Citation:

Dietz, E. & Weist, K. (2003). Meta-analysis in hospital and clinical epidemiology based on mixed generalized linear models. In R. Schulze, H. Holling & D. Böhning (Eds.), *Meta-analysis: New developments and applications in medical and social sciences* (pp. 179-195). Hogrefe & Huber.

12

Meta-Analysis in Hospital and Clinical Epidemiology based on Mixed Generalized Linear Models

Ekkehart Dietz
Working Group: Biometry and Epidemiology
Institute for International Health, Joint Center for Humanities and Health Sciences
Free University Berlin

Klaus Weist
Institute of Hygiene and Environmental Health
Free University Berlin

Summary

Meta-analyses in the area of hospital and clinical epidemiology have been done for quite some time. Typically, the quantitative part of such analyses is to provide pooled estimators of new hygiene measures or clinical interventions, respectively. If the sizes of the effect in the studies of a meta-analysis were considered to be nearly identical, then the respective pooled estimator could be interpreted as an estimate of the common effect of the new measure. Otherwise, it could be interpreted as an estimate of a mean effect. To draw practical consequences from a mean effect estimation, a description of the heterogeneity of effects is necessary. Such a description is not provided by the standard random effect estimator, which assumes normal distributed study specific effects. As an alternative, a non-parametrical random effect estimator is suggested. This estimator is based on a finite mixed generalized linear model. These models have proven to be very flexible and useful to estimate mean effect sizes and to explain heterogeneity, because they allow for non-normal random

effects and to use covariables to explain baseline and effect heterogeneity for several effect measurements. The method is illustrated using data of two meta-analyses, which have been published recently by Thompson and Sharp (1999) as well as Veenstra, Saint, Saha, Lumley, and Sullivan (1999).

12.1 INTRODUCTION

A main task of hospital epidemiology is to evaluate the hygiene regulations in hospitals with respect to hospital acquired or nosocomial infections. For example, hospital epidemiological studies have to justify certain infection control measures such as modification of central venous catheters to reduce catheter related blood stream infections. For several reasons, hospital epidemiological studies are usually small with respect to the number of patients. Hence, a single study provides only little evidence to indicate that a certain hygiene regulation is better than a standard one. Therefore, meta-analyses in this area have been done for quite some time. The quantitative part of such an analysis is to provide a pooled estimator of the effect of the new hygiene measure. If the sizes of the effect in several studies of a meta-analysis were considered to be nearly identical, such a pooled estimator could be interpreted as an estimate of the *common effect* of the new measure. It could otherwise be interpreted as an estimate of a *mean effect*. To draw practical consequences from a mean effect estimation, a description of the heterogeneity of effects is necessary.

12.1.1 The Data Base

Meta-analyses in hospital epidemiology are typically based on count data obtained in several studies. These are mostly intervention studies. The nature of the count data that can be used does not only depend on the study design but also on the available study report. Two typical types of data layout of studies are shown in Tables 12.1 and 12.2.

Table 12.1 Prevalence Data

Hygiene Regulations	Number of Patients Infected	Number of Patients
Standard	n_0	N_0
New	n_1	N_1
Σ	n	N

Incidence data are published from cohort studies, whereas prevalence data are published either from cohort studies or from cross-sectional studies. In addition to this count data, characteristics of the studies like "year of the study" and "type of hospital" are also available. Mostly, the count data are also available for subgroups of the study populations. Such subgroups are defined by cross classifications by characteristics like "hospital ward", "gender", "sever-

Table 12.2 Incidence Data

Hygiene Regulations	Number of Infections	Patient Days
Standard	n_0	N_0
New	n_1	N_1
Σ	n	N

ity of disease", and "age group". These groups are called "units of the study". Their defining characteristics are called "first order variables", whereas characteristics of the studies are called "second order variables". The binary indicator variable "hygiene regulations" (1 = new regulations, 0 = standard regulations) is an example of a first order variable. This is usually the variable of main interest. Of course, it always has to be available. Sometimes, but very rarely, both outcome and explanatory variables of individuals (patients) can be obtained, that is, units are individuals. The meta-analysis can be considered then as the evaluation of a multi-center study.

