Chapter 6

Discussion

6.1 Comparison with Classical Approaches

It is becoming more important in medical research to carry out analyses which explains reasons for heterogeneity of the treatment effects in multicenter studies.

In a conventional approach, the weighted regression analysis assumes a normal distribution for the logarithmic relative risk where for each center the variance is gained from a first-order Taylor-series approximation. For the analysis, this variance is treated as a known value. To explain the heterogeneity of the treatment effects by covariate information, the idea of generalized linear model was applied and the identity link function was used to link covariate information to relative risk. Fixed and Random effects of weighted regression analysis have been discussed in many literatures [4, 11, 32, 36, 38]. The problem related to these approaches, is firstly, the potentially insufficient approximation of the normal distribution. This may be inadequate when the sample size or the number of events for the individual studies are small. Secondly, correlation between estimates of the log-relative risk and their variance estimates may produce bias in the estimates of regression coefficients. The third problem is the inappropriate identity link which is used to link covariate information to relative risk. This link does not guarantee that the relative risk estimates are positive, which would be an essential requirement for a relative risk.

In the second conventional approach, logistic regression is applied to overcome these problems. The approach directly uses the binomial structure of the binary data. The conventional logistic model assumes a fixed parameter and fixed effects to intercept parameter and treatment effects, respectively. Multi-level logistic regression model is

one way to allow heterogeneity of treatment effects in the models. The treatment effects are assumed to be random effects and have an independent normal distribution in this model. An alternative multi-level model, the intercept parameter, is regarded as random rather than fixed. The multi-level approach has been discussed in various contributions [3, 36, 37]. One disadvantage of the logistic regression approach is the strong influence of the nuisance parameter, that is the intercept parameter, on estimating the treatment effects and potential lack of power in identifying the structure of relative risk.

The profile likelihood approach eliminates the nuisance parameter before dealing with the inference for the parameter of interest and becomes therefore attractive in this situation. In this study, a generalized linear model was applied to describe covariate information. The canonical link is used to link covariate information to relative risk which guarantees that the relative risk estimates are positive. In practice there may be many such covariates that should be considered for the sources of heterogeneity. Collinearity between pairs of covariates can occur and lead to difficulty in interpretation [5]. This can be a problem in modelling covariate information using profile likelihood.

As in many references [14, 29, 36, 38], heterogeneity of the treatment effects cannot be completely explained by the covariate information. Thus, it is important to take into account the possibility of unobserved heterogeneity. Ignoring the unobserved heterogeneity will underestimate the standard error of the parameters in the model, and thus overstate the importance of the covariate [36]. However, the modelling of covariate information using the profile likelihood considered in this study does not include the unobserved heterogeneity of the treatment effects into the model. In other words, the unobserved heterogeneity of the treatment effects need to be considered.

6.2 Extensions of the Profile Likelihood Model

In this section, the possible extensions of the modelling covariate information using profile likelihood are described.

We consider again in the likelihood function for estimating the relative risk for each center

$$\exp(-\theta_i p_i^C n_i^T)(\theta_i p_i^C n_i^T)^{x_i^T}/x_i^T! \times \exp(-n_i^C p_i^C)(n_i^C p_i^C)^{x_i^C}/x_i^C! \tag{6.1}$$

and then inserting the MLE of p_i^C (3.10) into (6.1) leading to

$$\exp\left(-\theta_i n_i^T(\frac{x_i^C + x_i^T}{n_i^C + n_i^T\theta_i})\right)\left(\theta_i n_i^T(\frac{x_i^C + x_i^T}{n_i^C + n_i^T\theta_i})\right)^{x_i^T}/x_i^T!$$

$$\times \exp\left(-n_i^C\left(\frac{x_i}{n_i^C+n_i^T\theta_i}\right)\right)\left(n_i^C\left(\frac{x_i^C+x_i^T}{n_i^C+n_i^T\theta_i}\right)\right)^{x_i^C}/x_i^C! \tag{6.2}$$

and, finally, the parameter dependent part of the profile likelihood for risk ratio becomes

$$f(x_i^T, x_i^C, \theta_i) = \frac{\theta_i^{x_i^T}}{(n_i^C+\theta_i n_i^T)^{x_i^T+x_i^C}}. \tag{6.3}$$

In situations where heterogeneity is present, the mixture profile likelihood approach has been applied. It is assumed that the population consists of m different subpopulations with different relative risks $\theta_1, \theta_2, \ldots, \theta_m$. Some centers belong to the same subpopulation or cluster. However, it is assumed that the subpopulation or cluster to which the center belongs is not known. This situation is called unobserved or latent heterogeneity [7].

