_{C_i}, a standard least squares regression would give the correct slope. Unfortunately, we observe θ_{E_i} and θ_{C_i} only indirectly through Y_{E_i} and Y_{C_i} and the associated measurement error leads to a biased estimate of the regression function. Assume that the observed quantities are related to the true quantities by the *Observation model*:

$$\begin{pmatrix} Y_{E_i} \\ Y_{C_i} \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_{E_i} \\ \theta_{C_i} \end{pmatrix}, \begin{pmatrix} \sigma_{E_i}^2 & \sigma_{EC_i} \\ \sigma_{EC_i} & \sigma_{C_i}^2 \end{pmatrix}\right).$$

This model may be rewritten as $Y_{E_i} = \theta_{E_i} + \beta_{obs_i}(Y_{C_i} - \theta_{C_i}) + \varepsilon_{E_i}$ and $Y_{C_i} = \theta_{C_i} + \varepsilon_{C_i}$, where ε_{E_i} and ε_{C_i} are independent observation errors with $\varepsilon_{E_i} \sim N(0, \sigma_{E_i}^2)$ and $\varepsilon_{C_i} \sim N(0, \sigma_{C_i}^2)$. The term $\beta_{obs_i} = \sigma_{EC_i}/\sigma_{C_i}^2$ is the within-trial regression coefficient. When the dependent variable is the treatment rate, $\beta_{obs_i} = 0$ because the observed treatment rate and observed control rate are independent. For any of the other treatment effect measures, $\beta_{obs_i} \neq 0$ because of the induced correlation between the observed treatment effect and the observed control rate.

McIntosh[32] showed that regression using the observed Y_{E_i} and Y_{C_i} gives a biased estimate of β_{true} and that the expected bias may be either positive or negative. Only small bias occurs when the within-study variance of the control rate, σ_C^2, is small relative to the between-study variance, τ_C^2, but large bias often occurs when τ_C^2 is small because the control rates do not display sufficient variation to fit a regression line.

We estimate the structural model parameters $\{\mu_E, \mu_C, \beta_{true}, \tau_E^2, \tau_C^2\}$ in the model by an EM algorithm in which we treat the unobserved parameters θ_{E_i} and θ_{C_i} as missing data. Conditional on the current estimates of the structural model parameters, the E-step estimates the expected values of the mean and variance of the θ_{E_i} and θ_{C_i}. The M-step then updates iteratively the structural model parameters via standard bivariate normal maximum likelihood equations.

3. DATA ANALYSIS

To compare the hierarchical model with weighted least squares analysis of the observed data empirically and to examine other aspects of control rate meta-regression, we compiled a database of all meta-analyses of randomized controlled trials (RCTs) from the 1994 Cochrane Collaboration database of pregnancy and childbirth[33] and from the 1990–1995 issues of seven major medical journals that have published meta-analyses: *Annals of Internal Medicine*; *Archives of Internal Medicine*; *British Medical Journal*; *Circulation*; *Journal of the American Medical Association*; *Lancet*, and *New England Journal of Medicine*. Containing about 500 systematic overviews, the Cochrane database is the most comprehensive database of meta-analyses of RCTs in any one clinical field. The complete selection of meta-analyses from the seven journals broadens the clinical scope of the analysis, while attempting to ensure that the included meta-analyses are of high quality and maximal relevance.

For inclusion in our database, each meta-analysis had to satisfy the following criteria: (i) use of binary outcomes reported for two treatment groups; (ii) data available in the text of the document or already in our possession; (iii) inclusion of at least six RCTs each with at least one event in the control group; and (iv) report of at least five events on average in the control group (or, if the event rate was high, at least five non-events). We applied the third and fourth criteria so that the maximum likelihood estimates based on the assumptions of normally distributed random error in the hierarchical model would be well-estimated by the EM algorithm.[32] We excluded from its meta-analysis any RCT with fewer than 10 patients in either the treatment group or the control group before we applied the above inclusion criteria. Because some meta-analyses looked at multiple outcomes, we chose the outcome with the greatest clinical relevance, preferably mortality, as determined by one of us (JL) blinded to the results among those meeting the inclusion criteria. We excluded any outcome that related to patient or physician behaviour or to a choice of medical procedure (for example, Caesarean section) so that the control rate is directly related to patient health.

The final collection of meta-analyses included 45 from the Cochrane database[33] and 70 from the seven major medical journals.[14,21,34–68] Table I categorizes these two sources according to the size of the meta-analysis expressed by either the number of RCTs included or the total number of patients enrolled in the combined studies. The meta-analyses from the Cochrane database were generally smaller, tending to involve fewer patients and fewer studies.

For each meta-analysis, we fit the observed data model by weighted least squares weighting by the inverse of the variance of the dependent variable and the hierarchical model by EM using the risk difference, log relative risk and log odds ratio as dependent variables and the control rate as the independent variable. For each outcome, we recorded the slope estimate, its standard error, the z-test statistic defined as the ratio of the estimate to its standard error and whether this statistic reached statistical significance, defined as greater than 1·96 in absolute value. In addition, we took the estimate of the control rate and its variance from the hierarchical model, and also

Table I. Description of the size of the 45 selected meta-analyses in the Cochrane database and the 70 selected meta-analyses from seven major medical journals

Number of studies	Cochrane	Medical journals
6–9	29	22
10–14	12	21
15–19	4	12
20–62	0	15
Median	*8*	*11·5*
Number of patients		
2000 or less	24	24
2001–5000	13	13
5001–10 000	6	13
more than 10 000	2	20
Median	*1835*	*4396*
Median number per RCT	*177*	*265*

obtained the random effects estimate of the pooled treatment effects[14,69] from a standard meta-analysis that did not consider the control rate.

In the following analysis, we excluded three meta-analyses whose control rate variance estimates from the hierarchical model were less than 0·0001 because they gave unstable regression estimates. All three were from the Cochrane database and each consisted of six RCTs. For these three meta-analyses, the EM algorithm converged to extremely large slope or slope variance estimates or to negative slope variance estimates. This is not surprising because these small variance estimates indicate that the true control rates in the individual RCTs in these meta-analyses vary only within a narrow range, precluding reliable estimation of a regression function on the true control rates.

3.1. Slope Estimates: Comparison of Hierarchical and Observed Data Models

3.1.1. Risk Ratio

We first consider analyses under the observed data model and the hierarchical model using the log risk ratio as the outcome. Table II(*a*) shows that 15 of 112 (13 per cent) meta-analyses gave a significant control rate slope (either positive or negative) by the hierarchical model compared with 23 of 112 (21 per cent) by the observed data model. More often than not, the significant slopes were negative, indicating that the treatment was more likely to show benefit as the control rate increases. Nineteen of the 23 results significant under the observed data model showed a negative correlation, whereas 9 of 15 significant under the hierachical model did. There were 19 meta-analyses in which the sign of the slope differed under the two models, but none of these disagreed as to statistical significance. In 18 of the 19, the observed data slope was negative. Therefore, there were no meta-analyses in which the two models gave significant results in opposite directions.