12.1.2 Effect Measurements and Baseline Heterogeneity

Let

$$R_S = \frac{n_0}{N_0} \quad \text{and} \quad R_N = \frac{n_1}{N_1}$$

denote the infection rates under the standard and the new hygiene regulations, respectively. These two quantities can be used to measure the effectiveness of the new regulations. An overview of measures that are in some use is given in Olkin (1999). The most popular ones are the logarithm of the relative risk ($\log(RR)$), the logarithm of the odds ratio ($\log(OR)$) and the risk difference (RD):

$$\begin{aligned} \log(RR) &= \log(R_N) - \log(R_S) \\ \log(OR) &= \log(R_N/(1-R_N)) - \log(R_S/(1-R_S)) \\ RD &= R_S - R_N. \end{aligned}$$

Another often used measure and one which can be derived from the risk difference is the number needed to treat (NNT):

$$NNT = 1/(R_S - R_N).$$

In cohort studies, an intuitively appealing measure of efficacy of new hygiene regulations is

$$ef = RD/R_S. \tag{12.1}$$

It is just the probability that a certain patient who would be infected in a certain period of time under standard hygiene regulations will not be infected in the same period of time under the new hygiene regulations. From Equation 12.1 it

follows that

$$OR = \frac{1 - ef - R_S \cdot (1 - ef)}{1 - R_S \cdot (1 - ef)}$$
$$RR = 1 - ef,$$
$$RD = R_S \cdot ef,$$

and

$$NNT = 1/(R_S \cdot ef).$$

Thus, one reason for heterogeneity of OR, RD, and NNT in the several studies could be *baseline heterogeneity*, which is the heterogeneity of the rates under standard regulations R_S. Baseline heterogeneity is very common in hospital epidemiologic meta-analyses. The advantage of relative risk RR is that it does not depend on baseline rates. The common efficacy ef can be estimated by the common relative risk RR without considering the baseline heterogeneity and explanatory variables of baseline heterogeneity.

In the case of cross-sectional studies, the odds ratio is preferred. Thereby, the odds ratio is considered as an estimation of the relative risk. This is justified by assuming the mean duration of an infection under the new regulations to be about the same as under the standard regulations. Let D denote the common mean duration of an infection and $R_S^{'}$ and $R_N^{'}$ the underlying incidence under the standard regulations and the new regulations, respectively. Under steady state conditions it holds

$$\frac{R_S}{1 - R_S} = R_S^{'} \cdot D$$

and

$$\frac{R_N}{1 - R_N} = R_N^{'} \cdot D.$$

Therefore,

$$RR^{'} = \frac{R_N^{'}}{R_S^{'}} = \frac{R_N \cdot (1 - R_S)}{(1 - R_N) \cdot R_S} = OR.$$

Thus, in the case of cross-sectional studies, we have an estimation of the odds ratio OR, which is an estimation of the relative risk RR in the study population. To compute asymptotical confidence intervals in the case of small sample sizes, it is advantageous to use $\log(RR)$ and $\log(OR)$ instead of RR and OR, respectively.

12.1.3 Heterogeneity of Effect Size and Standard Methods of Meta-Analysis

Let $\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_J$ denote the effect size estimates in the J studies considered and let $w_1, w_2, \ldots, w_J$ denote their respective inverse variances. Common effect size (m) and its standard error (SE) are usually estimated by

$$\hat{m} = \frac{\sum_{j=1}^{J} w_j \hat{b}_j}{\sum_{j=1}^{J} w_j}$$

and

$$SE(\hat{m}) = \left(\sum_{j=1}^{J} w_j \right)^{-\frac{1}{2}},$$

respectively. The null hypothesis $H_0 : m = 0$ is rejected if the absolute value of

$$T = \frac{\sum_{j=1}^{J} w_j \hat{b}_j}{\sqrt{\sum_{j=1}^{J} w_j}}$$

is larger than the $(1 - \alpha)$-quantile of the standard normal distribution, where α is the test level chosen (Thompson, 1993). If the assumption of a common effect size does not hold, a random effect has to be assumed. The standard random effect model is

$$b_j \sim \mathcal{N}(m, \tau^2)$$

and

$$\hat{b}_j \sim \mathcal{N}(b_j, w_j^{-1}).$$

It is assumed that the effect sizes of the selected studies are normally distributed with a certain unknown mean value m and a certain unknown variance τ^2. As a respective estimate of the mean effect size m,

$$\hat{m}^* = \frac{\sum_{j=1}^{J} w_j^* \hat{b}_j}{\sum_{j=1}^{J} w_j^*}$$

is used, where $w_j^* = (w_j^{-1} + \hat{\tau}^2)^{-1}$ and $\hat{\tau}^2$ is a suitable estimator of effect variance τ^2. Confidence intervals and significance tests of the mean effect size can be obtained in a similar fashion as for the common effect estimation. One has simply to replace the w_j by the w_j^* in the respective formulas above. As an estimator of τ^2,

$$\hat{\tau}^2 = \max\left(\frac{Q - J + 1}{\sum w_j - \sum w_j^2 / \sum w_j}\ ,\ 0\right) \tag{12.2}$$

can be used (DerSimonian & Laird, 1986). The quantity

$$Q = \sum_{j=1}^{J} w_j(\hat{b}_j - \hat{b})^2$$

in Equation 12.2 can also be used as a test statistic of heterogeneity. If the null hypothesis is true (no heterogeneity), then Q is distributed as χ^2 with $J - 1$ degrees of freedom.

