Efficient Bayesian Sample Size Calculation for Designing a Clinical Trial with Multi-Cluster Outcome Data

KELLY H. ZOU*1,2, FREDERIC S. RESNIC3, ADHEET S. GOGATE1, SILVIA ONDATEGUI-PARRA1,4 and LUCILA OHNO-MACHADO1

1 Department of Radiology, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA 02115, USA
2 Department of Health Care Policy, Harvard Medical School, Boston, MA 02115, USA
3 Division of Cardiovascular Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA 02115, USA
4 Harvard Center for Risk Analysis, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA 02115, USA

Abstract

Health care utilization and outcome studies call for hierarchical approaches. The objectives were to predict major complications following percutaneous coronary interventions by health providers, and to compare Bayesian and non-Bayesian sample size calculation methods. The hierarchical data structure consisted of: (1) Strata: PGY4, PGY7, and physician assistant as providers with varied experiences; (2) Clusters: \(k_s\) providers per stratum; (3) Individuals: \(n_s\) patients reviewed by each provider. The main outcome event illustrated was mortality modeled by a Bayesian beta-binomial model. Pilot information and assumptions were utilized to elicit beta prior distributions. Sample size calculations were based on the approximated average length, fixed at 1%, of 95% posterior intervals of the mean event rate parameter. Necessary sample sizes by both non-Bayesian and Bayesian methods were compared. We demonstrated that the developed Bayesian methods can be efficient and may require fewer subjects to satisfy the same length criterion.

Key words: Bayesian statistics; Sample size calculation; Cluster analysis; Cardiovascular study; Risk assessment.

1. Introduction

A randomized controlled trial (RCT) is one of the simplest, most powerful and revolutionary tools of research. In light of the call for evidence-based medicine, important, albeit isolated, research efforts have used RCTs as the subject rather than the tool of research. These studies are usually designed to generate empirical evidence to improve the design, reporting, dissemination, and use of RCTs in health care (see SILVERMAN and CHALMERS, 1992; JADAD and RENNIE, 1998).

* Corresponding author: zou@bwh.harvard.edu

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 0323-3847/03/0710-0826
Over the past decade, health organizations in many countries have sponsored an increasing number of multi-center clinical trials and health observational studies in medical research. For example, in the past decade, the Radiological Diagnostic Oncology Group (RDOG; see http://www.acr.org/links/rdog_intro.html) studies, supported by the National Institutes of Health and the National Cancer Society of the US, evaluated emerging and practical imaging modalities in the management of patients with cancer. The format chosen for these studies was multi-institutional clinical trials. The RDOG group has evaluated separately and prospectively over ten years the relative accuracy of CT, MRI, and less frequently ultrasound in staging cancers in the following 9 anatomical sites: prostate, lung, colorectal, pancreas, musculoskeletal, and head & neck, ovarian, pediatric, and breast. Currently, the American College of Radiology Imaging Network (ACRIN; see http://www.acrin.org), a National Cancer Institute-funded cooperative group, has sponsored clinical trials of diagnostic imaging and image-guided therapeutic technologies. The ACRIN studies have generated information on cervical, brain, prostate, breast, bone, lung cancers, with the goal of lengthening and improving the quality of the lives of cancer patients.

Multi-center clinical studies or quality of care studies call for complex and efficient study designs and analyses. Health care utilization and outcome data often have a multi-level structure with individual patients at the first level, and then with subsequent physicians, hospitals, and geographic regions forming several higher level clusters. Donner, Birkett and Buck (1981) and Donner (1985) showed that cluster sample size may vary substantially at each level of the hierarchy. In addition, both individual- and cluster-level covariates are often available for measuring, for example, disease severity and co-morbidity for individual patients, as well as location, size, and organizational characteristics of a particular hospital.

The Food and Drug Administration of the US (Campbell, 2000) has recently been promoting Bayesian methods in health economics and outcomes analysis. The objective of this study is to present statistical Bayesian methods for efficient sample size calculation based on count and proportion data. Following Parmigiani and Berry (1994) the method has an additional advantage in the multiple cluster setting because it lends better “borrowing strength” to the estimation of the parameters or hyper parameters among clusters.

