

# **Mathematisches Forschungsinstitut Oberwolfach**

## Tagungsbericht 8 / 1995

# Medical Statistics: Multivariate Models for Longitudinal Data

19.2. - 25.2.1995

The conference was organized by Martin Schumacher (Freiburg; Germany) and Niels Keiding (København; Denmark). There were 37 participants from Great Britain, Germany, The Netherlands, Sweden, Switzerland, Denmark, Norway, France, USA, Taiwan and Australia. Longitudinal data arise in many different situations whenever the dynamic of a process is of interest or time afflicts certain associations. Repeated assessment over time of one or multiple outcome variables and/or covariates calls for models which take into account the various dependencies among the obtained measurements and with time. The 29 presentations of this conference were mainly concerned with two different kinds of longitudinal data: In the first part of the conference the majority of presentations dealt with models for multivariate outcome variables that are measured at several occasions: marginal models, GEE, graphical models for conditional associations and independence, random effects models. In the second part situations were considered where the outcome variables are times to certain events. Special emphasis is put on modelling of multivariate failure times, time-dependent covariates and confounders, frailty models for correlated or timedependent frailties and recent developments in causal inference in structural nested failure time models.

Application of the statistical methods was illustrated using various examples from clinical and epidemiological research.

The fruitful and encouraging discussions at this conference have certainly inspired future international collaboration and new research activities.



#### **Abstracts**

#### Garrett Fitzmaurice (Oxford; Great Britain)

<u>A Likelihood-Based Approach for Analysing Longitudinal</u>
<u>Binary Responses</u>

We discuss a likelihood-based method for analysing longitudinal binary responses based on a log-linear representation of the joint probabilities of the binary responses. We focus on "marginal models" in which the marginal expectation of the response variable is related to a set of covariates. The association between binary responses is considered to be a nuisance characteristic of the data and is modelled in terms of conditional log odds-ratios. One important advantage of this "mixed parameterization" is that the maximum likelihood estimates of the marginal mean parameters are robust to misspecification of the time dependence. We describe an iterative two-stage procedure for obtaining the maximum likelihood estimates when there are incomplete or missing responses. Two examples are presented to illustrate this methodology.

## Edgar Brunner (Göttingen, Germany)

# Nonparametric Methods for Longitudinal Data

Parametric models for longitudinal data have been studied extensively in literature. With the exception of rank methods that have been applied to some special semiparametric models, a real nonparametric approach to longitudinal data does not exist. Based on some recent theoretical results on the asymptotic distribution of linear rank statistics for independent vectors (Brunner & Denker; 1994), some ideas to develop nonparametric procedures for such data shall be introduced in this paper. The hypotheses are formulated in terms of the distribution functions (Akritas & Arnold; 1994) and pure rank tests are derived to test these hypotheses where the general asymptotic results (Akritas & Brunner; 1995) are applied to special models for longitudinal data. Since no continuity of the distribution functions is assumed, the proposed procedures can also be applied to pure ordinal data such as quality scales, e.g. The procedures are also designed for an unequal number of replications at each time point per subject.





Approximations for small samples are shortly discussed and the results are supported by a simulation study. The application of the proposed procedures is demonstrated by some worked out examples and the statements for the computation of the statistics by the SAS-procedure PROC MIXED are given.

#### Rolf Meinert (Mainz; Germany)

# Time-Varying Covariates in Cross-Sectional Studies

Besides information on current exposure- and disease status, sometimes additional data on prior exposure status is also gathered in cross-sectional studies. There are several ways how to adjust for exposure status at different points in time in a regression model. However, careful interpretation of the resulting estimates taking into account a detailed consideration of the possible exposure patterns is necessary. In addition, if changes in exposure are caused by the occurrence of the disease, then adjusting for multiple measurements of exposure can give misleading results. A regression analysis on dummy variables describing possible patterns of changes in exposure is proposed as an alternative approach. This approach is equivalent to the incorporation of interaction terms between different exposure measurements but the resulting estimates have a simpler interpretation. The above considerations are illustrated with data from a study on the relationship between bronchial hyperreactivity in children and maternal smoking.