The assumption of a normal distribution of the effect size in the standard random effect model is needed to theoretically justify the mean and variance estimator of the effect size. On the other hand, this assumption provides the "mean effect size" with a certain statistical meaning. However, this assumption is rather restrictive. It is usually not provable, because of the relative small number of studies in a meta-analysis. Consequently, the scientific value and the clinical relevance of the result of meta-analysis based on the standard model are limited. Therefore, more general approaches have been considered recently (Thompson & Sharp, 1999; Aitkin, 1999a; Böhning, 2000a).

In this chapter, a certain generalization of the standard method above is presented and applied to example data. The generalizations are

1 the allowance for non-normal random effects, and

2 the use of covariables to explain baseline and effect heterogeneity.

12.2 THE MODEL

Let y_{ij} denote the value of the count number observed at the ith unit in the jth study, $j = 1, 2, \ldots, J;\ i = 1, 2, \ldots, n_j$. The observations y_{ij} are assumed to be independent random variables having expectations

$$E(y_{ij}) = \mu_{ij}, \quad i = 1, \ldots n_j; \quad j = 1, \ldots, J.$$

In the case of prevalence data

$$y_{ij} \sim \text{Binomial}(\mu_{ij}, N_{ij})$$

is assumed. If the units are individuals, $N_{ij} = 1 \ \forall i, j$. In the case of incidence data

$$y_{ij} \sim \text{Poisson}(\mu_{ij})$$

is assumed. In both cases, the mean parameter μ_{ij} is linked with a linear predictor LP

$$g(\mu_{ij}) = LP_{ij}$$

by a suitable link function $g(\cdot)$, as usual in generalized linear models. For prevalence data (binomial models), we consider the linear predictor

$$LP_{ij} = \beta_1^T X_{ij} + \beta_2^T X_j + z_j + b_j x_{ij},$$

where X_{ij} is a vector of first order variables, X_j is a vector of second order variables, x_{ij} is the binary indicator of the new hygiene regulations (1 = new hygiene regulations, 0 = standard hygiene regulations), and β_1 and β_2 are unknown parameter vectors. z_j and b_j are random effects with a joint distribution

$$\begin{pmatrix} z_j \\ b_j \end{pmatrix} \sim \phi(z, b) \quad \forall j.$$

ϕ remains completely unspecified. The expectation of b_j is the mean effect size m^*, which we are particularly interested in. If there is no effect heterogeneity, then the linear predictor can be simplified by replacing b_j by the fixed effect parameter m in the linear predictor above. If the logit link function $g(\mu_{ij}) = \log(\mu_{ij}/(1 - \mu_{ij}))$ is used, then m^* is just the mean log odds ratio.

This model is very flexible and more general than the standard random effect model of meta-analysis. Baseline heterogeneity is explained by the covariables and by the random effect z_j. Effect heterogeneity is explained by first level covariables and by the random effect b_j of this model. Note, that in case of the logit link the log odds ratio of the r_1th unit versus r_0th unit of the jth study is

$$\text{logit}(\mu_{r_1 j}) - \text{logit}(\mu_{r_0 j}) = \beta_1^T (X_{r_1 j} - X_{r_0 j}) + b_j (x_{r_1 j} - x_{r_0 j}),$$

where the second term of the right side simplifies to b_j if the r_1th unit is a treatment unit and the r_0th unit is a control unit. Another possibility to explain effect heterogeneity is to augment the linear predictor by interaction terms of second level explanatory variables and the treatment indicator x_{ij}.