We design and illustrate our methods on a hypothetical multi-cluster study proposed by Resnic et al. (2001) of a predictive model following cardiac interventions. The pilot study data used to illustrate our study design methods are derived from our published investigation of the interventional cardiology service at a major affiliated teaching hospital. The proposed study attempts to assess the risk of major in-hospital complications following percutaneous coronary interventions under both subjective and objective risk score models. In the proposed study, the authors plan to integrate the predictive models into the workflow of an interventional cardiology clinic in order to study their advantages over subjective physician estimation of risk for in-hospital complications after angioplasty. We refer to subjective assessments made by physicians or physician assistants (providers), as
opposed to objective assessments calculated under computerized predictive models, which the authors have developed.

Our main design goals are: (1) To minimize sample size enabling more efficient study designs, compared with a frequentist non-Bayesian design, (2) to better utilize prior information, and (3) to incorporate complex structures (e.g., individuals, clusters, and strata) and variability in the data analysis.

2. A Motivating Clinical Example

2.1 Subjective assessment of mortality rate

This research is intended to design a multi-cluster study in order to evaluate the performance of subjective binary assessments by health providers of major complications following percutaneous coronary interventions. The setting of this study is a tertiary care academic teaching hospital affiliated with a major medical school. Its Cardiac Catheterization Service performs more than 2,000 interventional procedures each year.

The sub-domain of angioplasty and the study of in-hospital complications is ideal for our illustrated study design because: (1) the domain is very focused, with a complete yet limited collection of items that are currently captured in structured format, and (2) the follow-up data are complete with essentially no censored items.

We have developed predictive models for major complications following angioplastic procedures based on a small number of critical pre-procedural variables. We wish to design a complex study to systematically evaluate the predictive ability of major in-hospital complications.

Outcomes modeled in this study included in-hospital complications such as death, as well as the combined outcome of death, need for coronary bypass surgery and myocardial infarction. Each patient is assessed by one of the following participating providers during a 24-month period: postgraduate fellows and physician assistants. Each individual provider constitutes a cluster and has his or her subjective binary assessments of event, such as death, for each patient.

2.2 Motivation for the study design

We focus exclusively on the statistical methodological aspects of designing the study to evaluate and compare the predictive abilities of the objective and subjective models described above. The incentive of adopting the proposed method to efficiently design such a study relies on available pilot information. The main purpose of conducting such a study led by a group of cardiology investigators is to evaluate the accuracy of providers’ subjective binary prediction using an event rate.

We adopt the “average length criterion” of JOESPH, DU BERGER, BELISLE (1997) because of the complex hierarchical structure in this study. In contrast with the
sample size calculation based on statistical hypothesis and power analysis, the length criteria are closely associated with the concept of confidence interval, and are often employed in practice. It is the pre-specified length of 95% central posterior or highest probability density interval. See the articles by Joseph, Wolfson and du Berger (1995), Joseph et al. (1997), Zou and Normand (2001), and Normand and Zou (2002) on the argument for using such a criterion in Bayesian inference and study designs. Alternative criteria such as coverage have also been considered as presented in Joseph et al. (1997).

3. Methods

3.1 Hierarchical nature of the design

The observed event data (the indicators for the events) for the individual is treated as the outcome variable. Denote each patient as an individual, each provider as a cluster, and a category of providers as a stratum. The hierarchical nature of this study is quite explicit, with higher to lower levels in the hierarchy of the following data structure:

(1) Strata: There are 3 strata (post-graduate year PGY4, PGY7, and physician assistant PA), indicating the category of provider experience.

(2) Clusters: For the $s$th strata ($s = 1, 2, 3$), there are a total of $k_s$ clusters (providers). Specifically, $k_1 = 28$ PGY4s, $k_2 = 10$ PGY7s, and $k_3 = 12$ PAs in a two-year study, which are pre-specified due to the workflow.

(3) Individuals: The number of cluster-specific individuals, $n_s$, is assumed to be balanced; i.e., $n_1$ individuals/PGY4, $n_2$ individuals/PGY7, and $n_3$ individuals/PA, with a total $N = \sum_{s=1}^{3} n_s k_s$ per length of study (two years). The numbers of patients (e.g., all of the $n$’s and the total sample size, $N$) were computed and compared using both non-Bayesian and proposed Bayesian methods.