Table II. Cross-tabulation of the significance of the control rate slope as fit by the observed data model (WLS) and the hierarchical model (EM). Each estimate is classified as significant and negative ($Z \leqslant -1{\cdot}96$) labelled 'Sig −', non-significant ($-1{\cdot}96 < Z < 1{\cdot}96$) labelled 'NS' or significant and positive ($Z \geqslant 1{\cdot}96$) labelled as 'Sig +'. Below each table are the results of the Jonckheere–Terpstra exact test[70] for testing if significance of the WLS results are independent of the significance of the EM results

(*a*) *Risk ratio*

		WLS				
		Sig −	NS −	NS +	Sig +	Total
EM	Sig −	8	1	0	0	9
	NS −	11	50	1	0	62
	NS +	0	18	17	0	35
	Sig +	0	0	2	4	6
	Total	19	69	20	4	112

Jonckheere–Terpstra exact test statistic = 8·145 ($p < 0{\cdot}0001$)

(*b*) *Odds ratio*

		WLS				
		Sig −	NS −	NS +	Sig +	Total
EM	Sig −	13	1	0	0	14
	NS −	19	40	0	0	59
	NS +	0	18	19	0	37
	Sig +	0	1	0	1	2
	Total	32	60	19	1	112

Jonckheere–Terpstra exact test statistic = 7·978 ($p < 0{\cdot}0001$)

(*c*) *Risk difference*

		WLS				
		Sig −	NS −	NS +	Sig +	Total
EM	Sig −	33	1	0	0	34
	NS −	26	26	0	0	52
	NS +	1	14	10	0	25
	Sig +	0	0	0	1	1
	Total	60	41	10	1	112

Jonckheere–Terpstra exact test statistic = 7·917 ($p < 0{\cdot}0001$)

The overall rank correlation between the z-statistics for the slopes of the two models was 0·83. When the hierarchical model indicated that the control rate was significantly correlated with the treatment effect, the observed data model also gave a significant result in 12 of 15 meta-analyses. Among the three meta-analyses with significant hierarchical model slopes, but non-significant observed data model slopes, the hierarchical and obsereved data slopes had the same sign (two positive and one negative). About half of the significant results found with the observed data

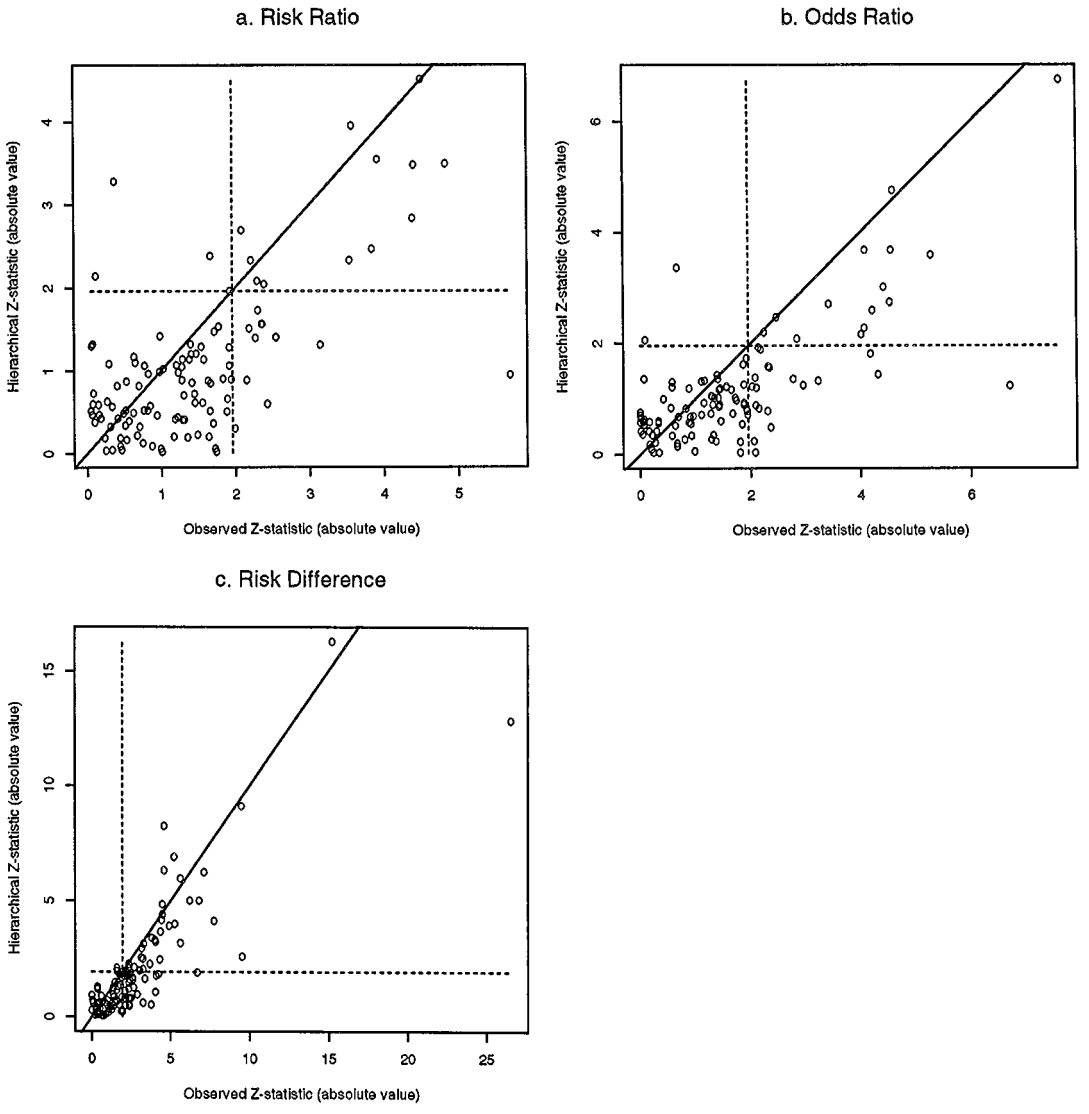

Figure 2. Plots of the absolute value of the z-statistics for testing the significance of the slope of the control rate meta-regression for the hierarchical model (y-axis) versus the observed data model (x-axis) using three treatment effect measures. Solid line is line of equality ($y = x$). Dotted lines are drawn at $x = 1{\cdot}96$ and $y = 1{\cdot}96$ to indicate regions of statistical significance

model, however, were actually false positives (11 of 23). All 11 of these false positives had negative slopes. As a test for significance of the control rate as a predictor of the treatment effect using the hierarchical model results as the gold standard, weighted least squares with observed data had a high negative predictive value (86/89 or 0·97), but a low positive predictive value (12/23 or 0·52). Its sensitivity was 12/15 (0·80) compared with a specificity of 86/97 (0·89). Figure 2(*a*) shows these results graphically, plotting the absolute values of the two sets of z-statistics.