For Poisson models, we consider the linear predictor

$$LP_{ij} = \beta_1^T X_{ij} + \beta_2^T X_j + z_j + b_j x_{ij} + \log(N_{ij}).$$

The difference to the former linear predictor is the additional offset term $\log(N_{ij})$. Using the log link $g(\cdot) = \log(\cdot)$ leads immediately to

$$m^* = E(\log(RR)).$$

12.3 ML-ESTIMATION

To obtain maximum likelihood estimates of β_1, β_2, and the parameter vector θ of ϕ, the likelihood function

$$L(\beta_1, \beta_2, \theta) = \prod_{j=1}^{J} \int \left\{ \prod_{i=1}^{n_j} f(y_{ij} \mid z_j, b_j, \beta_1, \beta_2, X_j, X_{ij}) \right\} \phi(z_j, b_j) \partial z_j \partial b_j \quad (12.3)$$

has to be maximized , where

$$f(y_{ij} \mid z_j, b_j, \beta_1, \beta_2, X_j, X_{ij}) = f(y_{ij} \mid LP_{ij})$$

denotes the respective conditional probability density distribution of y_{ij} given the linear predictor. Because we have not specified the distribution $\phi(z_j, b_j)$, we have to look for its nonparametric estimate. For this purpose, it is sufficient to consider two-dimensional discrete distributions with less than $J + 1$ mass points

$$\phi(z, b) = \begin{cases} p_k & \text{if } \begin{pmatrix} z \\ b \end{pmatrix} = \begin{pmatrix} z_k \\ b_k \end{pmatrix}, \\ 0 & \text{otherwise.} \end{cases}$$

$$k = 1, \ldots, K; \quad K \leq J.$$

(see Aitkin, 1999a, 1999b). This distribution has $(3 \cdot K - 1)$ parameters, which are $z = (z_1, z_2, \ldots, z_K)$, $b = (b_1, b_2, \ldots, b_K)$, and $p = (p_1, p_2, \ldots, p_{K-1})$, where $p_K = 1 - \sum_{k=1}^{K-1} p_k$. When using such a distribution, Equation 12.3 simplifies to

$$L(\beta_1, \beta_2, p, b, z) = \prod_{j=1}^{J} \sum_{k=1}^{K} p_k \prod_{i=1}^{n_j} f(y_{ij} \mid LP_{ijk})$$

and the respective log likelihood function is obtained as

$$LL(\beta_1, \beta_2, p, b, z) = \sum_{j=1}^{J} \log \sum_{k=1}^{K} p_k \prod_{i=1}^{n_j} f(y_{ij} \mid LP_{ijk}),$$

where

$$LP_{ijk} = \beta_1^T X_{ij} + \beta_2^T X_j + z_k + b_k x_{ij}$$

or

$$LP_{ijk} = \beta_1^T X_{ij} + \beta_2^T X_j + z_k + b_k x_{ij} + \log(N_{ij})$$

for the binomial model and for the Poisson model, respectively. These are the likelihood function and the log likelihood function of a finite mixed generalized linear model. An EM-algorithm and respective GLIM programs to com-

pute the ML-estimation of such models are described in Dietz and Böhning (1994, 1995) as well as in Aitkin (1999b).

Note that K is an unknown parameter of the log likelihood function above. To find the ML-estimate of K, we maximize LL for a fixed sufficiently large value K. Next, we systematically reduce the value of K to that value K^-, where the maximum of LL decreases for the first time. Then, we consider $\hat{K} = K^- + 1$ as the ML-estimate and the respective estimates of $p = (p_1, p_2, \ldots, p_{\hat{K}})$, $z = (z_1, z_2, \ldots, z_{\hat{K}})$, and $b = (b_1, b_2, \ldots, b_{\hat{K}})$ as the nonparametric estimate of ϕ. An estimate of the mean effect can be obtained by

$$\hat{m^*} = \sum_{k=1}^{\hat{K}} \hat{p}_k \hat{b}_k$$

and its variance by

$$\hat{\tau}^2 = \sum_{k=1}^{\hat{K}} \hat{p}_k \hat{b}_k^2 - \left(\sum_{k=1}^{\hat{K}} \hat{p}_k \hat{b}_k \right)^2 .$$

The posterior probability that the jth study comes from the kth mixture component (C) can be computed by

$$pr(j \in C_k \mid y_{1j}, y_{2j}, \cdots, y_{n_j j}, \hat{p}, \hat{z}, \hat{b}, \hat{\beta}_1, \hat{\beta}_2) = \frac{\hat{p}_k \prod_{i=1}^{n_j} f(y_{ij} \mid \widehat{LP}_{ijk})}{\sum_{r=1}^{\hat{K}} \hat{p}_r \prod_{i=1}^{n_j} f(y_{ij} \mid \widehat{LP}_{ijr})}, \quad (12.4)$$

where

$$\widehat{LP}_{ijr} = \hat{\beta}_1^T X_{ij} + \hat{\beta}_2^T X_j + \hat{z}_r + \hat{b}_r x_{ij} + \log(N_{ij})$$

for the Poisson models and without the last term for the binomial models. These probabilities can be used to obtain a classification of the studies. Such a classification is useful not only for a description of the heterogeneity but also for identification of further explanations of the heterogeneity in addition to the explanatory variables in the model. We now illustrate the method on data of two recently published meta-analyses.