## Hans van Houwelingen (Leiden; The Netherlands)

# <u>Fitting Normal-Mixture-Random-Effects Models to Repeated</u> <u>Non-Normal Measures</u>

The EM-algorithm for fitting normal models to non-normal (longitudinal) data involves integration over high-dimensional spaces. In the talk approximative methods will be discussed based on normal approximations for the posterior densities that play a role in the EM-algorithm.

An example will be discussed of dichotomous follow-up data from a randomized clinical trial. It will be shown that the approximations yield reasonable results that enable the comparison of different models and give some insight in the hidden covariance structure.



 $\odot$ 

#### Werner Vach (Freiburg; Germany)

# A Remark on the Preconditions for Using the GEE Approach in Analysing Longitudinal Studies

Let  $(Y_{it}, X_{it})$   $(t=1,..., T_i, i=1,..., n)$  be pairs of measurements of a covariate and an outcome variable for n individuals at  $T_i$  time points. Often the interest is in the parameters of the marginal regression models

$$E(Y_{it}|X_{it} = x_{it}) = \mu(\beta_{0t} + \beta_t x_{it}).$$

A first approach is to estimate the parameters by a cross-sectional analysis. The GEE approach promises a more efficient estimation by incorporating assumptions on the dependencies among the outcome variables. However, this approach requires that

$$\mathbf{E}\left(Y_{it}\middle|X_{i1} = x_{i1},...,X_{iT_{i}} = x_{iT_{i}}\right) = \mu\left(\beta_{0t} + \beta_{t}x_{it}\right).$$

If the true law generating the data is a Markov process, one can show that this condition is not satisfied in the case of time dependent covariates. We present some results on the asymptotic bias implied by this type of violation (joint work with U Grömping and K Schulz).

## Juni Palmgren (Stockholm; Sweden)

## Regression Models for Bivariate Binary Outcomes

BACKGROUND: A discrete time survival model is used for analysing data from the Helsinki Invalid Foundation Hospital on time from hip operation to possible failure of hip prosthesis. Hips were replaced both unilaterally and bilaterally. The analysis is driven by prior indication that different failure mechanisms may operate for the left and the right hip and that unilateral and bilateral arthroplastics may differ. Both person-specific and hip-specific covariates are involved.

THE MODEL: For the bilaterally treated patients a bivariate binary response is specified for hips at risk at the start of a time interval. Effects of covariates are addressed through logistic linear models for the marginal failure probabilities and a log linear model for the odds ratio. For the unilaterally operated patients one response is treated as missing.





LIKELIHOOD INFERENCE: The parameterisation constitutes the mixed parameterisation for a distribution belonging to the exponential family. Likelihood equations are set up using the connection between the multinomial and Poisson distributions.

#### Douglas Altman (London; Great Britain)

# Regression Modelling of Curved Relationships Using Fractional Polynomials

In regression models continuous variables are usually entered untransformed and represented as having a straight-line relationship with the outcome variable. The most common approach to examine possible curvature is polynomial regression, but this method has well-recognised difficulties, notably inflexibility, a propensity to produce artefacts, and the inability to model relations with asymptotes.

These problems can be largely overcome by the use of an extended family of polynomials known as fractional polynomials (Royston & Altman; 1994). Fractional polynomial models are similar to conventional polynomials in that the model comprises powers of X, but noninteger and negative powers (and log X) are allowed. In general a fixed set of candidate powers is considered, and the method involves evaluating the deviances of all models with a single power term, then with two terms, and so on.

These models can be used in any regression context. Various issues related to model choice are discussed, and several examples are presented. Special consideration is given to the case of multiple covariates

Fractional polynomials usually give better fit than conventional polynomials, often with fewer terms in the model. They offer a flexible system of simple parametric models of wide applicability.