3.2 Non-bayesian sample size calculation method

The sample size $N$ concerning the mean risk of death was based on pilot data provided by Resnic et al. (2001). By pre-specifying the target length, $\Delta$, of a $(1 - \alpha)$% confidence interval, the sample size was computed below. Conventionally, the significance level is fixed at $\alpha = 5\%$. The advantages of interval-criterion based approaches to sample size calculations over the statistical hypothesis and statistical power analysis based approaches are discussed in Joseph et al. (1995; 1997). It is assumed based on data from another medical domain that the accuracy of risk estimates can be different based on the category of medical training (stratum). The sample size calculations assume that the risk assessment results are similar among different providers within each stratum.

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
We assume a binomial model for this design problem. For each stratum, \( s \) (\( s = 1, 2, 3 \)), let the estimated number of event \( Y_{sj} \) for provider \( j \) have and independent binomial distribution with the corresponding event rate parameter of interest, i.e.,

\[
(Y_{sj} | n_s, \theta_{sj}) \sim \text{binomial} (n_s, \theta_{sj}).
\]

Furthermore, we assume a common underlying \( \theta \) for all of the above \( \theta_s \). This assumption is made for our illustrated application because typically the event rate, such as mortality, could be quite low.

The lower- and upper-bounds of a \( 100(1 - \alpha)\% \) confidence interval is constructed for the mean event rate \( \frac{1}{k_s} \sum_{j=1}^{k_s} \theta_{sj} \) is defined by

\[
\pm \Phi^{-1}(1 - \alpha/2) \left\{ \frac{\theta(1 - \theta)}{k_sn_s} \right\}^{1/2},
\]

where the first term \( \Phi^{-1}(1 - \alpha/2) \) is inverse standard normal quantile with cumulative probability of \( 1 - \alpha/2 \). Typically \( \alpha \) is fixed at 0.05, with \( \Phi^{-1}(0.975) = 1.96 \). Thus, if we pre-specify a target length, \( \Delta \), of the 95% confidence interval, then the necessary sample size for each stratum required to satisfy this length criterion can be computed from Equation (1):

\[
n_s = \frac{2 \{ \Phi^{-1}(0.975) \}^2 \theta(1 - \theta)}{k_s \Delta^2}.
\]

We further specify that \( \Delta = 1\% \) because the assumed underlying event rate is typically low (\( \theta = 2\% \)) based on our pilot result as shown in Resnic et al. (2001).

3.3 Bayesian hierarchical sample size calculation method

In addition to the binomial model on the individual patient level given in Equation (1), which is assumed for the event rates by different providers (cluster) per stratum, we assume that there exist similarities between clusters which one could borrow-strength across different providers. Thus, for each stratum, \( s \) (\( s = 1, 2, 3 \)), we assume that the event rate \( \theta_{sj} \) has a common beta distribution with hyper parameters \( (a_s, b_s) \).

\[
(\theta_{sj} | a_s, b_s) \sim \text{beta} (a_s, b_s).
\]

There are several methods for eliciting the hyper-parameters. The beta distribution yields a flexible shape of its probability density function, taking values between 0 and 1, and thus convenient to model a proportion or rate parameter. In addition, such prior distribution may be constructed both by informative and non-informative method. Jeffreys’ prior is \( (a, b) = (0.5, 0.5) \). A uniform flat prior gives \( (1, 1) \). When \( a \) and \( b \) are equal, the density function is symmetric at 0.5. However, when \( a > b \), it’s right skewed with the tail pointing towards 1. Conversely, when \( a < b \), it’s left skewed with tails towards 0.