12.4 EXAMPLES

12.4.1 Central Venous Catheters

The first example is a meta-analysis to assess the efficacy of chlorhexidine-silver sulfadiazine-impregnated central venous catheters for the prevention of nosocomial catheter colonization (NCC) and catheter-related bloodstream infection, described in Veenstra et al. (1999). We will reanalyze the published NCC data. Table 12.3 contains the count data of the 12 studies included in this meta-analysis. Here, the units are the sets of impregnated catheters and the sets of non-impregnated catheters in the studies. Thus, we have two units per study, totaling 24 units. The outcome variable is the number of catheter col-

onizations identified by the same culture techniques of intravascular catheter segments.

Table 12.3 Count Data and Characteristics of 12 Studies on the Efficacy of Chlorhexidine-Silver Sulfadiazine-Impregnation of Central Venous Catheters for the Prevention of Nosocomial Catheter Colonization

	Impregnated			Non-Impregnated			Study Characteristics		
Study	n_1	N_1	MCD	n_0	N_0	MCD	YEAR	CEX	PP
1	8	137	5.1	32	145	5.3	1997	0	0
2	28	208	6.0	47	195	6.0	1997	1	0
3	4	28	6.6	10	26	6.8	1996	0	0
4	22	68	7.0	22	60	8.0	1996	-	0
5	0	14	7.0	4	12	7.0	1994	0	0
6	2	116	7.7	16	117	7.7	1996	0	2
7	60	151	8.5	82	157	9.0	1988	1	0
8	2	98	9.0	25	139	7.3	1999	1	3
9	15	124	9.6	21	127	9.1	1996	1	0
10	45	199	10.9	63	189	10.9	1994	0	0
11	16	123	11.2	24	99	6.7	1995	1	0
12	10	44	8.0	25	35	7.6	1997	1	1

Note. n_1 = number of colonized impregnated catheters, N_1 = number of impregnated catheters, MCD = mean catheter duration, n_0 = number of colonized non-impregnated catheters, N_0 = number of non-impregnated catheters, YEAR = year of study, CEX = catheter exchange, PP = patient population.

Besides the count data, several characteristics of the studies and of the units were obtained and could be used as first and second level variables in our analysis. Table 12.3 contains one first level variable, which is the mean catheter duration (MCD) in the unit, and three second level variables, which are "year of the study", a binary variable, which indicates whether catheter exchange took place within the study (CEX), and a 4-categorical characteristic of the patient population of the study (PP). Their categories are: 1 = transplant ward, 2 = surgical ward, 3 = emergency department, and 0 = other wards. The count data in Table 12.3 yield prevalence rates. Therefore, it is reasonable to use the odds ratio as measure of efficacy. Application of the standard method described in Section 12.1.3 yields an estimate of a common odds ratio $OR = 0.47$ (0.38, 0.57), where the numbers within the brackets are the lower and the upper bound of its 95% confidence interval.

Since the heterogeneity of the 12 studies is highly significant ($Q = 26.7$, $p = .005$), a common efficacy of all studies cannot really be assumed. Therefore, confidence intervals cannot be interpreted. One has to switch to a random effect model. After computing the effect heterogeneity as $\tau^2 = 0.202$ on the basis of the standard random effect model assumptions, the odds ratio estimate $OR^* = 0.39$ (0.27, 0.55) as a mean effect estimate is obtained. This es-

timate indicates an even higher effect of the chlorhexidine-silver sulfadiazine-impregnation than the common effect size estimate because its value and also its upper confidence bound is lower.

Nevertheless, the standard random effect estimate is doubtful because its assumption of a normal distributed random effect cannot be verified by the data available. If this assumption is not true, and if nothing is known about the real distribution of the odds ratios, then few practical conclusions can be drawn from this result. Generally, it holds that a mean prevention effect does not contradict the possibility that the prevention measure actually increases the infection risk in some of the studies. So, a general recommendation of the measure cannot be given.

To overcome the drawback of the standard method, the nonparametric maximum likelihood approach described in Sections 12.2 and 12.3 is used. In a first step, mixed logistic regression models with a fixed treatment effect ($b_k = m,\ \forall k$) and without further covariables were fitted. The intercept was the only random effect in these models. We call this step "analysis of baseline heterogeneity". We started with $K = 8$ mixture components and reduced this number systematically. The first increase of the deviance (decrease of log likelihood) can be observed from $K = 4$ (deviance $= 69.4$) to $K = 3$ (deviance $= 78.3$). So 4 was considered as the maximum likelihood estimate of K. The mean treatment effect estimate in this 4-component mixture model is $OR^* = 0.44$ (0.33, 0.59).

In a second step, called "analysis of effect heterogeneity", the fixed treatment effect in the model is replaced by a random effect. Again, the respective nonparametric maximum likelihood estimate of this model can be obtained by the maximum likelihood estimation of a finite mixture model. As an estimate of the number of components, $\hat{K} = 5$ with a deviance 49.1 is obtained. On the basis of the 5-component mixture model, the mean effect size estimate is $OR^* = 0.33$ (0.22, 0.49). Note that, although this approach uses weaker model assumptions, its effect estimator indicates a slightly stronger treatment effect than those obtained as fixed model estimate and as standard random effect estimate. However, as already mentioned above, it is difficult to interpret a mean effect size if nothing is known about the distribution of the random effect. One nice property of our approach is that we have an estimate of the whole random effect distribution as a byproduct. This is a finite mixture distribution with 5 components in this case. Each component has its own effect size estimate and each study can be classified into one of these components by Equation 12.4. The results are shown in Table 12.4.

Each of the 5 component effect size estimates are smaller than one, although they are not statistically significant in the third and fourth component. Thus, our meta-analysis provides some evidence that the use of chlorhexidine-silver sulfadiazine-impregnation generally reduces the risk of catheter colonization.

There are two mixture components (1 and 5) with an especially high treatment efficacy. Their odds ratio estimates are $OR_1 = 0.1$ and $OR_5 = 0.11$, respectively, whereas the odds ratios of the other components are about 0.5. Three studies (6, 8, 12) are allocated to these components. In order to describe

Table 12.4 Nonparametric ML-Estimate of the Treatment Effect Distribution and Classification of the Studies by Their Posteriori Component Membership Probability

Component (k)	$OR_k = \exp(b_k)$	CI	p_k	Allocated Studies
1	0.10	(0.03, 0.42)	.19	6, 8
2	0.44	(0.28, 0.69)	.42	1, 2, 5, 9, 11
3	0.62	(0.36, 1.04)	.21	3, 4, 10
4	0.60	(0.32, 1.14)	.09	7
5	0.11	(0.03, 0.50)	.08	12

Note. OR_k = odds ratio as effect size estimate, CI = confidence interval, p_k = posteriori component membership probability.

situations where the efficacy of the prevention measure is particularly high, one should try to characterize their study population. Study 12 is the only study which was exclusively performed in a transplant ward. The patients of the studies 6 and 8 were from a surgical ward and an emergency department, respectively. Thus, if the patient population of the studies could be considered as a representative sample of the all patients in the respective wards, then there would be some evidence that the chlorhexidine-silver sulfadiazine-impregnation should be recommended especially in transplant, surgical, and emergency wards. In order to obtain a more complete picture, we tried to explain the two kinds of heterogeneity in the studies not only by the variable "patient population" but also by all covariables available. Only two second level variables turned out to provide some significant explanation. These are the mean catheter duration in the control group ($MCD0$) and the binary indicator of the transplant ward (TU). In order to explain not only the baseline heterogeneity but also the effect heterogeneity, we additionally included the interaction terms $TU \cdot x_{ij}$ and $\Delta MCD \cdot x_{ij}$ into the model, in which x_{ij} is the treatment indicator and ΔMCD denotes the difference of the mean catheter duration between control group and treatment group. The latter quantity is a derived second level variable which serves to adjust the effect estimate for the potential bias introduced by non-zero differences.

The inclusion of these variables explains a great deal of the heterogeneity in the data. The respective nonparametric ML-estimate of the random effect model is only a two-component mixture. Table 12.5 shows the effect estimates of the covariables in this model.

Both TU and its interaction with the treatment are significant. Consequently, the transplant ward study is a main source of both baseline heterogeneity and effect heterogeneity.

Also, the baseline mean catheter duration has a significant positive effect. That is, the larger the catheter duration, the higher the risk of catheter colonization. The variable $MCD0$ can only explain baseline heterogeneity. The

Table 12.5 ML-Estimates of the Coefficients of the Covariables in a 2-Component Mixed Logistic Regression Model

Variable	Estimate	Standard Error
TU	1.156	0.531
$MCD0$	0.134	0.040
$TU \cdot x_{ij}$	−1.698	0.722
$\Delta MCD \cdot x_{ij}$	0.010	0.086

Note. TU = binary indicator of the transplant ward, $MCD0$ = mean catheter duration in the control group, x_{ij} = treatment indicator, ΔMCD = difference of the mean catheter duration between control group and treatment group.

term $\Delta MCD \cdot x_{ij}$ is not significant and provides only inconsiderable explanation of effect heterogeneity.