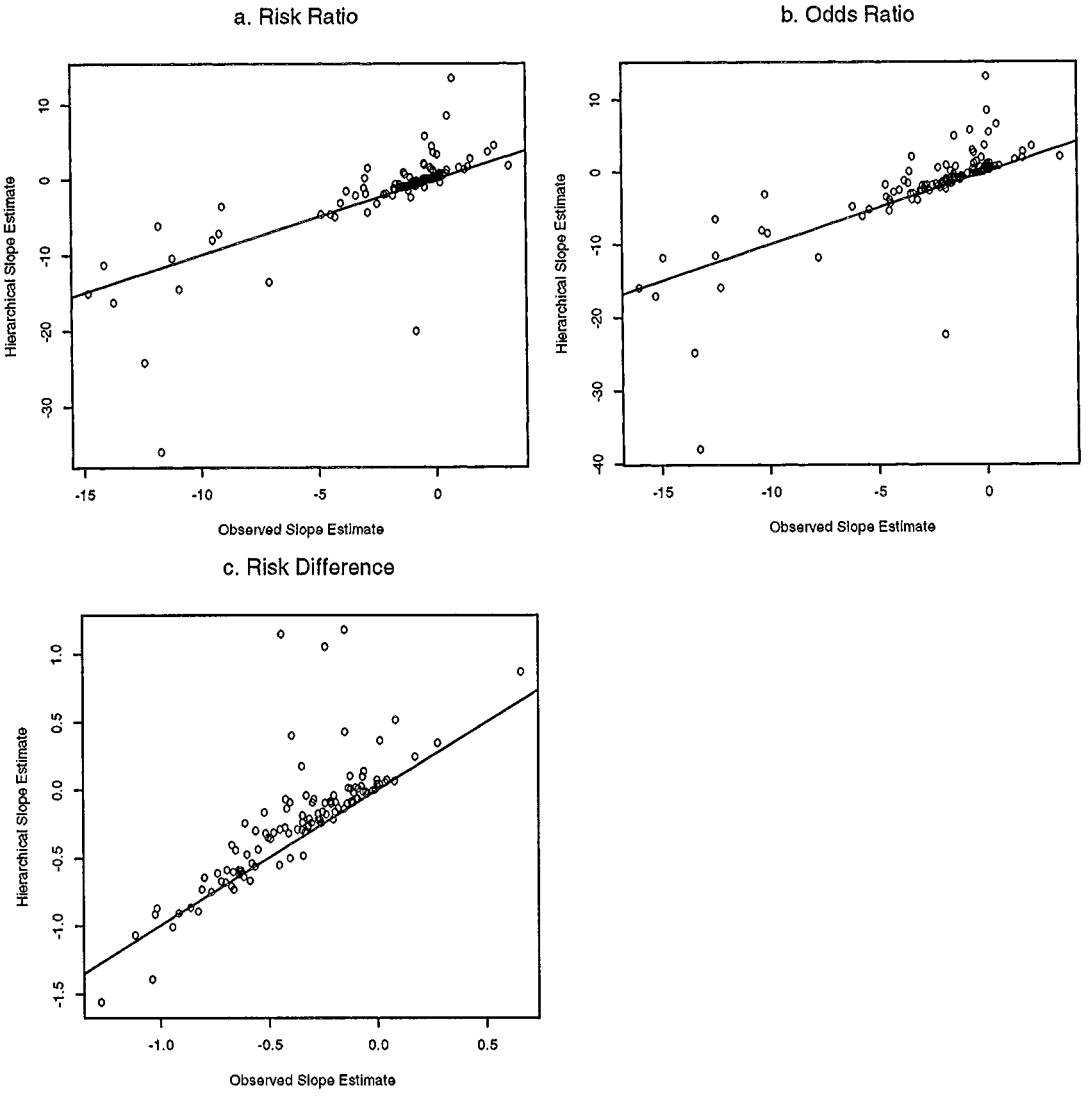

Figure 3. Slope estimates for hierarchical model (y-axis) versus observed data model (x-axis) for three treatment effect measures. Solid line indicates equal slopes

Figures 3(a) and 4(a) show that the slope estimates and standard errors from the two models are highly correlated (both correlations exceed 0·85 and the exact test of independence[70] in Table II is highly significant), but also that the hierarchical model gave higher standard error estimates and lower point estimates for the slope. Failure to account for the control rate variation using weighted least squares leads to underestimation of the uncertainty in the slope estimate as well as to a slight measurement bias toward the null in the slope estimates.

The z-test statistics associated with the test of significance for the control rate slopes measured by the hierarchical model were not related to the corresponding test statistics of the treatment