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
In the cardiovascular application, our pilot study showed that the mortality rate is 2% (95% confidence interval between 1.3% to 3.4% with a length of approximately 2%) in the moderate risk group. In order to constructing reasonable priors for the three strata of different types of providers, we further assume that all are unbiased with mean of the beta distribution as 2%. The variances of the priors decrease from PGY4, PGY7, and to PA, which reflects lower variability for providers with greater clinical experiences. To specify realistic values for the hyperparameters \((a_s, b_s)\), we use the following method-of-moment. Note that the mean and variance of each rate parameter, \(\theta_{sj}\), can be expressed in terms of \((a_s, b_s)\) based on the beta prior distributions with the mean and variance, respectively, as:

\[
\mu_{sj} = E(\theta_{sj} | a_s, b_s) = \frac{a_s}{a_s + b_s} \quad \text{and} \quad \Var(\theta_{sj} | a_s, b_s) = \frac{1}{(a_s + b_s + 1)} \mu_{sj}(1 - \mu_{sj}),
\]

where \(\mu_{sj}\) is the expected rate parameter and \(\Var(\theta_{sj} | a_s, b_s)\) is the variance. In our application, we know that the expected rate parameter is around 2%, similar to the non-Bayesian approach given in Equation (3). Thus, from Equation (5), the variance decreases in \((a_s + b_s)\), which is often known as the precision of the rate parameter. Because the mean rate is rather low as 2%, we realistically consider these values for the hyperparameters at \((a_s, b_s) = (1.0, 49.0)\) for PGY4s with a prior standard deviation (SD, i.e., square root of the variance) of 1.96%, \((1.1, 55.0)\) for PGY7s with a SD of 1.83%, and \((1.2, 59.0)\) for PAs with a SD of 1.79%, suggesting less variability among the PAs, as in this setting PAs tend to have greater clinical experience in subjective risk assessment.

It can be shown that the posterior distribution of the rate parameter is:

\[
(\theta_{sj} | y_{sj}, a_s, b_s) \sim \text{beta}(c_{sj}, d_{sj}),
\]

where \(c_{sj} = y_{sj} + a_s\) and \(d_{sj} = n_s - y_{sj} + b_s\).

Since the mean rate, \(\frac{1}{k_s} \sum_{j=1}^{k_s} \theta_{sj}\), over all providers within each stratum is of interest, the posterior mean and variance of such mean rate becomes based on the posterior beta distributions are, respectively:

\[
\mu_s = E\left(\frac{1}{k_s} \sum_{j=1}^{k_s} \theta_{sj} | y_s, a_s, b_s\right) = \frac{1}{k_s} \sum_{j=1}^{k_s} \frac{c_{sj}}{c_{sj} + d_{sj}}
\]

\[
\sigma_s^2 = \Var\left(\frac{1}{k_s} \sum_{j=1}^{k_s} \theta_{sj} | y_s, a_s, b_s\right)
= \frac{1}{k_s^2} \sum_{j=1}^{k_s} \frac{c_{sj}d_{sj}}{(c_{sj} + d_{sj})^2 (c_{sj} + d_{sj} + 1)},
\]

with \(c_{sj}\) and \(d_{sj}\) as given below Equation (6).
The lower- and upper-bounds of a $100(1 - \alpha)\%$ central posterior interval for estimating $\frac{1}{k_s} \sum_{j=1}^{k_s} \theta_{sj}$ are constructed by approximating the posterior beta distribution by a normal distribution:

$$\pm \Phi^{-1}(1 - \alpha/2) \{\sigma_s^2\}^{1/2},$$

with $\sigma_s$ given in Equation (7). Again $\alpha$ is typically 0.05 and $\Phi^{-1}(0.975) = 1.96$. However, based on Equations (6) and (7), the variance term $\sigma_s^2$ also depends on $y_{sj}$ as in the expressions of $c_{sj}$ and $d_{sj}$ given in Equation (6). Thus, in order to compute the necessary sample size for each stratum, we need to simulate the $y_{sj}$ data from the specified beta-binomial model given in Equations (1) and (4) using the Monte-Carlo method. A previous investigation by Zou and Normand (2001) found that for the mean function, the normal approximation of the posterior distribution performs well, analogous to the Central Limit Theorem in the non-Bayesian setting. Alternatively, the posterior beta distribution can be used although it is less computationally convenient to implement the methods on the basis of the beta distribution.