Suppose the distribution of subpopulations in the population follows a discrete distribution

$$Q = \begin{pmatrix} \theta_1, & \ldots, & \theta_m \\ q_1, & \ldots, & q_m \end{pmatrix} \tag{6.4}$$

where $q_1, \ldots, q_m$ is the probability of belonging to the subpopulation j, $j = 1, \ldots, m$. Additionally, let a latent indicator z_{ij} describe which of the m subpopulations the center belongs:

$$z_{ij} = \begin{cases} 1 & \text{center } i \text{ belongs to subpopulation } j \\ 0 & \text{otherwise} \end{cases}$$

and these indicator variables z_{ij} have the constrains that

$$\sum_{j=1}^{m} z_{ij} = 1. \tag{6.5}$$

The conditional likelihood, conditional that observed x_i^T, x_i^C in center i comes from subpopulation j, is given as

$$f(x_i^T, x_i^C, \theta_j) = \frac{\theta_j^{x_i^T}}{(n_i^C+\theta_j n_i^T)^{x_i^T+x_i^C}}. \tag{6.6}$$

Note that contrary to (6.3), several centers might come from the same subpopulation j with parameter θ_j. With (6.6) and given data (x_i^T, x_i^C, z_{ij}), the associated likelihood can be written as

$$\prod_{j=1}^{m} \left\{ f(x_i^T, x_i^C, \theta_j) q_j \right\}^{z_{ij}}. \tag{6.7}$$

Therefore, the unconditional or marginal likelihood over the unobserved variable z_{ij} becomes

$$\sum_{(z_{i1},z_{i2},\ldots,z_{im})} \prod_{j=1}^{m} \left\{ f(x_i^T, x_i^C, \theta_j) q_j \right\}^{z_{ij}} = \sum_{j=1}^{m} f(x_i^T, x_i^C, \theta_j) q_j. \tag{6.8}$$

Finally, the mixture profile log-likelihood overall centers can be written as

$$L^*(Q) = \sum_{i=1}^{k} \log \left(\sum_{j=1}^{m} f(x_i^T, x_i^C, \theta_j) q_j \right). \tag{6.9}$$

In situations where observed heterogeneity in the form of covariates is considered, the profile likelihood (6.3) becomes

$$L^*(\beta) = \sum_{i=1}^{k} \log \frac{\exp(\beta_0 + \eta_i)^{x_i^T}}{(n_i^C + \exp(\beta_0 + \eta_i) n_i^T)^{x_i^T + x_i^C}} \tag{6.10}$$

with $\eta_i = \beta_1 z_{i1} + \beta_2 z_{i2} + \ldots + \beta_p z_{ip}$.

One method to incorporate unobserved heterogeneity into the likelihood corresponding to (6.10) is to enter linearly an unobserved heterogeneity into the model. Here, we suppose that the intercept parameter is mixing and follows a discrete distribution

$$Q = \begin{pmatrix} \beta_0^{(1)}, & \ldots, & \beta_0^{(m)} \\ q_1, & \ldots, & q_m \end{pmatrix} \tag{6.11}$$

and then, the conditional likelihood for this situation is given as

$$f(x_i^T, x_i^C, \theta_{ij}) = \frac{\exp(\beta_0^{(j)} + \eta_i)^{x_i^T}}{(n_i^C + \exp(\beta_0^{(j)} + \eta_i) n_i^T)^{x_i^T + x_i^C}}. \tag{6.12}$$

Therefore, the mixture profile log-likelihood for incorporating covariate information and unobserved heterogeneity over all centers can be written as

$$L^*(\beta, Q) = \sum_{i=1}^{k} \log \left(\sum_{j=1}^{m} f(x_i^T, x_i^C, \theta_{ij}) q_j \right). \tag{6.13}$$

There are numerous ways how an unobserved heterogeneity can enter the model, however equation (6.13) illuminates only one of them. The intercept parameter might be fixed ($\beta_0^{(1)} = \beta_0^{(2)} = \ldots = \beta_0^{(m)} = \beta_0$) and mixing might occur in one, several or all covariate parameters, as it occurs in the intercept parameter similarly to our case. In other words, a new range of models need to be considered.

In this study, we consider only covariates on the study level, for example, average cholesterol reduction, percentage of males that could explain the differences between

the studies. We did not go into covariates on the individual level. The use of individual data would of course become necessary in the developing area where covariates on the individual level rather than on the study level are investigated [31], the profile likelihood approach naturally is applied to this case.