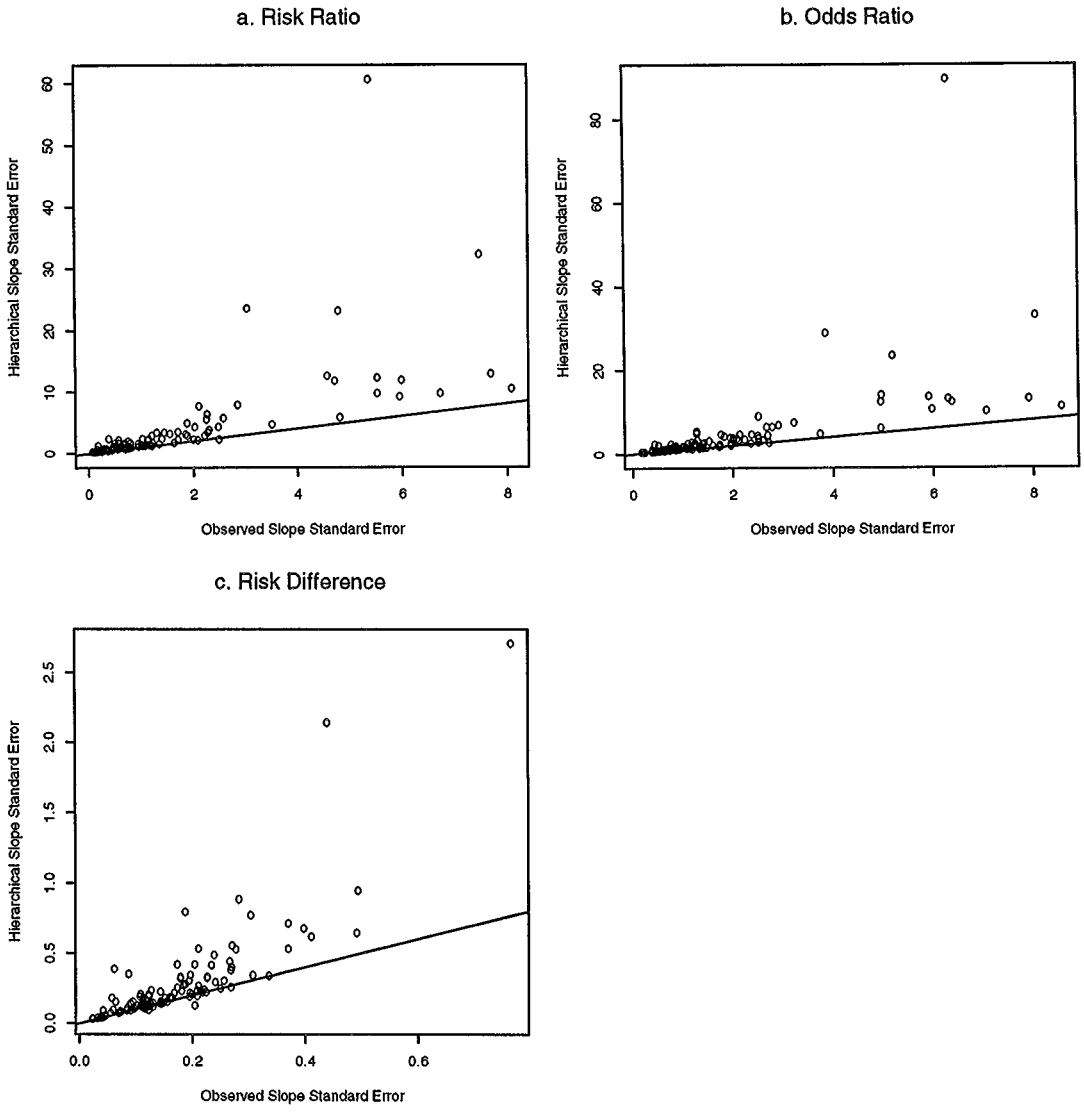

Figure 4. Standard error estimates for hierarchical model (y-axis) versus observed data model (x-axis) for three treatment effect measures. Solid line indicates equal standard errors

effects measured by the random effects risk ratio ($r = -0{\cdot}09$). Fifteen per cent (9 of 59) of meta-analyses with a significant treatment effect had a significant control rate slope compared with 11 per cent (6 of 53) of meta-analyses with a non-significant treatment effect ($p = 0{\cdot}54$). We found a similar lack of agreement when we fit the control rate slope by weighted least squares using the observed data.

Figure 5 gives an example, based on a meta-analyses of 34 studies of streptokinase for the treatment of myocardial infarction, of how control-rate regression based on the observed data

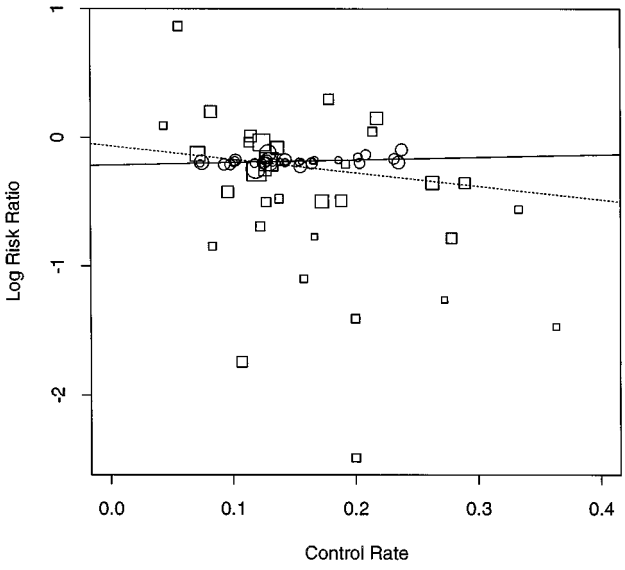

Figure 5. Log risk ratio versus control rate for 34 studies of streptokinase therapy for acute myocardial infarction. The squares denote the observed rates and the circles represent the hierarchical model estimates of the true rates. The dotted line is the regression using WLS on the observed data; the solid line is the regression from the hierarchical model using EM. The size of the symbols is proportional to the natural logarithm of the total size of the study.

can be misleading. While the plot of observed log risk ratio versus observed control rate appears to suggest a negative correlation between the treatment effect and control rate, the plot of the true treatment effect and control rate using the shrunken estimates from the hierarchical model shows that this correlation is spurious.

3.1.2. Odds Ratio

Table II(*b*) and Figures 2–4 give the results using the log odds ratio as the outcome. Results are very similar to those for the risk ratio except that the number of meta-analyses that show a significant control rate slope from the hierarchical model is slightly higher (16/112 = 14 per cent) and the significant slopes are more likely to be negative. Among the significant slopes identified by application of weighted least squares to the observed data model are many not truly significant (sensitivity, 0·88 (14/16); specificity, 0·80 (77/96); negative predictive value, 0·97 (77/79); positive predictive value, 0·42 (14/33)).

Again, significant control rate slopes did not correlate with significant treatment effects. Seventeen per cent (10 of 59) of meta-analyses with a significant treatment effect had a significant control rate slope compared with 11 per cent (6 of 53) of meta-analyses with a non-significant treatment effect ($p = 0{\cdot}72$). The correlation between the control rate slopes and the pooled treatment effect estimates was 0·15.

3.1.3. Risk Difference

Table II(*c*) and Figures 2–4 show results using the risk difference. Again, because the observed data model underestimated the variance of the slope estimates, the number of significant slopes

were nearly doubled relative to the hierarchical model (61 (54 per cent) versus 35 (31 per cent)), even though the correlation between their z-statistics was 0·89. The observed data model again performed well as a screening test with a high negative predictive value of 0·98 (50/51), but a low positive predictive value of 0·56 (34/61). The sensitivity was 0·97 (34/35) and the specificity was 0·65 (50/77).

Significant treatment effects were much more closely related to significant control rate slopes on the risk difference scale. Among meta-analyses with a significant pooled treatment effect, 28 of 57 (49 per cent) also had a significant control rate slope compared with only 7 of 55 (13 per cent) among meta-analyses without a significant pooled treatment effect ($p < 0{\cdot}0001$). The correlation between the control rate slopes and the pooled treatment effect estimates was 0·54 ($p < 0{\cdot}0001$).