Table 12.6 shows the nonparametric ML-estimate of the treatment effect and the respective classification of the studies. Now, most of the studies are classified into one component with an adjusted effect estimate of $OR = 0.43$ (0.30, 0.62). Thus, some evidence has been obtained to recommend the prevention measure in the patient population of all studies in this component. The high efficacy of the catheter impregnation in the study 12 is already shown by the significance of the term $TU \cdot x_{ij}$ of the model.

Table 12.6 Nonparametric ML-Estimate of Treatment Effect Distribution Adjusted for Covariables and Classification of the Studies by Their Posteriori Component Membership Probability

Component (k)	$OR_k = \exp(b_k)$	CI	p_k	Allocated Studies
1	0.43	(0.30, 0.62)	.73	1, 2, 3, 5, 6, 8, 9, 10, 11
2	0.64	(0.37, 1.10)	.27	4, 7, 12

Note. OR_k = odds ratio as effect size estimate, CI = confidence interval, p_k = posteriori component membership probability.

The only populations where the efficacy of the prevention measure remains questionable are those of studies 4 and 7, which are allocated to the second component. The strategy of further research could be to look for specific characteristics of these two studies, which can explain their worse results.

12.4.2 Ischaemic Heart Disease Events

The data of the second example are given by Thompson and Sharp (1999). They are taken from 28 randomized trials in which ischaemic heart disease events are considered as a response variable. An ischaemic heart disease event is defined as a fatal ischaemic heart disease or a non-fatal myocardial infarction. Another response variable of these trials is the average serum cholesterol

reduction. However, this variable will be used as an explanatory variable of the effect and baseline heterogeneity.

This meta-analysis is not a very typical one in hospital and clinical epidemiology because the prevention measures studied within the trials are very different. The spectrum of applied prevention measures includes dietary interventions, drugs, and even surgery. However, this meta-analysis is very suitable to demonstrate certain potentials of our method. It will be shown that a meta-analysis can be accomplished sensibly even when the studies considered have different study factors.

Trial-specific count data and cholesterol reductions are given in Table 12.7. Also, the count data in Table 12.7 are prevalence-type data. Therefore, the odds ratio is used again as measure of the efficacy. By applying the standard method, an estimate of a common odds ratio $OR = 0.82$ (0.77, 0.88) can be obtained.

The heterogeneity of the 28 studies is highly significant ($Q = 49.1$, $p = .006$), which is expected because of the different study factors of the studies. The estimated effect variance is $\tau^2 = 0.202$ and the respective mean odds ratio estimate based on the standard random effect model is $OR^* = 0.81$ (0.72, 0.90).

Now, one could proceed as in the previous example and estimate the random effects distribution. The aim of this study is not to describe the random effect distribution but to explain the variation of the odds ratios by the variable "mean cholesterol reduction". Therefore, we computed the nonparametric ML-estimation of a random effect model with the explanatory variables "study" and "mean cholesterol reduction" and with a random treatment effect.

By assuming that the value of the mean serum cholesterol reduction is equal to zero in the control groups, this variable can be considered a first level variable. The categorical variable "study" provides the complete explanation of baseline heterogeneity. $\hat{K} = 1$ is obtained as estimate of the number of mixture components. Consequently, the respective table of the effect distribution estimate and of the study classification has one line only (see Table 12.8).

Thus, the whole effect heterogeneity is explained by the variable mean cholesterol reduction. Its estimated logistic regression coefficient is -0.479 (0.14), where the number within the brackets is the respective standard error. The adjusted common treatment effect estimate is $\ln(OR) = 0.122$ (0.10). It is not significant. The following conclusions can be drawn from these results now:

1 The heterogeneity of the effect sizes in this meta-analysis can be explained by the variable "mean cholesterol reduction".

2 The significant mean treatment effect size estimate can be explained by the cholesterol reduction attained by the prevention measures.

3 The heterogeneity of the effect size can be explained by the heterogenous effects of the several prevention measures on cholesterol reduction.

4 Serum cholesterol reduction should be a main goal for ischaemic heart disease prevention.

Table 12.7 Count Data and one Study Characteristic of 28 Clinical Trials on the Efficacy of Diverse Prevention Measures to Reduce the Risk of Ischaemic Heart Disease

	Control group		Treatment Group		Cholesterol
Trial	n_0	$N_0 - n_0$	n_1	$N_1 - n_1$	Reduction(mmol/l)
1	210	5086	173	5158	0.55
2	85	168	54	190	0.68
3	75	292	54	296	0.85
4	936	1853	676	1546	0.55
5	69	215	42	103	0.59
6	101	175	73	206	0.84
7	193	1707	157	1749	0.65
8	11	61	6	65	0.85
9	42	1087	36	1113	0.49
10	2	28	2	86	0.68
11	84	1946	56	1995	0.69
12	5	89	1	93	1.35
13	121	4395	131	4410	0.70
14	65	357	52	372	0.87
15	52	142	45	154	0.95
16	81	148	61	168	1.13
17	24	213	37	184	0.31
18	11	41	8	20	0.61
19	50	84	47	83	0.57
20	125	292	82	339	1.43
21	20	1643	62	6520	1.08
22	0.5	52.5	2	92	1.48
23	0.5	29.5	1	22	0.56
24	5	25	3	57	1.06
25	144	871	132	886	0.26
26	24	293	35	276	0.76
27	4	74	3	76	0.54
28	19	60	7	69	0.68

Note. IHD events = fatal ischaemic heart disease and non-fatal myocardial infarction, n_0 = number of patients with an IHD event in the control group, N_0 = number of patients in the control group, n_1 = number of patients with IHD event in the treatment group, N_1 = number of patients in the treatment group.

Table 12.8 Nonparametric ML-Estimate of Treatment Effect Distribution Adjusted for Covariables and Classification of the Studies by Their Posteriori Component Membership Probability

Component (k)	$OR_k = \exp(b_k)$	CI	p_k	Allocated studies
1	1.13	(0.92, 1.37)	1.00	all studies

Note. OR_k = odds ratio as effect size estimate, CI = confidence interval, p_k = posteriori component membership probability.

12.5 CONCLUSION

Baseline and effect heterogeneity are almost always present in hospital and clinical epidemiological meta-analyses. This makes the evaluation of a treatment effect difficult and the use of some standard methods questionable. Attempts to restrict the meta-analysis on a homogenous selection of studies can never be completely successful. On the other hand, such attempts usually mean renouncing valuable information. It is now generally agreed that meta-analysis can and should go further than simply producing overall summaries of effects. In particular, understanding the possible causes of any heterogeneity can increase both the scientific value and clinical and epidemiological relevance of the results from a meta-analysis. In this chapter, an appropriate method for addressing this issue is presented. This method is based on finite mixed generalized linear models (FMGLMs), which have proven to be very flexible tools to estimate mean effect sizes and explain heterogeneity. Different effect measurements can be considered simply by changing the link function of the model. For example, using the identity link instead of the logit link leads to meta-analyses of risk differences instead of odds ratios.

The analysis of the example data shows that much more information can be gained by this approach than by the standard method. The second example clearly shows that the heterogeneity itself can be the most interesting subject of analysis.

There are of course some disadvantages and limitations to this approach. The first limitation is the fact that the count data of the studies must be available, which is not always the case. A second limitation is that the number of studies in the meta-analysis should be larger then 10. Otherwise, the nonparametric maximum likelihood estimation will be doubtful. A third limitation concerns the number of covariables. This number should not be too large in comparison to the number of studies. Finally, it should be noted that misinterpretation of the effect of second level variables is possible, especially if they are mean values of the study population.

REFERENCES

Aitkin, M. (1999a). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, *55*, 117–128.

Aitkin, M. (1999b). Meta-analysis by random effect modelling in generalized linear models. *Statistics in Medicine*, *18*, 2343–2351.

Böhning, D. (2000). *Computer-assisted analysis of mixtures and applications: Meta-analysis, disease mapping, and others.* Boca Raton, FL: Chapman & Hall/CRC.

DerSimonian, R., & Laird, N. M. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188.

Dietz, E., & Böhning, D. (1994). Analysis of longitudinal data using a finite mixture model. *Statistical Papers*, *35*, 203–210.

Dietz, E., & Böhning, D. (1995). Statistical inference based on an general model of unobserved heterogeneity. *Lecture Notes in Statistics*, *104*, 75–82.

Olkin, I. (1999). Diagnostic statistical procedures in medical meta-analysis. *Statistics in Medicine*, *18*, 2231–2341.

Thompson, S. G. (1993). Controversies in meta-analysis: The case of the trials of serum cholesterol reduction. *Statistical Methods in Medical Research*, *2*, 173–192.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, *18*, 2693–2708.

Veenstra, D. L., Saint, S., Saha, S., Lumley, T., & Sullivan, S. D. (1999). Efficacy of antiseptic-impregnated central venous catheters in preventing catheter-related bloodstream infection: A meta-analysis. *Journal of the American Mathematical Association*, *281*, 261–267.