#### Patrick Royston (London; Great Britain)

# Parametric Methods for Calculating Age-Specific Reference Centiles from Longitudinal Data

The problem of calculating two types of reference centiles from a few serial measurements on many individuals is considered. The context is the screening of fetuses for poor growth according to ultrasound measurements of fetal size. A simple linearizing approach is proposed whereby a power transform Z of the response variable (for example, estimated fetal weight) is linear in a monotonic function X of age. A suitable time transformation on X is a fractional polynomial. The distribution of Z is estimated by using a random-coefficients regression model with predictor X. This model determines both the unconditional centiles of fetal size at a given gestational age, and, of particular clinical reference, the distribution of present size conditional on size at earlier ages. The latter allows the clinician to assess whether the growth of the fetus is within normal limits for the population. The method is applied to a study of 113 pregnancies in which each woman was scanned up to 5 times by ultrasonography.

#### Theo Gasser (Zürich; Switzerland)

# Application of Structural Analysis of Curves to Human Growth

Most regression methodology deals with single curves and not with samples of curves (or regression problems), as is frequent in biomedical applications. To avoid artificial parametric models, non-parametric approaches would be desirable. Before doing any further analysis of the curves, they are aligned to a common dynamic (in growth, common maturational pace). This is done by estimating nonlinear monotone shift functions, determined by structural features common to all curves, such as extrema and inflection points. Afterwards curves can be averaged or treated in any other feasible way. Applications to longitudinal growth data show the potential of this method, in particular for bringing out fine structure in the data.



#### Ludwig Fahrmeir (München; Germany)

# Varying Coefficient Models, Roughness penalties and Gibbs Sampling for Event History Analysis

Varying coefficient - or dynamic - models for duration data with multiple states or risks allow simultaneous modelling and estimation of baseline hazards and time-varying effects in a flexible way. The focus is on discrete duration times, as they appear for example for unemployment duration in the German Socio-Economic Panel. For estimation two approaches are used: a) Penalized likelihood estimation, placing a roughness penalty on the sequence of time-varying parameters, thereby enforcing smooth variation. Various algorithmic solutions are available, e.g., Kalman-type smoothing. b) Bayesian posterior analysis, where a smoothner prior is imposed. Inference is possible via Markov chain Monte Carlo methods, e.g., Gibbs sampling. The methods are applied to analyse data on duration of unemployment, revealing significant time-varying effects of covariates such as sex, nationality and age.

## Uwe Feldmann (Homburg / Saar; Germany)

#### Linear Structural Regression (LISREG)

A multivariate regression model is suggested satisfying the properties: (i) all observable variables are subject to random errors, (ii) each observable variable can be predicted by means of the others, (iii) residual analysis decomposes the sample space into location and residual points, and, (iv) maximum likelihood as well as parameter estimation through moment and order statistics is feasible.

The concept of linear structural regression (LISREG) aims to combine the advantages of the multivariate regression model, for instance the predictability between observed variables, the identifiably and interpretability of all model parameters, with the advantages of the linear structural relationship (LISREL, Jöreskog, 1978) model, for instance the handling of straight linear relationships and the randomness of all observed variables. Linear structural regression avoids the concept of conditional expectations.

An one factorial error-in-the-variables regression model is established where all observed variables are assumed to be random. Each observable variable is predictable by means of the others simultaneously. Colinearities and straight linear relationships of the data are allowed in this approach. Structural variables are introduced which are the



analog to latent variables in the structural relationship model offering unique algebraic equations for the determination of the structural regression parameters and for model identification.

Maximum likelihood parameter estimation as well as parameter estimation via the moment method and order statistics are introduced. The latter leads to a median estimate of the model parameters which differs substantially from the least median of squares approach, proposed by Rousseeuw & Leroy (Psychometrika, 1987). It allows for a real robust multivariate data analysis.

The statistical properties of bivariate structural regression will be briefly discussed. A surprising result may be, that the structural regression parameter estimate is just the ratio of the standard deviations of the observable variables. This simple and useful estimate should be rediscovered for its presentation in statistical textbooks, see Freedman et al. (1978; p 122).

#### Norman Breslow (Seattle; USA)

# <u>Correcting the Bias in Penalized Quasi-Likelihood and Laplace</u> <u>Approximation Estimates in Generalized Linear Mixed Models</u>

Penalized Quasi Likelihood (PQL) is an approximate method for inference in generalized linear mixed models that may be motivated by first making a Laplace approximation to the likelihood that arises from integrating out the (normally distributed) random effects. Its principal advantage is that it may be implemented via repeated calls to standard procedures for REML estimation in normal theory mixed models, recalculating the working vector and GLM iterated weights at each cycle of iteration. After motivating the procedure by the problem of creating a map of smoothed lip cancer incidence rates for the 56 counties of Scotland, this talk discusses the biases that arise from application to binary outcome data and some possible methods of correcting them. Applications include a longitudinal study of respiratory disease in Indonesian children and a complicated experiment involving the mating of different species of Salamander.



### Raymond Carroll (College Station; USA)

# Asymptotics of the SIMEX Estimator in Nonlinear Measurement Error Models

The SIMEX estimator of Cook & Stefanski (JASA, 1994) has been advertised as a general purpose method for correcting the bias caused by measurement error in the predictor. The method is computationally complex, relying on a mixture of simulation and extrapolation, and takes ~400 times as much computation as standard estimators. This makes such methods as the bootstrap difficult to implement, and in any case it is not known if the estimator is normally distributed asymptotically. We provide a simple almost trivial proof of normality. The proof is extended to practically important variants of the methods.

# Svend Kreiner (København; Denmark)

# Graphical Modelling of Data from Panel Studies: Analysis of Conditional Homogeneity

Test statistics extending the usual test statistics for marginal homogeneity of repeated discrete measurements to test for conditional homogeneity are suggested. Their use in connection with graphical modelling of multivariate repeated measurements where the conditioning variables are given by collapsibility properties that may be read off the independence graph is illustrated by reanalysis of data from the Framingham Heart Study.

# Nanny Wermuth (Mainz; Germany)

# <u>A Longitudinal Study of Child Development - Some Analyses</u> <u>and Interpretation with Graphical Chain Models</u>

The lecture consisted of three parts. An introduction to graphical chain models is given by presenting examples of different situations with three and four variables. A further example of a large longitudinal study and some first analyses are presented to investigate the





question of early determinants of positive or negative development of young children regarding motoric, cognitive and socio-emotional aspects of development. Finally a theorem is presented and proven by which conditional associations and independencies may be directly read off a given directed acyclic graph, which describes a generating process for data.

#### James Robins (Boston; USA)

#### Causal Inference from Longitudinal Studies

The subject-specific data from a longitudinal study consist of a string of numbers. Calculations are performed on these strings and causal inferences drawn. Since the computer algorithms are well-defined mathematical objects, it is important to provide formal mathematical definitions for the sentences expressing the investigators' causal inferences.

In 1986, I proposed a formal theory of counterfactual causal inference that extended Rubin's theory of longitudinal studies. This talk summarizes recent developments in this theory.

In this theory, under the assumption of sequential randomization (equivalently, no unmeasured confounders) the causal effect of a time dependent treatment is non-parametrically identified by the G-computation algorithm formula. However, the G-computation algorithm formula requires one to evaluate high-dimensional integrals. To overcome this difficulty, I propose inverse probability of censoring weighted estimators and G-estimators of structural nested models to estimate causal effects. The relationship of these new methods to those based on the G-computation algorithm are discussed. The effect of aerosolized pentamidine on survival of AIDS patients in an AIDS randomized trial is estimated using the proposed methods.

## Jürgen Läuter (Magdeburg; Germany)

## Stable Statistical Inference under the One-Factor Structure

In the lecture the one-factor model is used for multivariate comparisons and for problems of discriminant, regression, and factor analysis. The one-factor model serves the stabilization of high-dimensional-decisions, especially in cases with highly correlated variables. It is





suitable for problems of longitudinal data analysis. The theory is based on the Bayesian approach. The results are different from the usual strategies of multivariate analysis which lead to ill-conditioned equations. In the calculus of the one-factor model, the p-dimensional inference attains an efficiency that is close to the efficiency of the underlying univariate situation (Läuter, 1992).

The one-factor model and some other stable methods have a great importance for the analysis of clinical studies with multiple endpoints. In the lecture high-effective statistical tests are treated which keep strictly the prescribed type I error rate  $\alpha$  for an arbitrary covariance structure (not only for one-factor structures).

### Per Kragh Andersen (København; Denmark)

## Frailty Models for Multivariate Survival Data

Our starting point is a study of survival among adoptive children and, on one hand, the survival among their biological parents and, on the other hand, the survival among their adoptive parents. To analyse these data we consider a gamma-frailty model for counting processes. Estimation in the model via the EM-algorithm is studied and we suggest using the non-parametric information matrix to obtain variance estimates for the parameter estimates. Some generalisations of the model inspired by the correlated frailty model (Yashin, Vaupel, Iachine, Math Pop Stud, 1994) are outlined.

## Chao Agnes Hsiung (Taipei; Taiwan)

## An Efficient Estimator for Proportional Hazards Models with Frailties

Proportional hazards models with time dependent frailties are constructed by glueing together classical proportional hazards models for counting processes, which include the frailty proportional hazards models in Nielsen et al. (1992) and Holt & Prentice (1974) as special cases. A regular efficient estimator for the relative risk coefficient is obtained. The theory is applied to discuss a matched-pair data of remission lengths of leukaemia patients. A simulation study is also presented.



2.,

#### Gabi Schulgen (Freiburg; Germany)

# <u>Application of Multistate Models in the Analysis of Occurrence</u> and Impact of Nosocomial Infections

Statistical methods for multistate models (Andersen et al., 1993) will be applied to construct scoring systems for hospital-acquired pneumonia and sepsis for patients in intensive care units (ICU), taking into account 'competing events' as death and discharge. The predictive value of the scores is derived by (nonparametric) estimation of the transition probabilities (Aalen & Johansen, Scand J Stat, 1978). Within the concept of 'predictive causality' (Arjas & Eerola, J Stat Plan Inf, 1993) the prediction process can be studied with respect to different patient histories, time of prediction and prediction interval. The impact of nosocomial infections on mortality and duration of ICU stay is evaluated within the same model. Two estimators for the prolongation of hospital stay will be proposed which are functionals of the transition probabilities in multistate models. Results will be compared to some ad-hoc approaches.

The methods are illustrated using data from a cohort study on 756 patients admitted to two different ICUs at the University Hospital in Freiburg.

## Odd O. Aalen (Oslo; Norway)

# Estimating Incidence of HIV Infection Using Information on Times of AIDS Diagnosis and Times of HIV Diagnosis

Backcalculation of HIV incidence rates has traditionally been based on the reported times of AIDS diagnoses. Recently, backcalculation procedures have been extended to make use of information on the times of HIV diagnosis, that is the time of an individual's first positive HIV test, for reported AIDS cases. An approach based on a simple Markov chain model will be discussed here. This model allows the causal influences of events to be considered. In discrete time the likelihood of the observed data can be computed by iteration, even though the Markov model is not homogeneous over time.

Estimation of HIV incidence and prediction will be illustrated on data for England and Wales. In order to judge the combined effect of sampling variation and model uncertainty, a scheme is proposed for combined resampling from the data (bootstrapping) and from a set of possible models.





#### Ian James (Murdoch; Western Australia)

#### On GEE and Multivariate (Cluster) Survival

We consider the use of generalized estimating equations (GEE) for the analysis of multivariate (cluster) survival data with specified parametric marginal distributions. Segal and Neuhaus (Stat Med, 1993) proposed an interesting synthesis of the Poisson likelihood formulation of Aitkin and Clayton and GEE with a patterned correlation structure to facilitate computation using readily available software. However, this approach imposes correlation essentially on the censoring indicators rather than the survival times and can be unstable as a result. A more direct GEE approach with approximations to the correlation structure is proposed. (Work in progress with M Segal and J Neuhaus, UCSF).

#### Thomas Scheike (København; Denmark)

### Modelling Longitudinal Data through Marked Point Processes

We present a general framework for the modelling of longitudinal data with random measurement times based on marked point processes. We construct quite general regression models for longitudinal data, which may in particular include both non-informative censoring that only depend on the past and outside random variation, and informative censoring that depends on the unobserved current measurement as well. The modelling also generalises statistical counting process models. We present a non-parametric Nadarya-Watson kernel estimator of the regression function, and a parametric analysis that is based on a conditional least squares (CLS) criterion. We conclude that the usual non-parametric and parametric regression modelling can be applied to this general set-up, with some small modifications. The presented framework provides an easily implemented and powerful tool for model building for repeated measurements.



## Dankmar Böhning (Berlin; Germany)

## Applications of Zero-Inflated Count Models in Epidemiology

For frequency counts of Poisson-type the situation of extra-zeros often arises. This is especially the case for the DMF-index used in dental epidemiology as will be demonstrated for data from the Belo Horizonte Caries Prevention Study (BEL-CAP). This extra-Poisson variation can be easily explained by a special mixture model, the Zero-Inflated Poisson (ZIP) model. We discuss moment and maximum likelihood estimation of the parameters of the Zip-model. A graphical device is presented which not only summarizes the mixing distribution, but also provides a visual information on the overall mean. This device can be exploited to evaluate and compare various groups. We discuss the inference problem for the overall mean. A confidence interval is developed which uses the properties of the ZIP-model. It compares favourable with the conventional method. Problems of the likelihood ratio test to test the hypothesis of homogeneity against the hypothesis of the ZIP-model are discussed. In this case, the conventional result that the limiting distribution is  $\chi^2$  with one df does not hold. Instead, the limiting distribution is shifted strongly to the left and is in fact given by the mixture  $0.5 \times \chi^2_{(0)} + 0.5 \times \chi^2_{(1)}$ . Alternative of the strong of the stro tive testing approaches and generalizations of the model to include covariates are discussed. A version of the EM-algorithm is provided to find the MLE. This model is then applied to covariates available in the BEL-CAP study.

# Walter Lehmacher (Hannover; Germany)

# <u>Multiple Testing and Interim Analysis in Clinical Trials with</u> <u>Repeated Measures</u>

The two sample design with repeated measures at T time points is considered. Suitable multivariate test statistics like Hotelling's T<sup>2</sup>, O'Brian's sum statistics (OLS and GLS) and rank sum statistics are described. Then, a short review of multiple testing procedures for this T-variate two-sample problem is given including the closed testing principle and the method of a priori ordered hypotheses. Interim analyses, like group sequential and adaptive designs, are available for this multivariable problem, but they can only handle multivariate global tests without multiple testing decisions. For this situation now, where multiplicity arises by the multivariate testing and the multistage interim analyses, methods for combining multiple and interim analy-



ses are presented. They control the experimentwise significance level, are easy and flexible to apply, and can gain more information than the usual global interim analyses, without loss of power concerning the overall hypothesis tests.

### Lutz Edler (Heidelberg; Germany)

# <u>Problems in Evaluating Toxicity Data in Clinical Trials and Some New Approaches</u>

Toxicity is an endpoint of increasing interest in clinical trials and more elaborate methods are required for the appropriate description as well as for the comparative assessment and questions of aetiology. Motivated by an ongoing analysis of toxicity in an adjuvant chemotherapy trial of colo-rectal cancer, basic problems of validity and reliability of the measurement will be addressed and the new Common Toxicity Criteria (CTC) will be presented. Toxicity can be modelled as a time-dependent process of which some characteristics can be estimated. Limited availability of data and restrictive model assumptions cause problems when event history models or finite state Markov processes are applied. More promising are time-to-toxicity proportional hazards models allowing the consideration of dose and accompanying treatment as covariates. Assuming synchronized data, a transition model for longitudinal categorical data is introduced and its restrictions with respect to synchronizing, baseline values and missing values is discussed.

#### Daniel Commenges (Bordeaux, France)

## Score Test of Homogeneity for Survival Data

Motivated by the study of familial aggregation of Alzheimer's disease we have developed a score test of homogeneity for different types of response. Since Alzheimer's disease is an age-dependent disease we may treat age-of-onset data, that is, survival data. The model is a frailty model in which members of the same family share the same frailty which may represent genetic determinants. Under the null hypothesis, the variance of this frailty is equal to zero. We test this hypothesis using a score test which is derived from the marginal partial likelihood. The score statistic can be written as a function of mar-

© (<del>)</del>

tingale residuals. We study the distribution of the score statistic using counting process theory.

We extend this approach to the case where each subject has a given frailty but the frailties are correlated. This model includes as a particular case, the case of groups. The same derivations can be applied to this more general case to compute the score statistic and its distribution. Application in genetics including testing familial aggregation and linkage tests are appealing.

#### John P. Klein (Milwaukee; USA)

# Synthesis of Time Dependent Covariates in a Cox Regression Model

We consider the problem of synthesis of a series of Cox models for intermediate events and terminal events in a complex survival experiment into predictive probabilities for patient outcome based on a dynamic patient history. Two models are considered. The first is based on a time dependent covariate model for intermediate events and assumes proportional hazards throughout. The second relaxes this assumption and fits distinct regressions to each transition. A small Monte Carlo study is presented which compares the two methods. These models are illustrated on a complex bone marrow transplant experiment using data from the International Bone Marrow Transplant Registry.

# Jochen Mau (Düsseldorf; Germany)

#### Multivariate Failure Times

Multivariate failure times arise with multicomponent systems in which each component may fail individually. Or, units of observation are grouped in clusters, and units within the same cluster have a common profile of exposure in comparison with units from other clusters. Such data occur with some dental restorations, and the failures of implants provide the particular context of application.

As a first approach towards data analysis, the suitability of two graphical methods was studied by simulations and in real data sets: a semi-circle diagram (constellation graph, Wakimoto & Taguri, 1978) and a generalized Lexis diagram.





As a first inferential approach, Nelson-Aalen type estimators for increased risk of a failure with more implants present were used in different populations of patients. This statistical approach rests upon partitioned counting processes with multiplicative intensities that are specific to the number of implants at risk.

Rapporteur: Gabi Schulgen





#### Tagungsteilnehmer

Prof.Dr. Odd O. Aalen Section of Medical Statistics University of Oslo P.O.Box 1122, Blindern

N-0317 Oslo 3

Dr. Douglas Altman Medical Statistics Laboratory Imperial Cancer Research Fund Lincoln's Inn Fields P.O.Box 123

GB-London WC2A 3PX

Prof.Dr. Per Kragh Andersen Department of Biostatistics University of Copenhagen Blegdamsvej 3

DK-2200 Kobenhavn N

Dr. Dankmar Boehning Institut für Sozialmedizin und Epidemiologie Freie Universität Berlin Augustastr. 37

12203 Berlin

Prof.Dr. Norman Breslow Dept. of Biostatistics SC-32 University of Washington

Seattle , WA 98195 USA Prof.Dr. Edgar Brunner Abteilung Medizinische Statistik Fachbereich Medizin Universität Göttingen

Humboldtallee 32

37073 Göttingen

Prof.Dr. Raymond Carroll Department of Statistics Texas A & M University

College Station , TX 77843 USA

Dr. Daniel Commenges INSERM U 330 Universite de Bordeaux 2 146 Rue Leo Saignat

F-33076 Bordeaux Cedex

Dr. Ekkehart Dietz Institut für Sozialmedizin und Epidemiologie Freie Universität Berlin Augustastr. 37

12203 Berlin

Dr. Lutz Edler Deutsches Krebsforschungszentrum Biostatistik 0820 Postfach 101949

69009 Heidelberg





Prof.Dr. Ludwig Fahrmeir Institut für Statistik Universität München Ludwigstr. 33/I

80539 München

Prof.Dr. Chao Agnes Hsiung Institute of Statistical Science Academia Sinica

Taipei , 11529 TAIWAN

Dr. Uwe Feldmann Institut für medizinische Biometrie Epidemiologie und Informatik Universitätskliniken des Saarlandes

66421 Homburg/Saar

Dr. Garrett Fitzmaurice Nuffield College

GB-Oxford OX1 1NF

Prof.Dr. Theo Gasser Department of Biostatistics ISPM University of Zürich Sumatrastr. 30

CH-8006 Zürich

Prof.Dr. Hans van Houwelingen Department of Medical Statistics University of Leiden P.O.Box 9604

NL-2300 RC Leiden

Prof.Dr. Ian James Mathematics Murdoch University South Street

Murdoch , WA 6150

AUSTRALIA

45122 Essen

Prof.Dr. Karl-Heinz Jöckel Institut für Medizinische Inf., Biometrie und Epidemiologie Universitätsklinikum Essen

Prof.Dr. Niels Keiding Department of Biostatistics University of Copenhagen Blegdamsvej 3

Prof.Dr. John P. Klein Medical College of Wisconsin Department of Biostatistics 8701 Watertown Plank Road

Milwaukee , WI 53226 USA

DK-2200 Kobenhavn N



Dr. Svend Kreiner Department of Sociology University of Copenhagen 22 Linnesgade

DK-1361 Copenhagen K

Prof.Dr. Jochen Mau Institut für Statistik in der Medizin Universität Düsseldorf Postfach 101007

40001 Düsseldorf

Prof.Dr.Dr.h.c. Klaus Krickeberg U.F.R. de Mathematiques, Statistique et Informatique Universite Rene Descartes 45, rue des Saints Peres

F-75270 Paris Cedex

Rolf Meinert
Institut für Medizinische Statistik
und Informatik
Universität Mainz
Obere Zahlbacherstr. 69

55131 Mainz

Prof.Dr. Jürgen Läuter Abteilung für Biomathematik und Medizinische Informatik Medizinische Akademie Magdeburg Leipziger Str. 44

39120 Magdeburg

Dr. Juni Palmgren Dept. of Cancer Epidemology Radiumhemmet Karolinska University Hospital

S-10401 Stockholm

Prof.Dr. Walter Lehmacher Institut für Biometrie und Epidemiologie Tierärztliche Hochschule Hannover Postfach 711180

30545 Hannover

Dr. James Robins
Dept. of Biostatistics
Harvard School of Public Health
677 Huntington Ave.

Boston , MA 02115

USA

Prof.Dr. Markus Löffler Institut für Medizinische Inform. Statistik u. Epidemiologie Universität Leipzig Liebigstr. 27

04103 Leipzig

Prof.Dr. Patrick Royston Medical Statistics Unit Royal Postgraduate Medical School Ducane Road

GB-London WI2 0NN





Wilhelm Sauerbrei Institut für Medizinische Biometrie und Medizinische Informatik Klinikum der Universität

79104 Freiburg

Stefan-Meier-Str. 26

Werner Vach Freiburger Zentrum für Datenanalyse und Modellbildung Universität Freiburg Albertstr. 26 - 28

79104 Freiburg

Dr. Thomas Scheike Department of Biostatistics University of Copenhagen 3, Blegdamsvej

DK-2200 Copenhagen N

Prof.Dr. Nanny Wermuth Psychologisches Institut Universität Mainz

55099 Mainz

Gabi Schulgen Institut für Medizinische Biometrie und Medizinische Informatik Klinikum der Universität Stefan-Meier-Str. 26

79104 Freiburg

79104 Freiburg

Prof.Dr. Alice S. Whittemore Dept. of Health Research and Policy Stanford University School of Medicine Stanford , CA 94305-5092

Prof.Dr. Martin Schumacher Institut für Medizinische Biometrie und Medizinische Informatik Klinikum der Universität Stefan-Meier-Str. 26