3.1.4. Comparison of Treatment Effect Measures

Compared with the analysis on the log risk ratio scale, the analysis on the risk difference scale gave significant control rate slopes more than 2·5 times as often by either model (hierarchical, 35 versus 15; observed data, 61 versus 23). Some overlap between the two scales did occur: 7 of 15 cases significant on the risk ratio scale were also significant on the risk difference scale with the hierarchical model, while the corresponding numbers were 19 of 23 with the observed data model. Conversely, only 7 of 35 slopes significant on the risk difference scale were significant on the risk ratio scale using the hierarchical model (19 of 61 for the observed data model). Significant correlations between the treatment effect and the control rate were even more likely to be negative on the risk difference scale than on the risk ratio or odds ratio scales (60 of 61 by the observed data model and 34 of 35 by the hierarchical model).

3.2. Subanalyses

3.2.1. Relationship to variance of true control rates

One very useful feature of the hierarchical model is that it provides an estimate of the variability of the true unobserved control rates. Treatment effects can only be differentiated by control rates that have sufficient variation. Otherwise, regression slope and standard error estimates become widely inflated. Meta-analyses based on a small number of RCTs are particularly prone to this problem. Because a test of heterogeneity based on the observed control rates ignores the distinction between measurement error and true variation, the hierarchical model estimate is more appropriate. We found that when the variance of the unobserved true control rates (τ_C^2) was very small, only rarely was the treatment effect related to the control rate. Among the 14 meta-analyses with $\tau_C^2 < 0{\cdot}001$ (that is, standard deviation less than about 0·03), the control rate was significantly correlated once with the risk ratio and risk difference and never with the odds ratio. This reflects the strong negative correlations found between the control rate variance estimate and the variance of the slope estimate ($-0{\cdot}80$ for RR; $-0{\cdot}81$ for OR; $-0{\cdot}51$ for RD).

3.2.2. Relationship to size of meta-analysis and length of follow-up

Table III classifies the control rate significance by the type of meta-analysis, subdividing by the number of trials in the meta-analysis (where the cut-off of 10 roughly divides the total in half) and by the duration of follow-up (all short-term of about the same duration, heterogeneous mix of different amounts of follow-up and long-term). Studies with homogeneous short-term follow-up and studies with 10 or more RCTs more often showed a significant relationship on the risk ratio

Table III. Number (per cent) of meta-analyses with significant control rate slopes classified by effect measure, number of studies and duration of follow-up

	Number of studies < 10			Number of studies $\geqslant 10$			All
	Short follow-up	Mixed follow-up	Long follow-up	Short follow-up	Mixed follow-up	Long follow-up	
Total MAs	37	7	4	36	19	9	112
RD	12 (32)	3 (43)	0	12 (33)	6 (32)	2 (22)	35 (30)
RR	4 (11)	0	0	10 (28)	0	1 (11)	15 (13)
OR	5 (14)	0	0	10 (38)	1 (5)	0	16 (14)

and odds ratio scales. Of the 36 meta-analyses with both of these characteristics the control rate was significantly related to the treatment effect for 10 (28 per cent) on the risk ratio and odds ratio scales, compared with 5 (7 per cent) and 6 (8 per cent), respectively, among the remaining 76 meta-analyses. This difference was not present on the risk difference scale, however, where 33 per cent of the short-term, homogeneous follow-up meta-analyses showed significant correlations of control rate with treatment effect compared with 30 per cent of the remaining meta-analyses. The small number of meta-analyses with long follow-up and the potential confounding of duration of follow-up and number of trials by the disease and treatment studied preclude drawing any definitive conclusions, however.

4. DISCUSSION

Meta-analysts disagree about how to handle between-study heterogeneity. Some advocate using homogeneous collections and fixed-effects analyses so that the pooled effect has a precise meaning.[71] Others argue that sources of between-study variation provide critical information that must be uncovered.[10] Identifying these sources is complicated by the lack of information historically included in clinical trial reports, although efforts are now being made to improve such summaries.[72]

The control rate may serve as a useful partial surrogate for this missing information because of its ready availability and reflection of study or patient heterogeneity. Any non-random variation in treatment effects across studies may relate to study or patient differences that could be reflected in the control rates. A few of the many possible sources of this type are length of follow-up, concommitant drug dosage and subject mix.

4.1. Use of a hierarchical model

In order not to introduce spurious associations between the treatment effect and the control rate when none exists, it is necessary when using the control rate as a predictor for treatment effects to adjust for the random nature of the control rate and for the functionally induced correlation between the observed treatment effect and observed control rate. To measure this bias empirically, we compare the fit of a hierarchical model that incorporates these features with that of the observed data model.

As expected, we found more frequent significant associations with the non-hierarchical observed data model, regardless of the treatment effect measure. Because the model does not take

into account all the sources of variability introduced by the stochastic control rates and therefore underestimates the standard errors of the slope estimates, about half of the significant associations it found were false positive. Conversely, significant results found with the hierarchical model were usually also significant with the observed data model.

Using the hierarchical model, we found significant correlations between the treatment effect and the control rate about 1 in 6 times on the risk ratio or odds ratio scales and about 1 in 3 times on the risk difference scale. When found, these correlations generally indicated that treatments were more effective on those populations with higher control rates, thus suggesting that the treatments might have worked better on the sicker patients.

These findings, however, do not necessarily reflect what would be found in a complete survey of all meta-analyses because the meta-analyses used come from a highly selected group of high quality meta-analyses. Lower quality meta-analyses may be more prone to exaggerated observed effects that would change the rates reported in this study.

4.2. Risk ratio or odds ratio preferable to risk difference

Our results show that significant correlations between the treatment effect and the control rate are much more frequent when we measure the treatment effect by the risk difference than when we measure by either the log risk ratio or log odds ratio. Studies with low control rates rarely have high treatment rates unless the treatment is toxic; studies with high control rates often have low treatment rates when the treatment is beneficial. Therefore, the size of the difference between the treatment rate and the control rate often increases as the control rate increases. On multiplicative scales such as the risk ratio and odds ratio, however, this arithmetic relationship does not necessarily induce any correlation with the size of the control rate.

This difference in frequency across treatment effect measures occurred despite the fact that the choice of treatment effect measure did not affect the frequency with which we found that the pooled effect differed significantly from the null value. In particular, we noted that on the risk ratio and odds ratio scales the significance of the control rate as a predictor of treatment effect was unrelated to whether or not the pooled treatment effect was significant. On the risk difference scale, though, we found control rate correlations much more often when the pooled treatment effects were significant. The increased number of control rate correlations with the risk difference and their tendency to occur more often when the pooled effect was significant suggest that the risk difference scale may overestimate the influence of the control rate.

4.3. Implications for clinical trial design

When the size of the treatment effect is linearly related to the size of the control rate, the correlation can mean either that treatment effects differ in magnitude but not sign for different subgroups of individuals (a quantitative interaction) or that treatment effects change sign across subgroups, showing benefit to some and harm to others (a qualitative interaction). The majority of meta-analyses that show correlation between the treatment effect and the control rate also have significant pooled treatment effects and thus suggest quantitative interaction. Those meta-analyses in which the pooled treatment effect is not significant, but the control rate is related to the treatment effect, suggest qualitative interactions and are more problematic. Peto and colleagues, in fact, discount qualitative interactions as spurious correlations and artifacts of the data.[71]

'Mega' trials such as GISSI-3[73] and ISIS-4[25] have addressed the issue of differential treatment effects through factorial designs and planned subgroup analyses. Though the results of these very large trials often agree with those of a meta-analysis of small trials, differences between them can sometimes relate to control rate variation.[74] The experience with magnesium and nitrates in the treatment of acute myocardial infarction is an example in which smaller studies showed large benefits of treatment that mega-trials did not support. We reported previously that one could have predicted the lack of benefit evidenced by the mega-trials with low control rates by the regression of the risk ratio on the control rate which demonstrated that trials with high control rates had lower risk ratios.[74]

This discussion suggests that the pooled effect may not necessarily be the best estimate of treatment efficacy for the individual patient. If control rate heterogeneity reflects differences in the average baseline risk between different patient populations, then the associations found between treatment effects and control rates indicate that some groups of patients respond better to some treatments than others do. It may not always be easy to find these subgroups, but, as an exploratory technique, control rate meta-regression may suggest study conditions for further examination. These explorations are only possible, of course, if the clinical trials provide different subgroups. Designing trials to be different may promote a better synthesis.

We have noted that correlations between control rates and treatment effects were more frequent in meta-analyses with homogeneous, short-term follow-up and in meta-analyses based on a large number of trials (Table III). One possible reason for this finding is that such meta-analyses exhibit the most variation among the control rates and therefore provide the widest range over which to measure this correlation. In trials with long-term follow-up, control rates will tend to converge as competing risks even out differences in the study cohorts. In collections of trials with heterogeneous follow-up, control rate differences based on cohort characteristics will become diluted by control rate differences based on variable follow-up. Varying follow-up can also dilute treatment effects if these diminish over time and therefore can complicate the process of finding correlations with control rates. Control rate meta-regression would then seem most useful when the control rates differ substantially because of study cohort features that could modify treatment effects.

It is essential that we interpret clinical trial results in the light of all other similar studies, and meta-analysis in its broader context is essential for this understanding. For these interpretations to be possible, clinical trialists must design their studies with future meta-analyses in mind[75] and must report in a structured manner detailed information about subpopulations and ideally individuals.[72] Such reporting will help identify any risk factors that affect control rate heterogeneity for which the control rate may be a surrogate predictor.

5. CONCLUSION

We need not view heterogeneity of treatment effect as an argument against the pooling of clinical trial data. When heterogeneity does not arise because of inadequate design or incomplete publication of results, we should welcome it as an opportunity to optimize treatment benefit. Regression analysis with the control rate as a predictor of treatment effects presents a useful tool to explore such heterogeneity. Though correlations between treatment effects and control rates in a series of clinical trials are overstated when using regression of observed rates, proper accounting for random error in the control rate and functional correlation between observed rates in a hierarchical model reveals that significant relationships do occur and are probably best measured as risk ratios or odds ratios rather than as risk differences. In such cases, we can

consider the control rate a surrogate for patient risk factors and study characteristics with which treatment benefits vary. Because the control rate is an aggregate measure that has no direct meaning at the patient level, we must view these regressions as exploratory tools that can help to determine whether further study into sources of heterogeneity is warranted.

REFERENCES

1. Bailar, III, J. C. 'When research results are in conflict', *New England Journal of Medicine*, **313**, 1080–1081 (1985).
2. Horwitz, R. I. 'Complexity and contradiction in clinical trial research', *American Journal of Medicine*, **82**, 498–510 (1987).
3. Chalmers, T. C., Celano, P., Sacks, H. S. and Smith, H. Jr. 'Bias in treatment assignment in controlled clinical trials', *New England Journal of Medicine*, **309**, 1356–1361 (1983).
4. Colditz, G. A., Miller, J. N. and Mosteller, F. 'How study design affects outcomes in comparisons of therapy. I: Medical', *Statistics in Medicine*, **8**, 441–454 (1989).
5. Yusuf, S., Collins, R. and Peto, R. 'Why do we need some large, simple randomized trials?', *Statistics in Medicine*, **3**, 409–420 (1984).
6. Thompson, S. G. 'Controversies in meta-analysis: the case of the trials of serum cholesterol reduction', *Statistical Methods in Medical Research*, **2**, 173–192 (1993).
7. Egger, M. and Smith, G. D. 'Misleading meta-analysis: Lessons from "an effective, safe, simple" intervention that wasn't', *British Medical Journal*, **310**, 752–754 (1995).
8. CLASP (Collaborative Low-Dose Aspirin Study in Pregnancy) Collaborative Group. 'CLASP: a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women', *Lancet*, **343**, 619–629 (1994).
9. Annello, C. and Fleiss, J. L. 'Exploratory or analytic meta-analysis: should we distinguish between them?', *Journal of Clinical Epidemiology*, **48**, 109–116 (1995).
10. Greenland, S. 'Invited commentary: A critical look at some popular meta-analytic methods', *American Journal of Epidemiology*, **140**, 290–296 (1994).
11. Greenland, S. 'Quantitative methods in the review of epidemiologic literature', *Epidemiologic Review*, **9**, 1–30 (1987).
12. Berlin, J. A. and Antman, E. M. 'Advantages and limitations of metaanalytic regressions of clinical trials data', *Online Journal of Current Clinical Trials*, **3**, Document 134 (1994).
13. Kasiske, B. L., Kalil, R. S. N., Ma, J. Z., Liao, M. and Keane, W. F. 'Effect of antihypertensive therapy on the kidney in patients with diabetes: a meta-regression analysis', *Annals of Internal Medicine*, **118**, 129–138 (1993).
14. Ioannidis, J. P., Cappelleri, J. C., Lau, J., Skolnik, P. R., Melville, B., Chalmers, T. C. and Sacks, H. S. 'Early or deferred zidovudine therapy in HIV-infected patients without an AIDS-defining illness: a meta-analysis', *Annals of Internal Medicine*, **122**, 856–866 (1995).
15. Langbein, L. I. and Lichtman, A. J. *Ecological Inference*, Sage Publications, Beverly Hills, California, 1976.
16. Morgenstern, H. 'Uses of ecologic analysis in epidemiologic research', *American Journal of Public Health*, **72**, 1336–1344 (1982).
17. Light, R. J. 'Accumulating evidence from independent studies: What we can win and what we can lose', *Statistics in Medicine*, **6**, 221–228 (1987).
18. Brand, R. and Kragt, H. 'Importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials', *Statistics in Medicine*, **11**, 2077–2082 (1992).
19. Glasziou, P. and Irwig, L. 'An evidence based approach to individualising treatment', *British Medical Journal*, **311**, 1356–1359 (1995).
20. Chalmers, T. C. and Lau, J. 'Changes in clinical trials mandated by the advent of meta-analysis', *Statistics in Medicine*, **15**, 1263–1268 (1996).
21. Smith, G. D., Song, F. and Sheldon, R. A. 'Cholesterol lowering and mortality: the importance of considering initial level of risk', *British Medical Journal*, **306**, 1367–1373 (1993).
22. Boissel, J. P., Collet, J. P., Lievre, M. and Girard, P. 'An effect model for the assessment of drug benefit: example of antiarrhythmic drugs in postmyocardial infarction patients', *Journal of Cardiovascular Pharmacology*, **22**, 356–363 (1993).

23. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. 'Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction', *New England Journal of Medicine*, **321**, 406–412 (1989).
24. The Cardiac Arrhythmia Suppression Trial II Investigators. 'Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction', *New England Journal of Medicine*, **327**, 227–233 (1992).
25. ISIS-4 Collaborative Group. 'ISIS-4: A randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction', *Lancet*, **345**, 669–685 (1995).
26. Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F. and Chalmers, T. C. 'Cumulative meta-analysis of therapeutic trials for myocardial infarction', *New England Journal of Medicine*, **327**, 248–254 (1992).
27. Teo, K. K., Yusuf, S., Collins, R., Held, P. H. and Peto, R. 'Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomised trials', *British Medical Journal*, **303**, 1499–1503 (1991).
28. Woods, K. L., Fletcher, S., Roffe, C. and Haider, Y. 'Intravenous magnesium sulphate in suspected acute myocardial infarction: results of the second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2)', *Lancet*, **339**, 1553–1558 (1992).
29. Antman, E. M., Seelig, M. S., Fleischmann, K., Lau, J., Kuntz, K., Berkey, C. and McIntosh, M. W. 'Magnesium in acute myocardial infarction: Scientific, statistical, and economic rationale for its use', *Cardiovascular Drugs Therapy*, **10**, 297–301 (1996).
30. Senn, S. 'Letters to the editor (with author's reply). Importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials', *Statistics in Medicine*, **13**, 293–296 (1994).
31. Sharp, S. J., Thompson, S. G. and Altman, D. G. 'The relation between treatment benefit and underlying risk in meta-analysis', *British Medical Journal*, **313**, 735–738 (1996).
32. McIntosh, M. 'The population risk as an exploratory variable in research synthesis of clinical trials', *Statistics in Medicine*, **15**, 1713–1728 (1996).
33. Enkin, M. W., Keirse, M. J. N. C., Renfrew, M. J. and Neilson, J. P. (eds). *Pregnancy and Childbirth Module*: *Cochrane Database of Systematic Reviews* (*Cochrane Updates on Disk*), Disk Issue 2, Update Software, Oxford, England 1994.
34. Garg, R. and Yusuf, S. 'Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor trials', *Journal of the American Medical Association*, **273**, 1450–1456 (1995).
35. Imperiale, T. F. and McCullough, A. J. 'Do corticosteroids reduce mortality from alcoholic hepatitis? A meta-analysis of the randomized trials', *Annals of Internal Medicine*, **113**, 299–307 (1990).
36. Antiplatelet Trialists' Collaboration. 'Collaborative overview of randomised trials of antiplatelet therapy – I Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients', *British Medical Journal*, **308**, 81–106 (1994).
37. Antiplatelet Trialists' Collaboration. 'Collaborative overview of randomised trials of antiplatelet therapy – II: Maintenance of vascular graft or arterial patency by antiplatelet therapy', *British Medical Journal*, **308**, 159–168 (1994).
38. Antiplatelet Trialists' Collaboration. 'Collaborative overview of randomised trials of antiplatelet therapy – III: Reduction in venous thrombosis and pulmonary embolism by antiplatelet prophylaxis among surgical and medical patients', *British Medical Journal*, **308**, 235–246 (1994).
39. Collins, R., Peto, R., MacMahon, S., Hebert, P., Fiebach, N. H., Eberlein, K. A., Godwin, J., Qizilbash, N., Taylor, J. O. and Hennekens, C. H. 'Blood pressure, stroke, and coronary heart disease. Part 2: Short-term reductions in blood pressure: overview of randomised drug trials in their epidemiological context', *Lancet*, **335**, 827–838 (1990).
40. Insua, J. T., Sacks, H. S., Lau, T. S., Lau, J., Reitman, D., Pagano, D. and Chalmers, T. C. 'Drug treatment of hypertension in the elderly: a meta-analysis', *Annals of Internal Medicine*, **121**, 355–362 (1994).
41. Yudkin, P. L., Ellison, G. W., Ghezzi, A., Goodkin, D. E., Hughes, R. A., McPherson, K., Mertin, J. and Milanese, C. 'Overview of azathioprine treatment in multiple sclerosis', *Lancet*, **338**, 1051–1055 (1991).
42. Pagliaro, L., D'Amico, G., Sorensen, T. I., Lebrec. D., Burroughs, A. K., Morabito, A., Tine, F., Politi, F. and Traina, M. 'Prevention of first bleeding in cirrhosis. A meta-analysis of randomized trials of nonsurgical treatment', *Annals of Internal Medicine*, **117**, 59–70 (1992).

43. Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V. and Mosteller, F. 'Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature', *Journal of the American Medical Association*, **271**, 698–702 (1994).
44. Gould, A. L., Rossouw, J.E., Santanello, N. C., Heyse, J. F. and Furberg, C. D. 'Cholesterol reduction yields clinical benefit. A new look at old data', *Circulation*, **91**, 2274–2282 (1995).
45. Wang, P. H., Lau, J. and Chalmers, T. C. 'Meta-analysis of effects of intensive blood-glucose control on late complications of type I diabetes', *Lancet*, **341**, 1306–1309 (1993).
46. Selective Decontamination of the Digestive Tract Trialists' Collaborative Group. 'Meta-analysis of randomised controlled trials of selective decontamination of the digestive tract', *British Medical Journal*, **307**, 525–532 (1993).
47. Leizorovicz, A., Simonneau, G., Decousus, H. and Boissel, J. P. 'Comparison of efficacy and safety of low molecular weight heparins and unfractionated heparin in initial treatment of deep venous thrombosis: a meta-analysis', *British Medical Journal*, **309**, 299–304 (1994).
48. Laine, L. and Cook, D. 'Endoscopic ligation compared with sclerotherapy for treatment of esophageal variceal bleeding. A meta-analysis', *Annals of Internal Medicine*, **123**, 280–287 (1995).
49. Early Breast Cancer Trialists' Collaborative Group. 'Effects of radiotherapy and surgery in early breast cancer. An overview of the randomized trials', *New England Journal of Medicine*, **333**, 1444–1455 (1995).
50. Beasley, C. M. Jr, Dornseif, B. E., Bosomworth, J. C., Sayler, M. E., Rampey, A. H. Jr, Heiligenstein, J. H., Thompson, V. L., Murphy, D. J. and Masica, D. N. 'Fluoxetine and suicide: a meta-analysis of controlled trials of treatment for depression', *British Medical Journal*, **303**, 685–692 (1991).
51. Sacks, H. S., Chalmers, T. C., Blum, A. L. Berrier, J. and Pagano, D. 'Endoscopic hemostasis. An effective therapy for bleeding peptic ulcers', *Journal of the American Medical Association*, **264**, 494–499 (1990).
52. Early Breast Cancer Trialists' 'Collaborative Group. 'Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy: 133 randomised trials involving 31000 recurrences and 24000 deaths among 75000 women', *Lancet*, **339**, 1–15, 71–85 (1992).
53. Abrutyn, E. and Berlin, J. A. 'Intrathecal therapy in tetanus. A meta-analysis', *Journal of the American Medical Association*, **266**, 2262–2267 (1991).
54. Leizorovicz, A., Haugh, M. C., Chapuis, F. R., Samama, M. M. and Boissel, J. P. 'Low molecular weight heparin in prevention of perioperative thrombosis', *British Medical Journal*, **305**, 913–920 (1992).
55. Prostate Cancer Trialists' Collaborative Group. 'Maximum androgen blockade in advanced prostate cancer: an overview of 22 randomised trials with 3283 deaths in 5710 patients', *Lancet*, **346**, 265–269 (1995).
56. Furberg, C. D., Psaty, B. M. and Meyer, J. V. 'Nifedipine. Dose-related increase in mortality in patients with coronary heart disease', *Circulation*, **92**, 1326–1331 (1995).
57. Williams, R. L., Chalmers, T. C., Stange, K. C., Chalmers, F. T. and Bowlin, S. J. 'Use of antibiotics in preventing recurrent acute otitis media and in treating otitis media with effusion. A meta-analytic attempt to resolve the brouhaha', *Journal of the American Medical Association*, **270**, 1344–1351 (1993).
58. Michels, K. B. and Yusuf, S. 'Does PTCA in acute myocardial infarction affect mortality and reinfarction rates? A quantitative overview (meta-analysis) of the randomized clinical trials', *Circulation*, **91**, 476–485 (1995).
59. Naylor, C. D. and Jaglal, S. B. 'Impact of intravenous thrombolysis on short-term coronary revascularization rates. A meta-analysis', *Journal of the American Medical Association*, **264**, 697–702 (1990).
60. Coplen, S. E.. Antman, E. M., Berlin, J. A., Hewitt, P. and Chalmers, T. C. 'Efficacy and safety of quinidine therapy of maintenance of sinus rhythm after cardioversion. A meta-analysis of randomized control trials', *Circulation*, **82**, 1106–1116 (1990).
61. Advanced Ovarian Cancer Trialists Group. 'Chemotherapy in advanced ovarian cancer: an overview of randomised clinical trials', *British Medical Journal*, **303**, 884–893 (1991).
62. Pignon, J. P., Arriagada, R., Ihde, D. C., Johnson, D. H., Perry, M. C., Souhami, R. L., Brodin, O., Joss, R. A., Kies, M. S. and Lebeau, B., Onoshi, T., Osterlind, K., Tattershall, M. H. N. and Wagner, H. 'A meta-analysis of thoracic radiotherapy for small-cell lung cancer', *New England Journal of Medicine*, **327**, 1618–1624 (1992).
63. Hommes, D. W., Bura, A., Mazzolai, L., Buller, H. R. and ten Cate, J. W. 'Subcutaneous heparin compared with continuous intravenous heparin administration in the initial treatment of deep vein thrombosis. A meta-analysis', *Annals of Internal Medicine*, **116**, 279–284 (1992).

64. Anderson, I. M. and Tomenson, B. M. 'Treatment discontinuation with selective serotonin reuptake inhibitors compared with tricyclic antidepressants: a meta-analysis', *British Medical Journal*, **310**, 1433–1438 (1995).
65. Non-small Cell Lung Cancer Collaborative Group. 'Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials', *British Medical Journal*, **311**, 899–909 (1995).
66. Wells, P. S., Lensing, A. W. and Hirsh, J. 'Graduated compression stockings in the prevention of postoperative venous thromboembolism. A meta-analysis', *Archives of Internal Medicne*, **154**, 67–72 (1994).
67. Langhorne, P., Williams, B. O., Gilchrist, W. and Howie, K. 'Do stroke units save lives?', *Lancet*, **342**, 395–398 (1993).
68. Glasziou, P. P. and Mackerras, D. E. 'Vitamin A supplementation in infectious diseases: a meta-analysis', *British Medical Journal*, **306**, 366–370 (1993).
69. DerSimonian, R. and Laird, N. 'Meta-analysis in clinical trials', *Controlled Clinical Trials*, **7**, 177–188 (1986).
70. *SatXact 3 For Windows*: *Statistical Software for Exact Nonparametric Inference User Manual*, MA, CYTEL Software Corporation, Cambridge, 1995.
71. Peto, R., Collins, R. and Gray, R. 'Large-scale randomized evidence: large, simple trials and overviews of trials', *Journal of Clinical Epidemiology*, **48**, 23–40 (1995).
72. Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F. Simel, D. and Stroup, D. F. 'Improving the quality of reporting of randomized controlled trials: The CONSORT statement', *Journal of the American Medical Association*, **276**, 637–639 (1996).
73. Gruppo Italiano per lo Studio della Sopravvivenza nell'infarto Miocardico. 'GISSI-3: effects of lisinopril and transdermal glycerol trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction', *Lancet*, **343**, 1115–1122 (1994).
74. Cappelleri, J. C., Ioannidis, J. P. A., Schmid, C. H., de Ferranti, S. D., Aubert, M., Chalmers, T. C. and Lau, J. 'Large trials versus meta-analysis of smaller trials: How do their results compare?', *Journal of the American Medical Association*, **276**, 1332–1338 (1996).
75. Chalmers, T. C. and Lau, J. 'Meta-analytic stimulus for changes in clinical trials', *Statistical Methods in Medical Research*, **2**, 161–172 (1993).