The statistical computation is conducted using S-Plus software program (http://www.insightful.com), and the code for implementation will be distributed in the future. During each of the $l$th ($l = 1, \ldots, M$) Monte-Carlo iteration with a guess of the target sample size $n_s$, the length of a 95% central posterior interval is computed by $\hat{\Delta}_{sl} = 2\Phi^{-1}(0.975)\sigma_{sl}$. Over $M$ iterations, the corresponded average estimated length is then $\hat{\Delta}_s = \frac{1}{M} \sum_{l=1}^{M} \hat{\Delta}_{sl}$. Finally, the target optimal sample size $n_s$ is the smallest integer number such that the average central posterior interval stays within the specified target length, i.e., $\Delta_s \leq \Delta$. In the cardiovascular application, the target interval length $\Delta$ is fixed at 1% for all strata. A detailed Monte-Carlo sampling scheme can be found in the related work conducted by Zou and Normand (2001). We also discovered previously that a reasonable initial guess of the sample size $n_s$ is to subtract the precision parameter from the non-Bayesian solution presented in Equation (2), i.e., $n_s[Initial] = n_s[Non-Bayesian] - (a_s + b_s)$. In addition, instead of arbitrarily specifying a large enough number of Monte-Carlo iterations, $M$, we monitor the convergence of the algorithm using the CODA software program (http://www.mrc-bsu.cam.ac.uk/bugs/classic/coda04/readme.shtml).

4. Results

The sample sizes required in a two-year 50-cluster cardiovascular clinical example designed by both the non-Bayesian and Bayesian methods are presented in Table 1. The significance level is $\alpha = 5\%$, the target average length is $\Delta = 1\%$, and the underlying predicted mortality rate is $\theta = 2\%$. Without using prior knowledge, the
non-Bayesian solution requires to observe a total of \( \sum N = 9,056 \) patients by a total of 50 providers (108 patients by each of the 28 PGY4’s, 302 patients by each of the 10 PGY7’s, and 251 patients by each of the 12 PA’s). Such sample size is very large and may be impractical.

In comparison, the Bayesian sample size calculations resulted in a total of \( \sum N = 5,458 \) patients by these providers (51 patients by each of the PGY4’s, 199 by each of the PGY7’s, and 170 by each of the PA’s). The difference of the two approaches shows a total of saving of 3,598 patients to be evaluated in a two-year study period.

The bias of the mean rate parameter in each stratum using the Bayesian Monte-Carlo approximation of the posterior was small (all within 0.1%). In addition, the convergence criterion in the Monte-Carlo iterations was satisfied.

5. A Simulation Study

We conducted additional simulations to investigate the effect of the choice of beta hyper-parameters on the resulting average length of a rate parameter in a clustered design but without having different strata. Thus the index for stratum is dropped in the notation for simplicity.

In this simulated study, the event rate may be more general than the low mortality rate of major complications following percutaneous coronary interventions predicted by health providers as in our clinical example. It may be success or failure rate of a certain event of interest. In other words, the expected rate, \((0, | a, b)\) include \{50%, 75%, 90%\} with hyper-parameters \((a, b) = \{(1, 1), (3, 1), (9, 1)\}\). A lower rate than 50% can be perceived by just studying the complement...
of such a rate. For example, a 75% success (survival) rate can be viewed as a 25% failure (mortality) rate. Note, when \((a, b) = (1, 1)\), essentially, the prior distribution is the uniform between 0 and 1. The cluster sample sizes are specified at \(n = \{5, 20, 50, 100, 500\}\) per cluster. The numbers of clusters are \(k = \{10, 20, 50, 100\}\). The clusters in a broader sense represent institution or provider.

Using the Bayesian sample size method described above, we conduct \(M = 1000\) Monte-Carlo iterations to compute the average length of the 95% central posterior confidence intervals for \(\hat{\theta}\) based on each of the combination of the cluster sample size \(n\), the number of clusters \(k\), and the hyper-parameters \((a, b)\). The convergence is monitored.

To compare the Bayesian result against the non-Bayesian result, we analytically compute the approximate average length of the 95% confidence intervals of the mean rate \(\frac{1}{k} \sum_{j=1}^{k} \theta_j\) using the lower- and upper bounds calculated in Equation (2).

In Table 2, we report both non-Bayesian and Bayesian estimate of the average 95% confidence intervals or 95% central posterior intervals for the mean rate with a specified cluster sample size \(n\) and the number of clusters \(k\), at the underlying rates of 50%, 75% and 90%. Directly comparing the two methods for estimating

<table>
<thead>
<tr>
<th>(\theta)</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n)</td>
<td>(k = 10)</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>0.277</td>
<td>0.196</td>
<td>0.124</td>
</tr>
<tr>
<td>20</td>
<td>0.139</td>
<td>0.098</td>
<td>0.062</td>
</tr>
<tr>
<td>50</td>
<td>0.088</td>
<td>0.062</td>
<td>0.039</td>
</tr>
<tr>
<td>100</td>
<td>0.062</td>
<td>0.044</td>
<td>0.028</td>
</tr>
<tr>
<td>500</td>
<td>0.028</td>
<td>0.020</td>
<td>0.012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>((a, b))</th>
<th>(1, 1)</th>
<th>(1, 3)</th>
<th>(1, 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n)</td>
<td>(k = 10)</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>0.189</td>
<td>0.136</td>
<td>0.086</td>
</tr>
<tr>
<td>20</td>
<td>0.106</td>
<td>0.077</td>
<td>0.048</td>
</tr>
<tr>
<td>50</td>
<td>0.069</td>
<td>0.050</td>
<td>0.032</td>
</tr>
<tr>
<td>100</td>
<td>0.049</td>
<td>0.036</td>
<td>0.023</td>
</tr>
<tr>
<td>500</td>
<td>0.022</td>
<td>0.015</td>
<td>0.010</td>
</tr>
</tbody>
</table>
the average interval length, the Bayesian solution yields much narrower length, especially for small \( n \) or small \( k \), suggesting that the method is more efficient under the same sample size consideration.

However, the average length also depends on the prior distribution specified. With the same combination of \( n \) and \( k \), however, the specified parameters yield quite different length. Thus, the elicitation of prior distribution is an important issue, similar to the design assumptions prior to conduct sample size calculations for a new clinical study.

6. Discussion

Health care utilization and outcome studies call for hierarchical approaches. Due to the high cost of clinical studies, non-Bayesian methods may be too conservative in terms of study designs. By employing Bayesian design methodology, one may better utilize prior information from a previous pilot to design a more compact study. For small sample sizes or small number of clusters (provider or institutions), we have demonstrated both in our cardiovascular application based on an artificial design example that the Bayesian methods are efficient with tremendous reductions in sample size. Our statistical Bayesian methodology may be adapted to study designs in a variety of other clinical domains. Estimation methods may also be applicable to data analysis.

Several authors have also examined functions other than the mean event rate over different clusters. For example, the difference between two event rates has been studied by Joseph et al. (1997). Functions such as range of all even rates have also been studied previously under either fully exchangeable or partially exchangeable assumptions made by Zou and Normand (2001) and Normand and Zou (2002), respectively.

The elicitation of the prior distribution is an important issue. Although anti-conservative priors may lead to a much reduced sample size, the assumptions should be investigated using careful pilot data. Formal rule-based prior elicitation is recommended. In addition, one can hope to further borrow strength over all clusters by including a higher hierarchical stage. However, by adding another stage the posterior distribution no longer has simple analytical expression, and the sample size estimation procedure becomes more computational burdensome and time consuming. Therefore, we do not consider such a hierarchical three-stage model.

In summary, we have presented a simple Bayesian hierarchical method to calculate sample sizes for cluster outcome data. Compared to the non-Bayesian methods, the described method can reduce sample sizes as illustrated in our hypothetical cardiovascular risk-assessment study and on the statistical simulation results.
Acknowledgement

This work was in part supported by the Agency for Healthcare Research and Quality (AHRQ) of the National Institutes of Health HS13234-01, USA.

References


Kelly H. Zou, Phd
Department of Heath Care Policy
Harvard Medical School
180 Longwood Avenue
Boston, MA 02115
USA
Email: zou@bwh.harvard.edu
Phone: 617-732-6536; Fax: 617-264-6887

Received January 2003
Accepted March 2003

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim