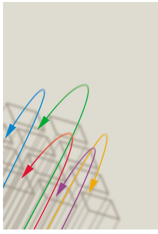

Progress Report 2010

IAP-Network in Statistics

Contract P6/03

March 31, 2011

http://www.stat.ucl.ac.be/IAP/PhaseVI

Contents

1 Accomplished Research Projects

1.1 Introduction

The research project has been built up around five work packages. Table 1 below gives the *main* contributors to each work package and indicates per package the partner that is coordinating the work.

Work package	Contributing partners
WP1: Multivariate data with qualitative constraints	UCL, KUL-1*, UH, UJF, UU, USC
WP2: Temporally and spatially related data	UCL*, KUL-1, UG, UJF, UU, USC
WP3: Incomplete data	UCL, KUL-1, KUL-2, UG, UH*, USC, LSHTM
WP4: Data with latent heterogeneity	UCL, KUL-1, KUL-2*, UH, UJF, UU, USC
WP5: Highdimensional and compound data	UCL, KUL-1, KUL-2, UG*, UH, UJF, UU, USC

Table 1: *Main contributors per work package, and coordinating partner per work package (indicated with a *).*

In the subsections below we describe the progress that has been made in the various work packages. For each of the work packages we indicate interactions with research results in other packages. The references mentioned in the text can be found at the end of this report.

1.2 Work package 1: Multivariate data with qualitative constraints

1.2.1 Boundaries, frontiers, and efficiency and productivity analysis

Estimation of a support of a (univariate or multivariate) density function, or estimation of a more general support is a topic of research in the network since the start.

Kneip, Simar and Van Keilegom (2010) propose a new approach to estimate the boundary of the support of a univariate density when measurement error is present. The novelty of their approach lies in the fact they do not assume that the variance of the noise is known. They propose a method for simultaneous estimation of the boundary point and the variance. The asymptotic consistency and the rate of convergence of the estimators are established.

A challenging issue when using envelopment type estimators is the robustness to extremes and outliers. Related to this aspect, Daouia and Gijbels (2011) deal with the important question of how two families of partial production frontiers, an expected maximum output frontier of order m and a conditional quantile-type frontier of order α, are linked and how robust they are. Insights from this study brought up a new class of quantities which is applied in estimating frontier cost models in Daouia and Gijbels (2010).

1.2.2 Nonparametric and semiparametric estimation of (irregular) curves and surfaces, and estimation under qualitative constraints

Unknown functions and surfaces are not always smooth, and may show different types of irregularities in different regions. In Desmet and Gijbels (2011) local linear fitting is a basic starting point for estimating regression curves with irregularities in the function itself as well as in the derivative. Desmet, Gijbels and Lambert (2010) deal with nonparametric estimation of irregular density functions.

Another way to deal with different degrees of smoothness of curves is to use penalized regression techniques with an appropriate choice of the penalty function. A specific penalty choice influences the optimization problem to be solved as well as the existence of a (unique) solution. In Antoniadis, Gijbels and Nikolova (2011) nonparametric regression estimation is considered in extended generalized linear models using penalized splines with non-quadratic penalties. This unified approach allows to study the optimization problem and asymptotic properties of the resulting estimator for a whole class of penalties. The papers Gijbels and Verhasselt (2010a,b) look into penalized regression with a data-driven choice of difference type of penalty. The powerful P-splines techniques are also used as a basic ingredient in a variable selection procedure for additive models and varying coefficient models in Antoniadis, Gijbels and Verhasselt (2010a,b).

Gijbels, Prosdocimi and Claeskens (2010) look into the problem of estimation of under- or overdispersion. Especially overdispersion of data is a well-known phenomenon when analysing data. A unified approach to nonparametrically estimating mean and dispersion function is provided. These techniques were used to analyze Italian abortion rate data in Gijbels and Prosdocimi (2011). Croux, Gijbels and Prosdocimi (2010) developed robust procedures for mean and dispersion function estimation.

In financial applications, one needs to deal with a lot of irregular behaviour (e.g. of stock market prices, or stock indices). Among the key issues is the study of the volatility function. In financial mathematics and engineering, the erratic behaviour is often modelled by using particular classes of stochastic processes, such as Lévy processes. See for example Schoutens and Van Damme (2010) and Fang *et al.* (2010), among others.

1.2.3 Nonparametric (location-scale) regression and semiparametric regression

The receiver operating characteristic curve (ROC curve) is a tool of extensive use to analyze the discrimination capability of a diagnostic variable in medical studies. González-Manteiga, Pardo-Fernández and Van Keilegom (2010) model the effect of the covariate over the diagnostic variable by means of nonparametric location-scale regression models. They propose a new nonparametric estimator of the conditional ROC curve and study its asymptotic properties.

Neumeyer and Van Keilegom (2010) estimate the error distribution in a nonparametric location-scale regression model with multivariate covariates. Modeling heteroscedasticity in semiparametric regression can improve the efficiency of the estimator of the parametric component in the regression function, and is important for inference problems such as plug-in bandwidth selection and the construction of confidence intervals. Van Keilegom and Wang (2010) propose a general method to estimate the dispersion function in a semiparametric way, and obtain generic conditions under which the proposed estimator satisfies certain asymptotic properties.

1.2.4 Multivariate data, robust analysis and nonparametric inference

Among the main research themes in recent years in the network, is the analysis of elliptically distributed multivariate data, possibly corrupted with outliers. For high-dimensional data, dimension reduction will often be applied by means of a principal component analysis (PCA) method. Work on robust PCA methods is reported on in WP 5.

1.2.5 Functional data analysis

Clustering techniques for functional data are studied in Claeskens, Hubert and Slaets (2010). Ferraty, Van Keilegom and Vieu (2010a) consider a functional nonparametric regression model, in which the response is univariate and the covariate is functional. They show the asymptotic validity of a naive and a wild bootstrap procedure, and study an application to the construction of confidence intervals. In Ferraty, Van Keilegom and Vieu (2010b) the challenging context of a functional nonparametric regression model, in which both the response and the covariate are functional, is studied.

1.2.6 Modelling and measuring of dependencies and copula functions

Often, the dependence structure between two (or more) random variables is itself influenced by another variable, a covariate. Such a dependence structure can be modeled via a so-called conditional copula function. Estimation of conditional copulas was discussed in Gijbels, Veraverbeke and Omelka (2010) and Veraverbeke, Omelka and Gijbels (2011). Bootstrapping procedures for conditional copulas are developed in Omelka, Veraverbeke and Gijbels (2010).

Among the important questions in modeling and measuring of dependencies, is the choice of an appropriate copula function or family. Some copulas exhibit interesting particular properties, such as positive quadrant dependence or negative quadrant dependence. In Gijbels, Omelka and Sznajder (2010) testing procedures for testing the null hypothesis of positive quadrant dependence are developed.

Hafner and Reznikova (2010) contributed towards efficient estimation of a semi-parametric dynamic copula model. Elliptical copulas were the focus of Manner and Segers (2011).

1.2.7 Extreme value theory and copulas

Among the typical distributions in extreme value theory are Pareto distributions. Hubert, Dierckx and Vanpaemel (2010) study how to detect influential data points in such distributions. Gudendorf and Segers (2010, 2011) paid attention to non-parametric estimation in extreme-value copulas of arbitrary dimension.

1.3 Work package 2: Temporally and spatially related data

Temporally and spatially related data are encountered in several sciences and their statistical analysis has a long history. However, the development and investigation of models for temporal or spatio-temporal correlated data under non standard assumptions, in discrete or continuous times and in one or higher dimensions, still encounters many important but unsolved problems. The main achievements of the research network for temporally and spatially related data can be subdivided into three categories: complex time series data, spatially related data and continuous time models. Most Belgian partners (UCL, KUL-1, KUL-3) and most European partners (UJF, USC) have contributed to WP2.

1.3.1 Statistical analysis of complex time series data

One type of complexity often present in real data is due to the irregularity of the spectral density function of a stationary time series. Nonparametric estimation of a spectral density with improved estimation at the peaks is dealt with in Desmet and Gijbels (2010). The proposed method relies on local linear fitting techniques. These techniques are also the basic ingredient for estimating nonparametrically the volatility function in Casas and Gijbels (2010). Local polynomial fitting (in a non time series setup) is among the methods studied in detail and applied in various contexts in WP1. An alternative very popular technique is thresholded wavelet estimation. Freyermuth et al. (2010) consider tree-structured wavelet estimation in a mixed effects model for spectra of replicated time series. An other source of complexity very often met in the applied sciences is the deviation from stationarity. In Roueff and von Sachs (2011) the estimation of the time-varying long memory parameter is considered in the presence of local stationarity. Bhm et al. (2010) propose a classification method of multivariate non-stationary signals. On the other hand using local linear/polynomial fitting Croux *et al.* (2010) derive methods for robust forecasting in nonstationary time series. Regularly varying time series in Banach spaces has been studied in Meinguet and Segers (2010).

Censoring of time series data is another problem sometimes encountered with real data. Teodorescu et al. (2010) study generalized time-dependent conditional linear models under left truncation and right censoring. In nonparametric location-scale regression models Heuchenne and Van Keilegom (2010) consider the estimation based on censored data, while Lambert (2010) assumes interval censored data. Antoniadis et al. (2010) study the Dantzig selector for regression with right censored data. On the other hand, the construction of confidence intervals and goodness-of-fit tests are also important to validate models on data. Smoothed empirical likelihood confidence intervals for the relative distribution with left-truncated and right-censored data are proposed in Molanez-Lopez et al. (2010). These papers are very much related to the research topics studied under WP1 and WP3.

A lot of work is done in the network in developing models for financial data. The GARCH and stochastic volatility (SV) models are two competing, well-known and often used models to explain the volatility of financial series. Hafner and Preminger (2010) propose strong decision rules for deciding between GARCH and stochastic volatility. By employing ploynomial fitting techniques Casas and Gijbels (2010) estimate nonparametrically the volatility function.

A domain where time series in discrete time arise naturally is that of ability testing: A set of items (e.g., questions) is given to a sample of examinees (e.g., pupils or students) with the instruction to solve them. Almost always there is a time limit, and this induces time pressure.

This time pressure typically plays a larger role towards the end of the test, with as a result that examinees may start to drop items. We have investigated whether it is possible to detect persons that are particularly sensitive to time pressure using local influence diagnostics (Goegebeur, De Boeck, and Molenberghs, 2010).

One important aspect of the research proposal was an extension of the univariate analysis to the case of multiple time series. Hafner and Reznikova (2010) consider efficient estimation of a semiparametric dynamic copula model. Segers (2010) derives weak convergence of empirical copula processes under non-restrictive smoothness assumptions. The development of working methods for analyzing the correlation between multiple series and its application to financial data has been of interest. For example in the context of insurance data, Manner and Segers (2010) characterize tails of correlation mixtures of elliptical copulas. Extreme values of a stationary, multivariate time series may exhibit dependence across coordinates and over time. Genest and Segers (2010) study the covariance of the asymptotic empirical copula process, a new and potentially useful tool.

1.3.2 Spatially related data

Another main objective of this work package is the analysis of spatial data. Crujeiras and Van Keilegom (2010) study the asymptotic and finite sample properties of an estimator of a nonlinear regression function when errors are spatially correlated, and when the spatial dependence structure is unknown. The proposed method is based on a weighted nonlinear least squares approach, taking into account the spatial covariance. Weak consistency of the regression parameters estimator is derived, along with its asymptotic Gaussian limit. There is a growing interest in improving the level of knowledge of spatial and spatio-temporal processes using spectral techniques. Crujeiras et al. (2010) suggest to extend two different goodness-of-fit testing techniques for the spatial spectral density. The first approach is based on a smoothed version of the ratio between the periodogram and a parametric estimator of the spectral density. The second one is a generalized likelihood ratio test statistic, based on the log-periodogram representation as the response variable in a regression model. As a particular case, they provide tests for independence. Asymptotic normal distribution of both statistics is obtained, under the null hypothesis. Finally, in Antoniadis et al. (2010) an application to real data is provided. The authors consider spatio-temporal prediction for West Africa monsoon.

1.3.3 Continuous time models

Concerning single level continuous time modes, we contributed to a theoretical analysis of Weibull distribution based models for time to event data (Tuerlinckx, 2010). Next, for multilevel continuous time models, in a first line of research, a multilevel or hierarchical Ornstein-Uhlenbeck process model has been developed together with a Bayesian inference algorithm to analyze irregularly spaced multivariate time series from different subjects simultaneously. Because of the complexity of the Bayesian inference algorithm, we have investigated in Oravecz and Tuerlinckx (in press) whether we could not make use of simpler algorithms that had been developed before to estimate linear mixed models. It turns out that there are some exact and approximate relations between the models, but that the Ornstein-Uhlenbeck process allows for some interesting sources of random variability that the linear mixed model cannot incorporate. Second, diffusion-based models for decision making have been extended hierarchically, so as to allow for analyses of experiments with many subjects and fewer decision trials. Given the complexity of some stimuli (e.g., words, pictures, etc.), it is not possible to make subjects undergo numerous identical trials; this makes a hierarchical model a necessity. The hierarchical models have been developed in a Bayesian context because of the computational tools that can be applied within this framework. In a first study, a hierarchical Ratcliff diffusion model is presented and explained from a psychometric point of view (Vandekerckhove, Tuerlinckx, and Lee, in press). This has led to a modeling framework that is highly flexible and easy to work with for applied researchers. Next, the model has been applied in Vandekerckhove, Verheyen, and Tuerlinckx (2010) to a speeded semantic categorization task in which both subjects and items were random (i.e., a crossed random effects model). Third, when latent quantities with dynamical properties are involved, state space models offer an interesting modeling framework. However, for simultaneous modeling of individual and population differences, a hierarchical extension of the basic state space model is necessary. We introduced a very versatile Bayesian hierarchical model with random effects for the system parameters (Lodewyckx, Tuerlinckx, Kuppens, Allen, and Sheeber, 2011). This model can be applied to any time series for multiple systems (where a system can represent any type of entity).

1.4 Work package 3: Incomplete data

The work on incomplete data can be captured under the broad headings of: (1) complex modeling approaches for missing data; (2) sensitivity analysis tools; (3) censored survival data; and (4) general incomplete data structures.

1.4.1 Complex modeling approaches for missing data

For complex modeling approaches for missing data, work has been done to provide a paradigm shift in the application of missing data methodology. In particular, we mention the work done under the auspices of the National Academy of Sciences (US) and commanded by the Food and Drug Administration (Little *et al* 2010), to which Geert Molenberghs has contributed. Likewise, Burzykowski and Carpenter contributed to the PSI working group on incomplete data (2010). A variety of contributions has been made to reference texts. Such work at the same times avoids continuing the use of sub-optimal methods and advocates the use of appropriate methods, such as direct likelihood and direct Bayesian inferences, inverse probability weighting methods, and multiple imputation. Molenberghs (2010) wrote an encyclopedia entry on incomplete data in clinical and epidemiological studies, as well as one on sensitivity analysis.

Aerts *et al* (2010) studied missingness and informative cluster sizes in a clustered-data context. Jansen and Molenberghs (2010) focused on pattern-mixture model for non-monotone missing data. Molenberghs *et al* (2010) published on pseudo-likelihood methodology for incomplete data, using doubly robust ideas. Pseudo-likelihood methodology was also the topic of Troxel *et al* (2010). Doubly-robust methodology was also put forward by Moerkerke, Vansteelandt, and Lange (2010) and Lee, Mangseok, and Kenward (2010). Sotto *et al* (2010) developed MCMC-methodology for incomplete, non-monotone longitudinal data. Tsonaka *et al* (2010) developed methods for non-randomly missing data. A broad study of unobservables was provided by Verbeke and Molenberghs (2010). Cortiñas *et al* (2010) studied software for incomplete data. Inverse probability weighting was the topic of Vansteelandt, Carpenter, and Kenward (2010). Partially linear models for incomplete data were studied by Bianco *et al* (2010). Multiple imputation was studied by Carpenter, Roger, and Kenward (2010), Spratt *et al* (2010), and White *et al* (2010). Carpenter and Plewis contributed to a book on non-response in longitudinal data. Daniel *et al* (2010) used graph theory in incomplete data problems.

1.4.2 Sensitivity analysis tools

Sensitivity analysis tools have been developed to assess the impact of untestable assumptions about the incomplete data distribution, given the observed data, on important scientific conclusions. Contributions were made by Poleto *et al* (2010). Creemers *et al* (2010) provided sensitivity analysis tools for the shared-parameter model.

1.4.3 Censored survival data

In the category of censored survival data, both parametric and non-parametric methods have been a topic of interest. Lambert (2011) considered an additive location-scale model for interval censored data. Cetinyurek and Lambert (2011) focused on smooth estimation of survival functions and hazard ratios from interval-censored data using Bayesian penalized B-splines. Gaddah and Braekers (2010) treated testing under the extended Koziol-Green model. The same model was also the focus of Veraverbeke (2010). Heuchenne and Van Keilegom (2010) and Van Keilegom *et al* (2011) focused on location-scale regression models for censored data. Molanez-Lopez, Cao, and Van Keilegom (2010) and Teodorescu and Van Keilegom (2010) studied the combination of left-truncated and right-censored data. Gijbels (2010) wrote an encyclopedia entry on the Kaplan-Meier estimator. Pérez González, Vilar Fernández, and González-Manteiga (2010) studied a non-parametric variance function with missing data.

Furthermore, attention has been given to longitudinal, recurrent, multivariate, and otherwise correlated survival times. To this end, work has been done in the copula and frailty paradigms. Meira-Machado (2010) *et al* considered non-Markov models for successive survival times. Hens and Wienke (2010) studied the correlated gamma frailty model for current status data. Markov models for successive survival times were treated in Meira-Machado (2010). Left truncated dependent data were of interest in de Uña Álvarez, Rodríguez-Casal, and Liang (2010).

1.4.4 General incomplete data structures

General incomplete data structures have received attention. In ability testing, often a set of items is given to a sample of examinees with the instruction to solve them. Almost always there is a time limit, and this induces time pressure. As a result, towards the end of the test examinees will typically start to drop items, giving rise to missing data. Goegebeur, De Boeck, and Molenberghs (2010) have developed a probabilistic model for test speededness and have investigated whether it is possible to detect persons that are particularly sensitive to time pressure using local influence diagnostics (see also WP2).

In the area of joint modeling of longitudinal and survival data, contributions were made by Dejardin, Lesaffre, and Verbeke (2010) and Rizopoulos, Verbeke, and Molenberghs (2010). Survival and incomplete dat were combined in Nur *et al* (2010).

1.5 Work package 4: Data with latent heterogeneity

In many situations, statistical models are used that assume the presence of latent, unobservable, structures to explain the variability observed in the data. Exactly the fact that those structures, by definition, can never be observed poses particular problems with respect to identifiability, as illustrated in Verbeke and Molenberghs (2010). Nevertheless, those models have proven to be very useful in various contexts, and some illustrations of this will be summarized in Section 1.5.1. Many statistical models employ latent structures to generate association structures in correlated data settings. One particular application is the construction of models for the joint analysis of multiple outcomes. An number of examples will be discussed in Section 1.5.2. Many standard statistical models are based on very strict underlying assumptions. So are many mixed models. In specific applications, model assumptions need to be relaxed in order for the models to be realistic for specific data sets at hand. Examples are presented in Section 1.5.3. An important class of latent heterogeneity models has been developed within the framework of item response theory (IRT). Specification and parameter estimation of some of those IRT models is discussed in Section 1.5.4. In the case that heterogeneity shows up at the person as well as the item side, crossed random effects models are most useful: see Section 1.5.5. Finally, in Section 1.5.6 we discuss several models for multivariate data on time dynamical processes; individual differences and trial-to-trial heterogeneity in such data are captured in those models through different kinds of hierarchical extensions.

1.5.1 Applications of models with latent structures

When the latent structures are assumed to be of a continuous nature, one traditionally refers to the models as mixed-effects models. Often the models are linear, generalized linear, or nonlinear regression models in which some parameters are assumed to be sampled from a continuous ditribution, often assumed (multivariate) normal. Those random effects can have various interpretations such as levels of general pracitioners in a multicenter study (Goderis et al. 2010), levels of hospitals in a study comparing various hospitals with respect to patient safety (Diya et al. 2010), or gene effects in the context of transgenic improvement of a commercial crop (De Wolf et al. 2010).

In Komárek et al. (2010), mixed models have been used to classify subjects into two or more prognostic groups using longitudinally observed values of markers related to the prognosis. The proposed method proceeds in two steps as described earlier in the literature. First, multivariate linear mixed models are fitted in each prognostic group from a training data set to model the dependence of markers on time and on possibly other covariates. Second, the fitted mixed models are used to develop a discrimination rule for future subjects. The method improves upon existing approaches by relaxing the normality assumption of random effects in the underlying mixed models. Namely, a heteroscedastic multivariate normal mixture is assumed and inference is performed within the Bayesian framework using the Markov chain Monte Carlo methodology. Software has been written for the proposed method and it is freely available. The methodology is applied to data from the Dutch Primary Biliary Cirrhosis Study.

1.5.2 Latent structures for joint modeling

Many subject-matter research questions can only be answered by modeling several outcomes jointly. Assuming a mixed or mixture model for each outcome separately and allowing the latent structures to be common or correlated for the various outcomes, implies joint models in a very straightforward way.

In Zhang et al. (2011) latent variables are used to derive the joint distribution of bivariate ordinal outcomes, measured longitudinally. Specifically, the ordinal outcomes are related to a bivariate latent variable, which is modeled over time using a linear mixed model. Random effects terms are used to tie together all repeated observations from the same subject. The cross-sectional association between the two outcomes is modeled through the correlation coefficient of the bivariate latent variable, conditional on random effects.

1.5.3 Model extensions and/or flexible models

Models with flexible assumptions about latent structures have been applied in many different contexts. In Ghidey, Lesaffre and Verbeke (2010), some recently suggested approaches to flexibly estimate the random effects distribution in linear mixed models are compared. While mixed models are particularly useful to model mean and/or dispersion functions flexibly, allowing for overdispersion, alternatives have been studied in Gijbels, Prosdocimi, and Claeskens (2010), who also compared their approach to mixed models.

1.5.4 Parameters of item response models

Item response models are probabilistic models for responses from a sample of persons to a set of items or questions. Latent heterogeneity is mostly situated at the person side to allow for differences in ability or proficiency of different persons. Regarding the parameters of such models, we have studied two groups of (partly interrelated) issues.

First, differential item function (DIF) refers to the phenomenon that questions of a test may mean something different to different groups. It is important to detect DIF in order to avoid test bias (and potential law suits), see for example Janssen (2010). We have constructed a novel model that incorporates DIF (based on a Rasch model with random item difficulties besides the common random person abilities) and that assumes a mixture model for the item difficulties such that the items may belong to one of two classes: a DIF or a non-DIF class (Frederickx, Tuerlinckx, De Boeck, and Magis, 2010). Moreover, we have built an R package to apply a variety of DIF methods and compared them to one another; see Magis, Bland, Tuerlinckx, and De Boeck (2010).

Second, we contributed a new method to estimate the ability levels of persons in an item response model in Magis and Raîche (2010), and proposed a correction to the Bayesian approach of maximum a posterior estimation in order to reduce the bias when estimating extremely large or small ability levels (Magis, Bland, and Raîche, in press).

1.5.5 Crossed random effects

Crossed random effects models are an exciting tool to deal with cases in which latent heterogeneity is present in two modes of the data. As an example, one may think of person by item data with latent heterogeneity in both the person and the item mode. Linking up with this example, a crossed random effects approach has been shown to be very fruitful in the context of item response models, see Janssen (in press), which includes a latent regression model for the item difficulties. Crossed random effects can also be included in the modelling of repeated decision trials, see Vandekerckhove, Verheyen, and Tuerlinckx (2010) and Vandekerckhove, Tuerlinckx, and Lee, in press.

1.5.6 Multilevel longitudinal data

Often, repeated measures are taken from a sample of persons, with the researcher wanting to learn something about the dynamic properties of the population from which the persons are sampled and about the individual differences. Models to address these issues have been studied, and the most important results will be summarized below. There is of course an obvious link between this work and WP2.

In a coordinated fashion, various types of multivariate multilevel continuous time models have been developed, and associated Bayesian inferential techniques have been proposed, see Oravecz and Tuerlinckx, in press. In this paper, the core model is the Ornstein-Uhlenbeck diffusion process, for which key parameters (mean, variances, autocorrelations) are allowed to be subject to latent heterogeneity across individuals. Another core model used, is the state space model that we extended hierarchically using a Bayesian hierarchical model with random effects for the system parameter (see Lodewyckx, Tuerlinckx, Kuppens, Allen, and Sheeeber, in press).

Latent heterogeneity is also allowed in another type of continuous time diffusion process models, called the Ratcliff diffusion model for decision making. In this model, the latent heterogeneity implies that some of the key parameters of the model may vary from trial to trial, through which a number of crucial experimental observations can be accommodated (e.g., both fast and slow error responses in a decision task). Regarding our contributions, we have extended the model to allow also for latent between-person heterogeneity (Vandekerckhove, Tuerlinckx, and Lee, in press).

Finally, we also dealt with multilevel longitudinal data within the context of item response models that capture speededness effects. Results in this area are reported under WP2 and WP3.

1.6 Work package 5: Highdimensional and compound data

1.6.1 Bioinformatics

We have continued our work on phylogenetic inference (i.e. inferring evolutionary models and trees on the basis of observed nucleotides of existing species, but not their ancestors). Most current approaches assume for computational reasons that the evolution at a given locus is independent of evolution at neighboring loci. We have therefore focused on the development of context-dependent evolutionary models which allow for the evolution of a given site to depends upon its ancestor and that ancestor's immediate flanking sites. Because such dependency pattern cannot be imposed on the root sequence in an evolutionary tree, we consider the use of different orders of

Markov chains to model dependence at the ancestral root sequence. MCMC analyses reveal strong support for second-order Markov chains at the ancestral root sequence, thus resulting in drastic model improvements. We further evaluated the importance of non-reversible evolutionary models when analyzing context-dependence. The relevance of this is motivated by the fact that context-dependence is typically thought to be present in the well-known CpG-methylation-deamination process in mammalian evolution, which is inherently non-reversible. We find that non-reversible context-dependent models can drastically increase model fit when compared to independent models and this on two primate non-coding datasets.

Handling outliers in high dimensional data is more difficult than in small to moderate dimensional data because it becomes unrealistic to assume that a majority of the observations is completely free of any contamination. Van Aelst, Vandervieren, and Willems (2011) address this problem for the Stahel-Donoho estimator of multivariate location and scatter by determining separate weights for each component of an observation instead of case weights that hold for all components of an observation simultaneously.

In a last bioinformatics application of the UG group, in plant breeding, Dewolf et al.. (2010) use mixed model techniques to deal with sources of variability in the data-analysis of phenotyping experiments with transgenic rice. Somaclonal variability and insertion variability are separated from each other to assess the effect of the inserted gene and its variability in a correct way. As an extension to this work, design aspects of such trials aiming at evaluating a specific gene effect and producing a productive mutant were studied.

1.6.2 Data mining

In many situations one observes a lot of variables, and a selection of the most important variables is a key task. Variable selection in an additive model as well as in a varying coefficient modelling structure has been accomplished in Antoniadis, Gijbels and Verhasselt (2010a,b) using a combination of the nonnegative garotte technique and P-splines estimation.

For high-dimensional data dimension reduction will often be applied by means of a principal component analysis (PCA) method. A new kernel PCA method and diagnostic tools to detect influential data points were proposed in Debruyne, Hubert and Van Horebeke (2010). Support-vector machine techniques are among the methods used for high-dimensional data. Such kernel-based regression methods and their robustness were studied in Debruyne *et al.* (2010).

1.6.3 Psychometrics

Highdimensional object by variable data
To deal with vast data sizes, we have been working on a broad range of methods that imply a reduction of the variables (and possibly also of the objects), with this reduction being either categorical or dimensional in nature (clustering resp. dimension reduction).

A particular focus of research was on models that simultaneously perform a clustering and a dimension reduction: We studied and compared such models from a theoretical point of view (Timmerman, Ceulemans, Kiers, and Vichi, 2010), and developed generic frameworks for them (see Timmerman and Ceulemans, 2010, and Van Mechelen and Van Deun, 2010).

For the case of clustering models that simultaneously group objects and variables in overlapping clusters, a particularly useful extension was proposed that allows to unveil the nature of interactions as present in the data (see Schepers and Van Mechelen, 2010).

The methods that have been developed were mainly deterministic, but in addition a generic approach has been proposed to expand deterministic models into their immediately related stochastic counterparts (Gelman, Leenen, De Boeck, and Poblome, 2010).

As regards algorithmic developments, special attention was paid to making user-friendly software with a graphical user interface (Wilderjans, Ceulemans, Van Mechelen, and Depril, in press).

Multiway data

Almost all methods for the analysis of multiway data assume that the comparison of any two entries in the data array under study reflects or represents meaningful content-specific information. Violations of this assumption may imply data-analytic results that are of doubtful quality at best and worthless in the worst-case scenario. In Van Mechelen and Smilde (in press), the assumption of comparability is clarified, a number of reasons of why the assumption is very often violated in practice are listed, and several possible approaches to deal with problems of comparability are reviewed. It is concluded that any satisfactory solution to comparability problems requires a very careful reflection about the data collection and the ultimate goal of the data analysis.

Coupled data

As a consequence of our information society, not only more and larger data sets become available, but also data sets that include multiple sorts of information regarding the same system. Such data sets can be denoted by the terms coupled, linked, or multiset data, and the associated data analysis can be denoted by the term data fusion.

Within the framework of this IAP program, we successfully developed a novel model for data fusion problems (see Wilderjans, Ceulemans, Van Mechelen, and van den Berg, in press). Our study of model interrelations further culminated in an overall conceptual framework for data fusion, along with a generic model that subsumes an important subset of specific (existing as well as to be developed) data fusion approaches as special cases, see Van Mechelen and Smilde, 2010.

On the level of data analysis we dealt with two types of issues:

1. Weighting: When dealing with the estimation of a global model to coupled data at hand, an important question reads how the information from the different blocks is to be integrated into an overall loss or objective function. This question becomes especially important if the data blocks differ rather strongly in size and/or if they are subject to rather different amounts of noise. To deal with this issue, we studied different possible weighting schemes and likelihood-based approaches, making use of synthetic data. This made clear that: (a) in a homogeneous error setting, it is advisable not to correct for block size by downweighting larger blocks (Wilderjans, Ceulemans, Van Mechelen, and van den Berg, in press), and (b) when data blocks differ considerably both in size and presumed noise level, one may benefit from using a likelihood-based downweighting of error-prone data blocks (Wilderjans, Ceulemans, Van Mechelen, and van den Berg, in press).

2. Model selection: We developed a numerical convex hull based procedure for selecting among several multilevel simultaneous component solutions (Ceulemans, Timmerman, and Kiers, in press).

Finally, we compared analyses of separate data blocks with a data fusion approach in a study on gene co-expression network analysis (van den Berg et al., 2010).

2 Network Activities

2.1 Web site and newsletter

All activities of the IAP-statistics network can be followed very closely from our web site. The address of the web site is

`http://www.stat.ucl.ac.be/IAP/PhaseVI`.

The web site contains e.g. the following information:

- Our logo
- Call for applications
- Description of the project
- List of scientific personnel working under the IAP project
- List of IAP members and their email addresses
- List of visitors
- Research activities (workshops, seminars, short courses,...)
- Defended theses and theses in preparation
- Downloadable technical reports, list of publications and list of books
- Annual reports and reports of scientific meetings
- Contact details

2.2 Scientific meetings

2.2.1 Annual workshop

The 2010 annual workshop entitled 'Cross-fertilization within the IAP network' was organized by KUL-1 (Leuven, 18-19 November 2010). The workshop focused on cross-links between the different work packages and on recurrent research themes that are being studied by several partners from different angles (such as copulas as a key tool to deal with dependencies, and variable selection / sparseness methods to deal with high dimensional data.) More information can be found on the web page of the workshop:

`http://ppw.kuleuven.be/okp/iapworkshop4/.`

The 2011 annual workshop will be organized by the UCL group and will take place in November 2011. The goal of this workshop will be to focus on major breakthroughs in the network over the last 5 years, and also on important and promising future research directions the network likes to take a lead in in the near future.

2.2.2 Meetings

The following meetings were organized by the network in 2010:

- The European partner of Santiago de Compostela (Spain) has organized, with IAP support, the METMAV Conference from 30 June to 2 July 2010, which is an international workshop on spatio temporal modelling. The purpose of this conference is to promote the development and application of spatio-temporal statistical methods in different fields related to environmental sciences. A total of 130 participants from 19 countries participated to the meeting.

- On October 14-15, 2010, the 18th Annual Meeting of the Belgian Statistical Society took place in Spa. The IAP network took part in the organization of this meeting, and organized one session.
- The University of Hasselt (UH) organizes on May 19-20, 2011 a conference on 'Recent Advances in Statistics and Probability'. The conference is in honor of Noël Veraverbeke, who retires at the end of this academic year The list of invited speakers includes four IAP members: L. Duchateau, I. Gijbels, W. González-Manteiga and I. Van Keilegom.

2.3 Organization of the network: administrative meeting

The annual administrative meeting took place on 18 November 2010 in Leuven, during the annual workshop. The meeting was attended by : A. Antoniadis (UJF), L. Duchateau (UG), P. Eilers (EMC), I. Gijbels (KUL-1), W. González-Manteiga (USC), J. Johannes (UCL), M. Kenward (LSHTM), I. Van Keilegom (UCL), I. Van Mechelen (KUL-1) and N. Veraverbeke (UH). The participants to this administrative meeting discussed issues related to past and future scientific activities organized by the network, scientific collaborations in the network, work valorization (such as web page, reports, ...), network organization, management and visibility. A detailed report of this meeting was sent to all participants.

2.4 Collaborations, working groups and seminars

2.4.1 Collaborations

The IAP network is working on a broad range of research topics in statistics. There is a large number of scientific collaborations within the network, as can be seen from the list of technical reports and publications (see Section 3, and in particular Subsection 3.10, where all joint technical reports and publications are collected). Below, we mention a few examples of ongoing collaborations between members of different teams of the network.

- G. Molenberghs (UH/KUL-2) and G. Verbeke (KUL-2) are preparing a book on mixed models.
- UH, KUL-2, and LSHTM are involved in a Taylor & Francis book project on missing data, together also with North Carolina State University.
- There is a continuing collaboration between KUL-1 and UH on the use of generalized linear mixed models in psychometric applications and test theory (including the study of test speediness).
- There exists a lot of collaboration between KUL-1 and UJF on variable selection and sparseness within the context of mixture, component, and unfolding models.

- P. Janssen (UH), L. Duchateau (UG), C. Legrand (UCL) and PhD students of their respective research groups are working together on projects related to frailty models and competing risks in survival analysis.
- KUL-1 and KUL-2 are both partners in a KUL Center of Excellence for computational systems biology. The teams have made several methodological contributions (of immediate relevance for WP5) in the context of statistical bioinformatics, which have led to new biological insights.
- R. Crujeiras Casais (USC) has stayed with the UCL group for 10 days in November 2010. She worked on spatial regression with censored observations.
- M. Garcia Magarios and W. González Manteiga (USC) have collaborated with A. Antoniadis (UJF).

2.4.2 Working groups

Below are a few examples of active working groups in the network. They are an important tool to stimulate interactions between network partners, and to stay informed of the research achievements of other partners of the network.

- *Variable selection and sparseness*
  Members of the KUL-1 (in particular I. Gijbels and A. Verhasselt) and of UJF (A. Antoniadis) have extensive collaborations on the development of semi- and nonparametric methods for complex data, such as heavy noisy data. One of the aspects of the work is exploring regularization techniques. These collaborations are situated in WP1 and on the interface between WP1 and WP4.
- *Frailty models*
  The working group on frailty models and related models with members from UH, UG, UCL, and KUL-2 continues to meet on a regular basis. The main research topics discussed in this working group are related to competing risks and to transformation models for survival data (where the transformed cumulative hazard is modeled as a function of covariates in a linear way and where such models are extended to time-varying covariates and frailties).
- *Bioinformatics*
  I. Van Mechelen (KUL-1) and G. Verbeke (KUL-2) are both partners in a KUL Center of Excellence for computational systems biology. Their teams have made several methodological contributions (of immediate relevance for WP5) in the context of statistical bioinformatics, which have lead to new biological and medical insights.
- *Copulas*
  The working group on 'Modeling dependencies and inference based on copulas' consisting of members of KUL-1 and UH has already led to several joint publications of members of these two universities. In 2010 the IAP-postdoc F. Abegaz (jointly hired by KUL-1 and UH) has been very active in this working group.

- *Sensitivity analysis*
  An incomplete data and sensitivity analysis working group with researchers from KUL-2, UH and EMC meets regularly and is hosted in alternation by UH and KUL-2. There are also contributions from LSHTM.

- *Goodness-of-fit tests*
  C. Heuchenne and I. Van Keilegom (UCL) have extensive collaborations with members of the USC partner (W. González-Manteiga and R. Crujeiras Casais) on goodness-of-fit tests in (semi)-parametric regression, when the data are or are not subject to right censoring. They also work together on a project dealing with ROC-curves in regression, and a project on inference for censored spatial data. The research is mostly situated in WP1, but also in WP2 and WP3. Recently, they started working on a project on semiparametric cure models in survival analysis.

2.4.3 Seminars

Each of the participating partners organizes on a regular basis statistics seminars at their universities. Announcements of these seminars are sent out to most Belgian statisticians, including those participating in the network.

Apart from the regular statistics seminars at the universities involved, several seminars have been organized by the network itself, around central themes of the network. They are on some occasions given by members of the network, in order to foster research interactions and exchange of ideas. These seminars are indicated by a star (*).

- 8 January, 2010: Yi Li (Harvard University), 'The Dantzig selector for censored linear regression models: identifying predictive genes for myeloma disease progression', at UG

- *January 22, 2010: Auguste Gaddah (UH), 'Flexible modeling in the generalized conditional Koziol-Green model by a copula function', at UCL

- *February 11, 2010: Jan Johannes, (UCL), 'Adaptive circular deconvolution by model selection under unknown error distribution, at KUL-1

- *March 18, 2010: Auguste Gaddah (UH), 'Flexible modeling in the generalized conditional Koziol-Green model by a copula function', at KUL-1

- *March 18, 2010: Fentaw Abegaz (postdoc IAP, joint at KUL-1 and UH), 'Semiparametric estimation of conditional copulas', at KUL-1. This seminar and the previous one were organized together around the common theme 'modeling dependencies in complex data using copulas'.

- May 21, 2010: Uschi Müller-Harknett (Texas A&M University),"Nonlinear regression with missing responses", at UCL

- May 21, 2010: Melanie Birke (University of Bochum), "Testing for monotonicity - an empirical process approach", at UCL

- November 11, 2010: Arne Kovac (Bristol University), 'Penalized regression on a graph', at UCL

- November 11, 2010: Juan Carlos Pardo Fernandez (University of Vigo), 'Nonparametric tests for risk-return relationship', at UCL

In addition, we like to mention that Anastasios Tsiatis (North Carolina State University) has been awarded the 2010-2011 Princess Lilian Visiting Professorship. The promoters (hosts) are Geert Molenberghs (UH and KUL-2), Geert Verbeke (KUL-2) and Marc Aerts (UH). During his visit, which is planned in May-June 2011, Anastasios Tsiatis will teach a doctoral-training course, he will provide guidance to the doctoral students of KUL-2 and UH, and he will give a sequence of research seminars at UH, KUL-2 and UCL.

Finally, Anthony Davison (EPFL, Lausanne), who is member of the follow-up committee of our IAP network, has been awarded the Francqui Chair 2011. He will give a series of lectures on 'Likelihood theory' and 'Statistics of extremes'. The host university is the University of Hasselt (UH). The lectures take place at UH, KUL-1 and UCL. There are 10 lectures in total between March 28 and May 5, 2011.

2.5 Short courses

Several short (intensive) courses have been organized. These courses were intended for all members of the network, and in particular (but not exclusively) for the PhD-students. The announcements were each time sent out to all members and posted on the web site. No (or reduced) registration fees were required for IAP-members.

A list of the short courses organized during the working year 2010 is given below.

- January 25-26, 2010: Short course on 'An introduction to statistical inverse problems', by Axel Munk (University of Goettingen, Germany), at UCL, jointly organized by the FNRS Graduate School in Statistics and the IAP network.

- February 17, 2010: Short course on 'Bayesian data analysis', by Andrew Gelman (Columbia University, New York, USA), at KUL-1. More than 180 participants registered for this course, including many PhD students and postdocs from all IAP partners. This activity can also be considered to have contributed significantly to the visibility of our network.

- June 1 and 3, 2010: Short course on 'Regression estimation under shape constraints - an overview', by Melanie Birke (Ruhr Universität Bochum, Germany), at UCL.

- October 18, 21 and 25, 2010: Short course on 'Long memory processes: theory and applications', by Francois Roueff (Telecom ParisTech), at UCL.

- November 24 and 25, 2010: Short course on 'Nonparametric methods in ROC curves' by Juan Carlos Pardo Fernández (University of Vigo), at UCL.

- December 20-21, 2010: Short course on 'The craft of smoothing' by Paul Eilers (EMC) and Brian Marx, at EMC.

2.6 PhD and postdoctoral researchers

The postdocs financed by the IAP network have been hired after a joint call, collecting all the positions offered by groups of the network in one single announcement, that was sent out on an international scale to many universities in the whole scientific community. Among those hired, we like to mention in particular the following two postdocs :

- Dan Lin is an IAP postdoc, who is for 50% financed by UH, and the other 50% by KUL-2 (contract from January 1 till December 31, 2010).
- Fentaw Abegaz is an IAP postdoc, who worked for 6 months at KUL-1 (October 2009-April 2010) followed by 6 months at UH (April 2010 - October 2010).
- Auguste Gaddah finished his PhD in September 2010 at UH, and started working as a postdoc at UCL from October on.

These types of postdoc positions are very interesting for the network, as it stimulates interactions among the different groups.

Several members of the network are members of PhD guidance committees of PhD students at other universities in the network. A list of PhD theses currently in preparation in the network can be found on the website:

`http://www.stat.ucl.ac.be/IAP/PhaseVI/research_theses.html`

The list mentions (among others) which members of partner universities of the network take part in the PhD committee. This participation is a very useful way to get familiar with the research carried out at other groups of the network. The website also contains a list of defended theses in the network since 2007.

2.7 Prizes obtained by network members

- Geert Molenberghs (UH and KUL-2): Drug Information Journal's Donald E. Francke Award for Overall Excellence in Journal Publishing, for article: Molenberghs (2009). Incomplete Data in Clinical Studies: Analysis, Sensitivity, and Sensitivity Analysis. *Drug Information Journal*, **43**, 409–446.
- Geert Molenberghs (UH and KUL-2): Award for Outstanding Contribution to the Development of the International Biometric Society. Awarded during the XXV International Biometric Conference, Florianópolis, Brazil, 07/12/2010.
- Geert Molenberghs (UH and KUL-2): Editor of Biostatistics, 2010-...
- Geert Verbeke (KUL-2): Editor of Biometrics, 2010-2012
- Geert Molenberghs (UH and KUL-2), Geert Verbeke (KUL-2) and Marc Aerts (UH) are promotors of the 2010-2011 Princess Lilian Foundation visiting professorship for Prof. Anastasios Tsiatis (University of North Carolina, U.S.A.).

- Geert Molenberghs (UH and KUL-2) and Geert Verbeke (KUL-2) received accredited as professional statistician by the American Statistical Association (ASA), 2010-2016.

- Anneleen Verhasselt (KUL-1): ‘Award for young researcher’ for the research paper ‘Regularisation and P-splines in generalised linear models” (joint with I. Gijbels), Journal of Nonparametric Statistics, 22, 271–295 (Journal of Nonparametric Statistics Student Paper Award 2010).

3 Technical Reports and Publications

Below we provide the scientific output related to the IAP-statistics network. We give both the technical reports and the publications of network members in 2010 :

- Technical Reports: These are manuscripts that have been written in 2010, and have been submitted for publication to an international journal. The reports are also available on our web site:

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_tr.html`.

  Each Technical Report has a number of the form TR10xxx, and we mention these reference numbers below. The web site also contains the pdf-file of many of the Technical Reports.

- Refereed publications: We list all published papers in international journals in 2010 (with refereeing system). We make the distinction between published papers and papers in press. See also the IAP-Statistics Reprints Series on our web site:

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_reprint.html`,

  for the published papers (reference numbers are of the form R10xxx). The papers in press have a label of the form RP10xxx.

- Non-refereed publications: We also include (an incomplete list of) papers that have been published without undergoing a peer review. The reference numbers are of the form NR10xxx (for the published ones) and NRP10xxx (for the ones in press).

- Books: These are books written by members of the network, that are published by international editors. They can also be found on the webpage

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_books.html`

  (reference numbers are of the form B10xxx and BP10xxx).

Below we list the research output of the IAP-network for each of the categories described above. We start with separate lists for each partner in the network, followed by a list of the technical reports and publications that are co-signed by researchers from at least two different groups from the network.

3.1 Université catholique de Louvain, UCL

3.1.1 Technical reports

[1] Autin, F., Freyermuth, J.-M. and R. von Sachs, Ideal denoising within a family of tree-structured wavelet estimators, 2011. TR11002.

[2] Badin, L., Daraio, C. and L. Simar, How to measure the impact of environmental factors in a nonparametric production model?, 2010. TR10054.

[3] Basrak, B., Krizmanic, D. and J. Segers, A functional limit theorem for partial sums of dependent random variables with infinite variance, 2010. TR10001.

[4] Breuning, C. and J. Johannes, Minimax-optimal local nonparametric instrumental regression, 2010. TR10007.

[5] Christiansen, M. and M. Denuit, First-order mortality rates and safe-side actuarial calculations in life insurance, 2010. TR10012.

[6] Christiansen, M., Denuit, M. and D. Lazar, The Solvency II square-root formula for systematic biometric risk, 2010. TR10038.

[7] Dahlke, M., Jay Breidt, F., Opsomer, J. and I. Van Keilegom, Nonparametric endogenous post-stratification estimation, 2011. TR11004.

[8] Daraio, C., Simar, L. and P.W. Wilson, Testing whether two-stage estimation is meaningful in non-parametric models of production, 2010. TR10035.

[9] Davydov, Y. and S. Liu, Estimation of parameters of regularly varying distributions on convex cones, 2010. TR10036.

[10] Davydov, Y. and S. Liu, Transformations of multivariate regularly varying tail distributions, 2010. TR10011.

[11] Denuit, M., Positive dependence of signals, 2010. TR10028.

[12] Denuit, M. and J. Dhaene, Convex order and comonotonic conditional mean risk sharing, 2010. TR10047.

[13] Denuit, M. and L. Eeckhoudt, A general index of absolute risk attitude, 2010. TR10016.

[14] Denuit, M. and L. Eeckhoudt, Bivariate stochastic dominance and common preferences of decision-makers with risk independent utilities, 2010. TR10021.

[15] Denuit, M. and L. Eeckhoudt, Stronger measures of higher-order risk attitudes: an extension, 2010. TR10051.

[16] Denuit, M. and M. Mesfioui, Dispersive effect of cross-aging with Archimedean copulas, 2010. TR10048.

[17] Denuit, M. and M. Mesfioui, Ordering functions of random vectors, with application to partial sums, 2010. TR10053.

[18] Denuit, M., Eeckhoudt, L. and O. Jokung, Transformations preserving stochastic dominance: Theory and applications, 2010. TR10027.

[19] Denuit, M., Eeckhoudt, L. and M. Menegatti, Correlated risks, bivariate utility and optimal choices, 2010. TR10018.

[20] Denuit, M., Eeckhoudt, L. and M. Menegatti, Adding independent risks in an insurance portfolio: Which shape for the insurer's preferences?, 2010. TR10023.

[21] Denuit, M., Eeckhoudt, L., Tsetlin, I. and R.L. Winkler, Multivariate concave and convex stochastic dominance, 2010. TR10039.

[22] Denuit, M., Haberman, S. and A. Renshaw, Longevity-indexed life annuities, 2010. TR10026.

[23] Denuit, M. and B. Rey, Transformations of multivariate regularly varying tail distributions, 2010. TR10024.

[24] Devolder, P. and F. Dominguez, NDC Dynamic Equilibrium model with financial and demographic risks, 2010. TR10041.

[25] Eeckhoudt, L. and M. Denuit, Stronger measures of higher-order risk attitudes, 2010. TR10013.

[26] Ferraty, F., Van Keilegom, I. and P. Vieu, Bootstrap and inference when both response and regressor are functional, 2010. TR10043.

[27] Guillotte, S., Perron, F. and J. Segers, Nonparametric Bayesian inference on bivariate extremes, 2010. TR10049.

[28] Hafner, C., Laurent, S. and F. Violante, The diffusion limit of dynamic conditional correlation models, 2010. TR10004.

[29] Hafner, C. and O. Reznikova, On the estimation of dynamic conditional correlation models, 2010. TR10005.

[30] Heuchenne, C. and I. Van Keilegom, Estimation of a general parametric location in censored regression, 2010. TR10010.

[31] Hunt, J. and M. Hahn, Estimation and calibration of a continuous-time semi-Markov switching model, 2010. TR10052.

[32] Hunt, J. and P. Devolder, Minimal entropy martingale measure in a semi Markov regime switching CIR model, 2010. TR10042.

[33] Jaeger, J. and P. Lambert, Bayesian generalized profiling estimation in hierarchical linear dynamic systems, 2011. TR11001.

[34] Johannes, J. and R. Schenk, On rate optimal local estimation in functional linear model, 2010. TR10006.

[35] Johannes, J. and M. Schwarz, Adaptive estimation in nonparametric instrumental regression, 2010. TR10008.

[36] Johannes, J. and M. Schwarz, Adaptive nonparametric instrumental regression by model selection, 2010. TR10029.

[37] Kneip, A., Simar, L. and I. Van Keilegom, Boundary estimation in the presence of measurement error with unknown variance, 2010. TR10050.

[38] Lambert, P., Additive location-scale models for interval censored data, 2010. TR10055.

[39] Meinguet, T. and J. Segers, Regularly varying time series in Banach spaces, 2010. TR10002.

[40] Meira-Machado, L., Roca-Pardinas, J., Van Keilegom, I. and C. Cadarso-Suarez, Estimation of transition probabilities in a non-Markov model with successive survival times, 2010. TR10057.

[41] Meinguet, T., Maxima of moving maxima of continuous functions, 2010. TR10030.

[42] Park, B.U., Simar, L. and V. Zelenuyk, Local maximum likelihood techniques with categorical data, 2010. TR10056.

[43] Pigeon, M. and M. Denuit, Composite Lognormal-Pareto model with random threshold, 2010. TR10017.

[44] Russo, F., Wunsch, G. and M. Mouchart, Inferring causality through counterfactuals in observational studies some epistemological issues, 2010. TR10033.

[45] Schaffer, A., Simar, L. and J. Rauland, Decomposing regional efficiency, 2010. TR10046.

[46] Segers, J.,Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions, 2010. TR10058.

[47] Simar, L. and A. Vanhems, Probabilistic characterization of directional distances and their robust versions, 2010. TR10044.

[48] Simar, L. and P. W. Wilson, Two-Stage DEA: Caveat Emptor, 2010. TR10045.

[49] Trufin, J., Albrecher, H. and M. Denuit, Properties of risk measures derived from ruin theory, 2010. TR10022.

[50] Trufin, J., Albrecher, H. and M. Denuit, Ruin problems in presence of underwriting cycles, 2010. TR10019.

[51] Trufin, J., Albrecher, H. and M. Denuit, Ruin problems under IBNR dynamics, 2010. TR10025.

[52] Van Keilegom, I. and N. Veraverbeke, Statistical models and methods for dependence in insurance data, 2011. TR11003.

[53] von Sachs, R. and C. Timmermans, BAGIDIS, a new method for statistical analysis of differences between curves with sharp discontinuities, 2010. TR10034.

[54] Wei, J., Carroll, R., Mller, U., Van Keilegom, I. and N. Chatterjee, Robust estimation for homoscedastic regression in the secondary analysis of case-control data, 2010. TR10037.

3.1.2 Refereed publications (published)

[55] Badin, M., Daraio, C. and L. Simar, Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. *European Journal of Operational Research*, **201**, 633-640, 2010. R10002.

[56] Bloemberg, T. G., Gerretzen, J., Wouters, H.J.P., Gloerich, J., van Dael, M., Wessels, H.J.C.T., van den Heuvel, L.P., Eilers, P.H.C., Buydens, L.M.C. and R. Wehrens, Improved parametric time warping for proteomics. *Chemometrics And Intelligent Laboratory Systems*, **104**, 65-74, 2010. R10245.

[57] Bhm, H., Ombao, H., Von Sachs, R. and J. Sanes, Classification of multivariate non-stationary signals: the SLEX-shrinkage approach. *Journal of Statistical Planning and Inference*, **140**, 3754-3763, 2010. R10052.

[58] Camargo, C.A., Gurner, D., Smithline, H., Chapela, R., Fabbri, L., Green, S., Malice, M.P., Legrand, C., Dass, S.B., Knorr, B. and T. Reiss, A randomized placebo-controlled study of intravenous montelukast for the treatment of acute asthma. *Journal of Allergy and Clinical Immunology*, **125(2)**, 374-38, 2010. R10037.

[59] Cardot, H. and J. Johannes, Thresholding projection estimators in functional linear models. *Journal of Multivariate Analysis*, **101(2)**, 395-408, 2010. R10050.

[60] Cetinyurek, A. and P. Lambert, Smooth estimation of survival functions and hazard ratios from interval-censored data using Bayesian penalized B-splines. *Statistics in Medicine*, **30**, 75-90, 2011. R11004.

[61] Comte, F. and J. Johannes, Adaptive estimation in circular functional linear models. *Mathematical Methods of Statistics*, **19**, 42-63, 2010. R10051.

[62] Crujeiras, R. and I. Van Keilegom, Least squares estimation of nonlinear spatial trends. *Computational Statistics and Data Analysis*, **54**, 452-465, 2010. R10004.

[63] Daouia, A., Florens, J.P. and L. Simar, Frontier estimation and extreme values theory. *Bernoulli*, **16(4)**, 1039-1063, 2010. R10058.

[64] Daskovska, A., Simar, L. and S. Van Bellegem, Forecasting the Malmquist productivity index. *Journal of Productivity Analysis*, **33**, 97-107, 2010. R10026.

[65] Degen, M., Lambrigger, D. and J. Segers, Risk concentration and diversification: Second-order properties. *Insurance: Mathematics and Economics*, **46**, 541-546, 2010. R10025.

[66] Denuit, M., Eeckhoudt, L. and B. Rey, Some consequences of correlation aversion in decision science. *Ann Oper Res*, **176**, 259-269, 2010. R10024.

[67] Denuit, M. and L. Eeckhoudt, A general index of absolute risk attitude. *Management Science*, **56**, 712-715, 2010. R10042.

[68] Denuit, M., Haberman, S. and A. Renshaw, Comonotonic approximations to quantiles of life annuity conditional expected present values: extensions to general ARIMA models and comparison with the bootstrap. *ASTIN Bulletin*, **40**, 331-349, 2010. R10044.

[69] Denuit, M. and M. Mesfioui, Generalized increasing convex and directionally convex orders. *Journal of Applied Probability*, **47**, 264-276, 2010. R10045.

[70] Deprins, D., "Probabilit et Incertitude" Responsabilit et environnement: recherches, dbats, actions, Faire face l'incertitude. *Annales des Mines (Ecole des Mines de Paris), Editions Eska*, **57**, 22-31, 2010. R10046.

[71] Devolder, P., Perspectives pour nos rgimes de pension lgale. *Revue Belge de Scurit Sociale*, **1**. R10047.

[72] El Ghouch, A., Van Keilegom, I. and I. W.McKeague, Empirical likelihood confidence intervals for dependent duration data. *Econometric Theory*, **27**, 178-198, 2011. R11006.

[73] Ferraty, F., Van Keilegom, I. and P. Vieu, On the validity of the bootstrap in non-parametric functional regression. *Scandinavian Journal of Statistics*, **37**, 286-306, 2010. R10029.

[74] Forget, P., Vandenhende, J., Berliere, M., Machiels, J.P., Nussbaum, B., Legrand, C. and M. De Kock, Do intraoperative analgesics influence breast cancer recurrence after mastectomy? A retrospective analysis. *Anesthesia Analgesia*, **110(6)**, 1630-1635, 2010. R10038.

[75] Freyermuth, J-M., Ombao, H. and R. von Sachs, Tree-structured wavelet estimation in a mixed effects model for spectra of replicated time series. *Journal of the American Statistical Association*, **105**, 490, 634-646, 2010. R10036.

[76] Gaddah, A. and R. Braekers, Testing under the extended Koziol-Green model. Durante, F., Hardle, W., Jaworski, P., Rychlik, T. (Eds.) Workshop on Copula Theory and its applications, Lectures Notes in Statistics 198. *Proceedings Springer, Berlin/Heidelberg*, 2010. R10195.

[77] Gaddah, A. and R. Braekers, Testing under the Koziol-Green model, Copula theory and its applications. Proceedings of the Workshop held in Warsaw 25-26 September 2009, ed. Piotr Jaworski, Fabrizio Durante, Wolfgang Karl Hrdle and Tomasz Rychlik. *Lecture Notes in Statistics*, **198**, 2, 279-288, Springer, 2010. R10202.

[78] Geenens, G. and L. Simar, Nonparametric test for conditional independence in two-way contingency tables. *Journal of Multivariate Analysis*, **101**, 765-788, 2010. R10030.

[79] Genest, C. and J. Segers, On the covariance of the aymptotic empirical copula process. *Journal of Multivariate Analysis*, **101**, 1837-1845, 2010. R10054.

[80] Gudendorf, G. and J. Segers, Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *Journal of Multivariate Analysis*, **102**, 37-47, 2011. R11003.

[81] Hafner, C. and A. Preminger, Deciding between GARCH and stochastic volatility via strong decision rules. *Journal of Statistical Planning and Inference*, **140**, 790-805, 2010. R10008.

[82] Hafner, C. and O. Reznikova, Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics and Data Analysis*, **54**, 2609-2627, 2010. R10049.

[83] Heuchenne, C. and I. Van Keilegom, Estimation in nonparametric location-scale regression models with censored data. *The Annals of the Institute of Statistical Mathematics*, **62**, 439-463, 2010. R10028.

[84] Hollevoet, K., Nackaerts, K., Thimpont, J., Germonpre, P., Bosquee, L., De Vuyst, P., Legrand, C., Kellen, E., Kishi, Y., Delanghe, J.R. and J.P. van Meerbeeck, Diagnostic performance of soluble mesothelin and megakaryocyte potentiating factor in mesothelioma. *American Journal of Respiratory and Critical Care Medicine*, **181**, 620-625, 2010. R10039.

[85] Jeong, S.O., Park, B.U. and L. Simar, Nonparametric conditional efficiency measures: asymptotic properties. *Annals of Operations Research*, **173**, 105-122, 2010. R10011.

[86] Lambert P., Smooth and semi- and nonparametric Bayesian estimation of bivariate densities from bivariate histogram data. *Computational Statistics and Data Analysis*, **55**, 429-445, 2011. R11005.

[87] Laurent, S., Further comments on the representation problem for stationary processes. *Statistics and Probability Letters*, **80**, 592-596, 2010. R10013.

[88] Laurent, S., On standardness and I-cosiness. *Sminaire de Probabilits XLIII, Lecture Notes in Mathematics*, **121**, 2010. R10055.

[89] Laurent, S., On Vershikian and I-cosy random variables. *Teoriya Veroyatnostei i ee Primeneniya*, **55**, 104-132, 2010. R10056.

[90] Linton, O., Mammen, E., Perch Nielsen, J. and I. Van Keilegom, Nonparametric regression with filtered data. *Bernoulli*, **17**, 1, 60-87, 2011. R11008.

[91] Manner, H. and J. Segers, Tails of correlation mixtures of elliptical copulas - Insurance: Mathematics and Economics. *Journal of Statistical Planning and Inference*, **48**, 153-160, 2011. R11002.

[92] Marx, B.D., Eilers, P.H.C., Gampe, J. and R. Rau, Penalized regression with individual deviance effects. *Statistics and Computing*, **20**, 191-202, 2010. R10031.

[93] Molanez-Lopez, E.M., Cao, R. and I. Van Keilegom, Smoothed empirical likelihood confidence intervals for the relative distribution with left-truncated and right-censored data. *The Canadian Journal of Statistics*, **38**, 3, 453-473, 2010. R10060.

[94] Motta, G., Hafner, C. and R. von Sachs, Locally stationary factor models: identification and nonparametric estimation. *Econometric Theory*, **27**, 6 , 2011. R11007.

[95] Mouchart, M. and M. Vandresse, A double-frontier approach for measuring market imperfection. *Special Issue of Annals of Operations Research on Efficiency and Productivity*, **173**, 137-144, 2010. R10057.

[96] Omey, E. and J. Segers, Generalised regular variation of arbitrary order. *Banach Center Publications*, **90**, 111-137, 2010. R10059.

[97] Park, B., Jeong, S.-O. and L. Simar, Asymptotic distribution of conical-hull estimators of directional edges. *Annals of Statistics*, **38(6)**, 1320-1340, 2010. R10041.

[98] Roueff, F. and R. von Sachs, Locally stationary long memory estimation. *Stochastic Processes and their Applications*, **121**, 813-844, 2011. R11009.

[99] Russo, F., Mouchart, M. and G. Wunsch, Causalit nelle scienze sociali: una sfida o una chimera?. *Sis-Magazine*, 2010. R10053.

[100] Schwarz, M. and S. van Bellegem, Consistent density deconvolution under partially known error distribution. *Statistics and Probability Letters*, **80**, 236-241, 2010. R10009.

[101] Simar, L. and P.W. Wilson, Inference from cross-sectional stochastic frontier models. *Econometric Review*, **29**, 1, 62-98, 2010. R10012.

[102] Surmont, V., Aerts, J.G.J.V., van Klaveren, R.J., Tournoy, K., Tan, K.Y., Vernhout, R.M., Schmitz, P.I.M, Legrand, C., Hoogesteden, H.C. and J.P.van Meerbeeck, A randomized phase II study comparing two schedules of the 21-day regimen of Gemcitabine and carboplatin in advanced Non-Small Cell Lung Cancer. *Oncology*, **78**, 267-270, 2010. R10040.

[103] Teodorescu, B., Cao, R. and I. Van Keilegom, Generalized time-dependent conditional linear models under left truncation and right censoring. *The Annals of the Institute of Statistical Mathematics*, **62**, 465-485, 2010. R10027.

[104] Teodorescu, B. and I. Van Keilegom, Goodness-of-fit test for generalized conditional linear models under left truncation and right censoring. *Journal of Nonparametric Statistics*, **22**, 547-566, 2010. R10035.

[105] Van De Putte, G., Vlasselaer, J., Teshome, B., Gaddah, A., Burzykowski, T., Schobbens, J.C., Benijts, G. and E.T.M. De Jonge, Outcome of per protocol best-evidence based routine breast cancer care in a large regional hospital in Belgium: the importance of a prospective database in quality assurance. *Facts, Views and Vision in Obstetrics and Gynaecology*, **2**, 119-124, 2010. R10163.

[106] van Eijk, R., Eilers, P.H., Natt, R., Cleton-Jansen, A.M., Morreau, H., van Wezel, T. and J. Oosting, MLPAinter for MLPA interpretation: an integrated approach for the analysis, visualisation and data management of Multiplex Ligation-dependent Probe Amplification. *BMC bioinformatics*, **11**, 67, 2010. R10032.

[107] Van Keilegom, I. and L. Wang, Semiparametric modeling and estimation of heteroscedasticity in regression analysis of cross-sectional data. *Electronic Journal of Statistics*, **4**, 133-160, 2010. R10001.

[108] Van Keilegom, I., De Ua-lvarez, J. and L. Meira-Machado, Nonparametric location-scale models for censored successive survival times. *Journal of Statistical Planning and Inference*, **141**, 1118-1131, 2011. R11001.

[109] Wunsch, G., Russo, F. and M. Mouchart, Do we necessarily need longitudinal data to infer causal relations? *BMS: Bulletin of Sociological Methodology*, **106**, 5-18, 2010. R10033.

3.1.3 Refereed publications (in press)

[110] Böhm, H., Ombao, H., von Sachs, R. and J. Sanes, Classification of multivariate non-stationary signals : The SLEX-shrinkage approach. *Journal of Statistical Planning and Inference*, 2010. RP10078.

[111] Florens, J.P., Johannes, J. and S. Van Bellegem, Identification and estimation by penalization in nonparametric instrumental regression. *Econometric Theory*, 2010. RP10063.

[112] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, 2010. RP10094.

[113] Geenens, G. and L. Simar, Single-index modelling of conditional probabilities in two-way contingency tables. *Statistics*, 2010. RP10077.

[114] González Manteiga, W., Pardo Fernández, J.C. and I. Van Keilegom, ROC curves in nonparametric location-scale regression models. *Scandinavian Journal of Statistics*, 2010. RP10009.

[115] Gudendorf, G. and J. Segers, Extreme-value copulas. *Workshop on Copula Theory and its Applications.* F. Durante, W. Haerdle, P. Jaworski and T. Rychlik, (eds.), Lecture Notes in Statistics - Proceedings. Springer, 2010. RP10081.

[116] Heuchenne, C. and I. Van Keilegom, Goodness-of-fit tests for the error distribution in nonparametric regression. *Computational Statistics and Data Analysis*, 2010. RP10010.

[117] Johannes, J. and S. Subba Rao, Nonparametric estimation for dependent data. *Journal of Nonparametric Statistics*, 2010. RP10065.

[118] Johannes, J., Van Bellegem, S. and A. Vanhems, Convergence rates for ill-posed inverse problems with an unknown operator. *Econometric Theory*, 2010. RP10064.

[119] Kneip, A., Simar, L. and P.W. Wilson, A Computational Efficient, Consistent Bootstrap for Inference with Non-parametric DEA Estimators. *Computational Economics*, 2011. RP11007.

[120] Mouchart, M. and F. Russo, Causal explanation: recursive decompositions and mechanisms. *Causality in the sciences*, McKay Illari, P., Russo, F. and J. Williamson (eds.), Oxford University Press, 2010. RP10071.

[121] Neumeyer, N. and I. Van Keilegom, Estimating the error distribution in nonparametric multiple regression with applications to model testing. *Journal of Multivariate Analysis*, 2010. RP10011.

[122] Sanchez, X., Lambert, P. and D.J. Llewellyn, Efficacy of pre-ascent climbing route visual inspection in indoor sport climbing. *Scandinavian Journal of Medicine and Science in Sport*, 2010. RP10141.

[123] Innovation and Export Activities in the German Mechanical Engineering Sector: An Application of Testing Restrictions in Production Analysis. *Journal of Productivity Analysis*, 2011. RP11009.

[124] Schwarz, M., Van Bellegem, S. and J.P. Florens, Nonparametric frontier estimation from noisy data. *Feestchrift in honour of Léopold Simar*, 2010. RP10125.

[125] Simar, L. and P.W. Wilson, Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*, 2011. RP11008.

[126] Simar, L. and V. Zelenyuk, Stochastic FDH/DEA estimators for frontier analysis. *Journal of Productivity Analysis*, 2010. RP10081.

[127] Varron, D. and I. Van Keilegom, Uniform in bandwidth exact rates for a class of kernel estimators. *Annals of the Institute of Statistical Mathematics*, 2010. RP10008.

3.1.4 Non-refereed publications (in press)

[128] Hafner, C. and H. Manner, Multivariate time series models for asset prices. *Handbook of Computational Finance*, Springer, 2010. NRP10002.

[129] Hafner, C., GARCH modelling, In: *Encyclopedia of Complexity and System Science*. R.A. Meyers (ed.), Springer, 2010. NRP10001.

3.2 Katholieke Universiteit Leuven, KUL-1

3.2.1 Technical reports

[130] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in additive models using P-splines, 2010. TR10064.

[131] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11007.

[132] Casas, I. and I. Gijbels, Unstable volatility: the break preserving local linear estimator, 2010. TR10065.

[133] Claeskens, G., Hubert, M. and L. Slaets, Phase and amplitude-based clustering for functional data, 2010. TR10066.

[134] Croux, C., Fried, R., Gijbels, I. and K. Mahieu, Robust forecasting of non-stationary time series, 2010. TR10067.

[135] Croux, C., Gijbels, I. and I. Prosdocimi, Robust estimation of mean and dispersion functions in extended generalized additive models., 2010. TR10068.

[136] Desmet, L. and I. Gijbels, Improved spectral density estimation at peaks. TR10069.

[137] Helsen, W.F., De Neve, J. D'Hooghe, P., Gijbels, I., Gram, M., Hubert, M., Peers K., Probst, S. and J. Van Winckel, A Ménage à Trois: Injury Prevention, Medical Attention and Time Loss Injuries, 2010. TR10070.

[138] Hubert, M., Dierckx, G. and D. Vanpaemel, Detecting influential data points in Pareto-type distributions, 2010. TR10071.

[139] Omelka, M., Veraverbeke, N. and I. Gijbels, Bootstrapping the conditional copula, 2010. TR10072.

3.2.2 Refereed publications (published)

[140] Beirlant J. and H.H.J. Einmahl, Asymptotics for the Hirsch Index. *Scandinavian Journal of Statistics*, **37**, 355-364, 2010. R10280.

[141] Beirlant J., Guillou A. and G. Toulemonde, Peaks-Over-Threshold modeling under random censoring. *Communications in Statistics - Theory and Methods*, **39**, 1158-1179, 2010. R10281.

[142] Ceulemans, E. and G. Storms, Detecting intra and inter categorical structure in semantic concepts using HICLAS. *Acta Psychologica*, **133**, 296-304, 2010. R10072.

[143] Debruyne, M., Christmann, A., Hubert, M. and J. Suykens, Robustness of reweighted least squares kernel based regression. *Journal of Multivariate Analysis*, **101** (2), 447-463, 2010. R10281.

[144] Debruyne, M., Hubert, M. and J. Van Horebeek, Detecting influential observations in kernel PCA. *Computational Statistics & Data Analysis*, **54** (12), 3007-3019, 2010. R10282.

[145] Desmet, L., Gijbels, I. and A. Lambert, Estimation of irregular probability densities. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková*. IMS Collections, **7**, 46-61, 2010. R10284.

[146] Fang, F., Jonsson, H., Schoutens, W. and C. Oosterlee, Fast valuation and calibration of credit default swaps under Lévy dynamics. *Journal of Computational Finance*, **14**, Number 2, 2010. R10283.

[147] Frederickx, S., Tuerlinckx, F., De Boeck, P. and D. Magis, RIM: A random item mixture model to detect Differential Item Functioning. *Journal of Educational Measurement*, **47**, 432-457, 2010. R10062.

[148] Gelman, A., Leenen, I., Van Mechelen, I., De Boeck, P. and J. Poblome, Bridges between deterministic and probabilistic models for binary data. *Statistical Methodology*, **7**, 187-209, 2010. R10063.

[149] Gijbels, I., Censored data. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**, 2, 178-188., 2010. R10285.

[150] Gijbels, I. and A. Verhasselt, Regularisation and P-splines in generalised linear models. *Journal of Nonparametric Statistics*, **22**, Issue 3, 271-295, 2010. R10286.

[151] Gijbels, I. and A. Verhasselt, P-splines regression smoothing and difference type of penalty. *Statistics and Computing*, **20**, 499-511, 2010. R10287.

[152] Gijbels, I. and I. Prosdocimi, Loess. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**, 590-599, 2010. R10288.

[153] Gijbels, I., Omelka, M. and D. Sznajder, Positive quadrant dependence tests for copulas. *The Canadian Journal of Statistics*, **38**, Issue 4, 555581, 2010. R10289.

[154] Gijbels, I., Prosdocimi, I. and G. Claeskens, Nonparametric estimation of mean and dispersion functions in extended generalized linear models. *Test*, **19**, Number 3, 580-608, 2010. R10290.

[155] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics & Data Analysis*, **55**, 1919-1932, 2010. R11010.

[156] Goegebeur, Y., De Boeck, P. and G. Molenberghs, Person fit for test speededness: normal curvatures, likelihood ratio tests and empirical Bayes estimates. *Methodology: European journal of research methods for the behavioral and social sciences*, **6**, 3-16, 2010. R10064.

[157] Goegebeur, Y., Beirlant J. and T. De Wet, Generalized kernel estimators for the Weibull-Tail Coefficient. *Communications in Statistics - Theory and Methods*, **39**, 3695-3716., 2010. R10292.

[158] Goegebeur Y., Beirlant J. and T. de Wet, Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, **140**, 2632-2652, 2010. R10293.

[159] Hubert, M. and M. Debruyne, Minimum Covariance Determinant. *WIREs Computational Statistics*, **2**, 36-43, 2010. R10294.

[160] Hubert, M. and S. Van der Veeken, Robust classification for skewed data. *Advances in Data Analysis and Classification*, **4**, 239-254, 2010. R10295.

[161] Jönsson, H., and Schoutens, W. (2010). Known and less known risks in asset-backed securities. In: Wigan D. (Eds.), *Credit Derivatives: The March to Maturity*, **9**, 113-137, 2010. R10296.

[162] Joossens, E. and W. Schoutens, Portfolio insurances, CPPI and CPDO, truth or illusion?. In: Kiesel R., Scherer M., Zagst R. (Eds.), *Alternative Investments and Strategies*, **11**, 259-294), 2010. R10297.

[163] Madan, D., and W. Schoutens, Simple Processes and the Pricing and Hedging of Cliquets. *Mathematical Finance*, 2010. R10298.

[164] Magis, D., Bland, S., Tuerlinckx, F. and P. De Boeck, A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods*, **42**, 847-862, 2010. R10065.

[165] Magis, D. and G. Rache, An iterative maximum a posteriori estimation of proficiency level to detect multiple local likelihood maxima. *Applied Psychological Measurement*, **34**, 75-90, 2010. R10066.

[166] Schepers, J. and I. Van Mechelen, Uniqueness of real-valued hierarchical classes models. *Journal of Mathematical Psychology*, **54**, 215-221, 2010. R10067.

[167] Schoutens, W. and S. Mangold, Single Name Credit Derivatives. In: Wigan D. (Eds.), *Credit Derivatives : The March to Maturity*, **8**, 95-112, 2010. R10299.

[168] Schoutens, W. and G. Van Damme, The Beta-Variance Gamma Process. *Review of Derivatives Research*, 2010. R10300.

[169] Timmerman, M.E. and E. Ceulemans, *The generic subspace clustering model* (Y. Lechevallier G. Saporta), Proceedings of COMPSTAT'2010. 19th International Conference on Computational Statistics (pp. 359-366). Heidelberg: Physica-Verlag, 2010. R10034.

[170] Timmerman, M.E., Ceulemans, E., Kiers, H.A.L. and M. Vichi, Factorial and Reduced K-means reconsidered. *Computational Statistics and Data Analysis*, **54**, 1858-1871, 2010. R10068.

[171] Tuerlinckx F., *Weibull distribution* (N. Salkind), Encyclopedia of Research Design. Thousand Oaks, CA: Sage, 2010. R10023.

[172] Vandekerckhove, J., Verheyen, S. and F. Tuerlinckx, A crossed random effects diffusion model for speeded semantic categorization decisions. *Acta Psychologica*, **133**, 269-282, 2010. R10069.

[173] van den Berg, R.A., Braaksma, M., van der Veen, D., van der Werf, M.J., Punt, P.J., van der Oost, J. and L.H. de Graaff, Identification of modules in Aspergillus niger by gene co-expression network analysis. *Fungal Genetics and Biology*, **47**, 539-550, 2010. R10070.

[174] Van Mechelen, I. and A.K. Smilde, A generic linked-mode decomposition model for data fusion. *Chemometrics and Intelligent Laboratory Systems*, **104**, 83-94, 2010. R10071.

[175] Van Mechelen, I. and K. Van Deun, *Multiple nested reductions of single data modes as a tool to deal with large data sets* (Y. Lechevallier G. Saporta), Proceedings of COMPSTAT'2010. 19th International Conference on Computational Statistics (pp. 349-358). Heidelberg: Physica-Verlag, 2010. R10010.

[176] Verboven, S. and M. Hubert, Matlab library LIBRA. *WIREs Computational Statistics*, **2** (4), 509-515, 2010. R10301.

3.2.3 Refereed publications (in press)

[177] Antoniadis, A., Gijbels, I. and M. Nikolova, Penalized likelihood regression for generalized linear models with nonquadratic penalties. *The Annals of the Institute of Statistical Mathematics*, RP11006.

[178] Ceulemans, E., Timmerman, M.E. and H.A.L. Kiers, The CHull procedure for selecting among multilevel component solutions. *Chemometrics and Intelligent Laboratory Systems*, 2010. RP10084.

[179] Claeskens, G. and M. Jansen, Nonparametric estimation. In: *International Encyclopedia of Statistical Science*, Springer, 2010. RP10014.

[180] Daouia, A. and I. Gijbels, Estimating frontier cost models using extremiles. *Festschrift in honor of Leopold Simar*, Editors: P. Wilson, I. Van Keilegom, 2010. RP10002.

[181] Daouia, A. and I. Gijbels, Robustness and inference in nonparametric partial-frontier modeling. *Journal of Econometrics*, 2011. RP11010.

[182] Desmet, L. and I. Gijbels, Curve fitting under jump and peak irregularities using local linear regression. *Communications in Statistics, Theory and Methods*, 2011. RP11011.

[183] Ding, H., Claeskens, G. and M. Jansen, Variable selection in partially linear wavelet models. *Statistical Modelling*, 2010. RP10015.

[184] Gijbels, I. and I. Prosdocimi, Smooth estimation of mean and dispersion function in extended Generalized Additive Models with application to Italian Induced Abortion data. *Journal of Applied Statistics*, 2011. RP11012.

[185] Gijbels, I., The Kaplan-Meier estimator. In: *International Encyclopedia of Statistical Science*, Springer, 2010. RP10020.

[186] Hubert, M. and S. Van der Veeken, Fast and robust classifiers adjusted for skewness. *Proceedings of Compstat*, 2010. RP10082.

[187] Jansen, M. and G. Claeskens, The Cramér-Rao inequality. In: *International Encyclopedia of Statistical Science*, Springer, 2010. RP10021.

[188] Lodewyckx, T., Tuerlinckx, F., Kuppens, P., Allen, N.B. and L. Sheeber, A hierarchical state space approach to affective dynamics. *Journal of Mathematical Psychology*, 2010. RP10085.

[189] Magis, D., Bland, S. and G. Rache, A test-length correction to the estimation of extreme proficiency levels. *Applied Psychological Measurement*, 2010. RP10086.

[190] Oravecz, Z. and F. Tuerlinckx, The linear mixed model and the hierarchical Ornstein-Uhlenbeck model: Some Equivalences and differences. *British Journal of Mathematical and Statistical Psychology*, 2010. RP10087.

[191] Vandekerckhove, J. and F. Tuerlinckx, MATLAB for behavioral scientists: A novice's guide to MATLAB. *Experimental Psychology*, 2010. RP10088.

[192] Vandekerckhove, J., Tuerlinckx, F. and M.D. Lee, Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 2010. RP10143.

[193] Van Mechelen, I., and A.K. Smilde, Comparability problems in the analysis of multiway data. *Chemometrics and Intelligent Laboratory Systems*, 2010. RP10090.

[194] Veraverbeke, N., Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scadinavian Journal of Statistics*, 2010. RP10093.

[195] Verdonck, T., Rousseeuw, P. and M. Hubert, DetMCD in a regression framework. *Proceedings of Compstat*, 2010. RP10083.

[196] Wilderjans, T.F., Ceulemans, E., Van Mechelen, I. and D. Depril, ADPROCLUS: A graphical user interface for fitting additive profile clustering models to object by variable data matrices. *Behavior Research Methods*, 2010. RP10091.

[197] Wilderjans, T.F., Ceulemans, E., Van Mechelen, I. and R.A. van den Berg, Simultaneous analysis of coupled data matrices subject to different amounts of noise. *British Journal of Mathematical and Statistical Psychology*, 2010. RP10092.

3.2.4 Books (in press)

[198] De Spiegeleer, J., and W. Schoutens, *The Handbook of Convertible Bonds: Pricing, Strategies and Risk Management*. Chichester: Wiley, 2010. BP11001.

3.3 Katholieke Universiteit Leuven, KUL-2

3.3.1 Technical reports

[199] Vannieuwenhuyze, J. and G. Molenberghs, A SAS macro to disentangle mode effects on proportions and the mean of a categorical variable in an extended mixed-mode dataset, 11 pp. Leuven: K.U.Leuven. Ceso, 2010. TR10065.

3.3.2 Refereed publications (published)

[200] Aerts, M., Faes, C., Hens, N. and G. Molenberghs, Handling missingness and informative cluster sizes when modelling prevalence and force of infection. *Joint Statistical Meetings 2010 Proceedings* 3491-3501, 2010. R10194.

[201] Alonso, A., Laenen, A., Molenberghs, G., Geys, H. and T. Vangeneugden, Reliability of Single Administered Tests: A Unified Approach. *Biometrics*, **66** (4), 1061-1068, 2010. R10236.

[202] Alonso, A., Litire, S. and A. Laenen, A note on the indeterminacy of the random-effects distribution in hierarchical models. *American Statistician*, **64** (4), 318-324, 2010. R10235.

[203] Alonso, A., Litire, S. and G. Molenberghs, Testing for misspecification in generalized linear mixed models. *Biostatistics*, **11**, 771-786, 2010. R10125.

[204] Ampe, M., Verbeke, G., Vanneste, E. and J. Vermeesch, Analysis of array CGH data for the detection of single-cell chromosomal imbalances. *Online Journal of Bioinformatics*, **11** (2), 224-244, 2010. R10231.

[205] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Using earlier measures in a longitudinal sequence as potential surrogate for a later one. *Computational Statistics and Data Analysis*, **54**, 1342-1354, 2010. R10137.

[206] Bruckers, L., Serroyen, J., Molenberghs, G., Slaets, H., and W. Goeyvaerts, Latent-class analysis of persistent disturbing behavior patients using longitudinal profiles. *Applied Statistics*, **59**, 495-512, 2010. R10119.

[207] Buysschaert, I., Carruthers, K., Dunbar, D., Peuteman, G., Rietzschel, E., Belmans, A., Hedley, A., De Meyer, T., Budaj, A., Van de Werf, F., Lambrechts, D. and K. Fox, A variant at chromosome 9p21 is associated with recurrent myocardial infarction and cardiac death after acute coronary syndrome: The GRACE Genetics Study. *European Heart Journal*, **31** (9), 1132-1141, 2010. R10219.

[208] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G., and M.G. Kenward, A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes. *Biometrical Journal*, **52**, 111-125, 2010. R10124.

[209] Creemers, A., Hens, N., Aerts, M. Molenberghs, G. Verbeke, G. and M.G. Kenward, A sensitivity analysis for shared-parameter models. *The Biometrical Journal*, **52**, 111-125, 2010. R10142.

[210] De Baets, E., Lambrechts, H., Lemiere, J., Diya, L. and G. Willems, Influence of self-esteem on quality of life in children seeking orthodontic treatment. *European Journal of Orthodontics*, **32**, 6, European Orthodontic Society Annual Meeting. Portoroz, Slovenia, e53, e53, Abstract No. 125, 2010. R10303.

[211] Debruyne, M., Christmann, A., Hubert, M. and J.A.K. Suykens, Robustness and stability of reweighted kernel based regression. *Journal of Multivariate Analysis*, **101**, 447-463, 2010. R10020.

[212] De Foer, B., Vercruysse, F., Bernaerts, A., Meersschaert, J., Kenis, C., Pouillon, M., De Beuckeleer, L., Michiels, J., Bogaerts, K., Deckers, F., Somers, T., Hermans, R., Offeciers, E. and J. Casselman, Middle Ear Cholesteatoma: Non-Echo-planar Diffusion-weighted MR Imaging versus Delayed Gadolinium-enhanced T1-weighted MR Imaging–Value in Detection. *Radiology*, **255** (3), 866-872, 2010. R10215.

[213] Dejardin, D., Lesaffre, E. and G. Verbeke, Joint modelling of progression-free survival and death in advanced cancer clinical trials. *Statistics in Medicine*, **29**, 1724-1734, 2010. R10228.

[214] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, A principal. *International*, **6**, Article 20, 2010. R10095.

[215] Diya, L., Lesaffre, E., Van den Heede, K., Sermeus, W. and A. Vleugels, Establishing the relationship between nurse staffing and hospital mortality using a clustered discrete-time logistic model. *Statistics in medicine*, **29** (7-8), 778-785, 2010. R10208.

[216] Dubois, C., Pappas, C., Belmans, A., Erven, K., Adriaenssens, T., Sinnaeve, P., Coosemans, M., Kayaert, P., Weltens, C. and W. Desmet, Clinical outcome of coronaroy stenting after thoracic radiotherapy: a case-control study. *Acta Cardiologica*, **65** (1), 134-134, 2010. R10241.

[217] Dubois, C., Liu, X., Claus, P., Marsboom, G., Pokreisz, P., Vandenwijngaert, S., Dpelteau, H., Streb, W., Chaothawee, L., Maes, F., Gheysens, O., Debyser, Z., Gillijns, H., Pellens, M., Vandendriessche, T., Chuah, M., Collen, D., Verbeken, E., Belmans, A., Van de Werf, F., Bogaert, J. and S. Janssens, Differential effects of progenitor cell populations on left ventricular remodeling and myocardial neovascularization after myocardial infarction. *Journal of the American College of Cardiology*, **55** (20), 2232-2243, 2010. R10216.

[218] Dubois, C., Pappas, C., Belmans, A., Erven, K., Adriaenssens, T., Sinnaeve, P., Coosemans, M., Kayaert, P., Weltens, C. and W. Desmet, Clinical outcome of coronary stenting after thoracic radiotherapy: a case-control study. *Heart*, **96** (9), 678-682, 2010. R10217.

[219] Flameng, W., Herregods, M., Vercalsteren, M., Herijgers, P., Bogaerts, K. and B. Meuris, Prosthesis-patient mismatch predicts structural valve degeneration in bioprosthetic heart valves. *Circulation*, **121** (19), 2123-2129, 2010. R10218.

[220] Forcheh Chiara, A., Verbeke, G., Valkenborg, D., Zhao, H., Marchal, K. and K. Engelen, Statistical approach to assign functional regulators in gene expression compendia. *Online Journal of Bioinformatics*, **11**, 145-154, 2010. R10226.

[221] Garcia-Zattera, M., Mutsvari, T., Jara, A., Declerck, D. and E. Lesaffre, Correcting for misclassification for a monotone disease process with an application in dental research. *Statistics in Medicine*, **29** (30), 3103-3117, 2010. R10238.

[222] Ghidey W., Lesaffre E. and G. Verbeke, A comparison of methods for estimating the random effects distribution of a linear mixed model. *Statistical Methods in Medical Research*, **19**, 575-600, 2010. R10209.

[223] Girma, M.A., Hens, N., Aets, M., Coenen, S., Uller, A., Adriaenssens, N., Beutels, P., Molenberghs, G. and H. Goossens, In Bowman A.C., editor. Time series and mixed models to study the country-specic outpatient antibiotic use in Europe. Proceedings of the 25th Workshop on Statistical Modelling, July 5-9, 2010 Glasgow. *Scotland, U.K.*, 53-56, 2010. R10201.

[224] Goegebeur, Y., De Boeck, P., Molenberghs, G., and M.G. Kenward, Person fit for test speededness: Normal curvature, likelihood ratio tests, and empirical Bayes estimates. *Methodology*, **6**, 3-16, 2010. R10125.

[225] Gyselaers, W., Molenberghs, G., Hanssens, M., and L. Peeters, Maternal renal interlobar vein impedance index is higher in early- than in late-onset preeclampsia. *Ultrasound in Obstetrics and Gynecology*, **36**, 69-75, 2010. R10113.

[226] Hens, N., Van Ranst, M., Aerts, M., Robesyn, E., Van Damme, P. and P. Beutels, Estimating the effective reproduction number for pandemic influenza from notification data made publicly available in real time: a multi-country analysis for influenza A/H1N1v 2009. *Vaccine*, 2010. R10148.

[227] Hermans, H., Tjahjono, M., Faes, D., Belmans, A., Meuris, B., Herijgers, P., Flameng, W. and M. Herregods, Mid-term follow up of triple valve surgery in a western community: predictors of survival. *The Journal of Heart Valve Disease*, **19** (5), 644-51, 2010. R10213.

[228] Jacobs, T., Straetemans, R., Molenberghs, G., Bouwknecht, A., and L. Bijnens, A latent pharmacokinetic time profile to model dose-response survival data. *Journal of Biopharmaceutical Statistics*, **20**, 759-767, 2010. R10118.

[229] Jansen, I. and G. Molenberghs, Pattern-mixture models for categorical outcomes with non-monotone missingness. *Journal of Statistical Computation and Simulation*, **80**, 1279-1296, 2010. R10107.

[230] Jarai, R., Huber, K., Bogaerts, K., Sinnaeve, P., Ezekowitz, J., Ross, A., Zeymer, U., Armstrong, P. and F. Van de Werf, Prediction of cardiogenic shock using plasma B-type natriuretic peptide and the N-terminal fragment of its pro-hormone concentrations in ST elevation myocardial infarction: an analysis from the ASSENT IV percutaneous coronary intervention trial. *Critical Care Medicine*, **38** (9), 2010. R10233.

[231] Jarai, R., Huber, K., Bogaerts, K., Droogne, W., Ezekowitz, J., Granger, C., Sinnaeve, P., Ross, A., Zeymer, U., Armstrong, P. and F. Van de Werf, Plasma N-terminal fragment of the prohormone B-type natriuretic peptide concentrations in relation to time to treatment and Thrombolysis in Myocardial Infarction (TIMI) flow: a substudy of the Assessment of the Safety and Efficacy of a New Treatment Strategy with Percutaneous Coronary Intervention (ASSENT IV-PCI) trial. *American Heart Journal*, **159** (1), 131-140, 2010. R10223.

[232] Komarek A., Hansen B.E., Kuiper E.M., van Buuren H.R. and E. Lesaffre, Discriminant analysis using a multivariate linear mixed model with a normal mixture in the random effects distribution. *Statistics in Medicine*, **29**, 3267-3283, 2010. R10206.

[233] Laenen, A. and A. Alonso, The Functional Living Index - Cancer: Estimating its Reliability based on Clinical Trial Data. *Quality of Life Research*, **19**, 103-109, 2010. R10232.

[234] Laenen, A., Alonso, A., Molenberghs, G., Vangeneugden, T. and C. Mallinckrodt, Coping with memory effect and serial correlation when estimating reliability in a longitudinal framework. *Applied Psychological Measurement*, **34** (4), 255-266, 2010. R10234.

[235] Lin, D., Shkedy, Z., Cortinas Abrahantes, J., Tilahun, A., Molenberghs, G., Talloen, W., Goehlmann, H., and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiments. *Online Journal of Bioinformatics*, **11**, 106-127, 2010. R10126.

[236] Lin, D., Shkedy, Z., Molenberghs, G., Talloen, W., Goehlmann, H.W.H. and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiment. *Online Journal of Bioinformatics*, **11** (1), 106-127, 2010. R10167.

[237] Little, R.J.A., D'Agostino, R., Dickersin, K., Emerson, S.S., Farrar, J.T., Frangakis, C., Hogan, J.W., Molenberghs, G., Murphy, S.A., Neaton, J.D., Rotnitzky, A., Scharfstein, D., Shih, W., Siegel, J.P. and H. Stern, The prevention and treatment of missing data in clinical trials. Panel on handling Missing Data in Clinical Trials. *Washington, D.C.: The National Academies Press*, 2010. R10205.

[238] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Design archetypes for phase II clinical trials in central nervous system disorders. *Drug Information Journal*, **44**, 421-430, 2010. R10110.

[239] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for using positive controls in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 431-441, 2010. R10111.

[240] Malhotra-Kumar, S., Van Heirstraeten, L., Lee, A., Cortinas Abrahantes, J., Lammens, C., Vanhommerig, E., Molenberghs, G., Aerts, M., Harbart, S., Goossens, H., on behalf of the MOSAR WP2 Study Team, Evaluation of molecular assays for rapid detection of methicillin-resistant Staphylococcus aureus. *Journal of Clinical Microbiology*, **48**, 4598-4601, 2010. R10122.

[241] Malhorta-Kumar, S., Cortinas Abrahantas, J., Sabiiti, W., Lammens, C., Vercauteren, G., Ieven, M. Molenberghs, G., Aerts, M., Goossens, H., on behalf of the MOSAR XP2 Study Team, Evaluation of chromogenetic media for detection of methicillin-reisistant Staphylococcus anreaus. *Journal of Clinical Microbiology*, **48**, 1040-1046, 2010. R10153.

[242] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for comparing a test drug with standard of care in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 443-452, 2010. R10112.

[243] Mebis, J., Jansens, H., Minalu, G., Molenberghs, G., Schroyens, W.A., Gadisseur, A.P., Van de Velde, A., Vrelust, I., Goossens, H., and Z.N. Berneman, Long-term epidemiology of bacterial susceptibility profiles in adults suffering from febrile neutropenia with hematologic malignancy after antibiotic change. *Infection and Drug Resistance*, **3**, 53-61, 2010. R10128.

[244] Molenberghs, G. and M.G. Kenward, Semi-parametric marginal models for hierarchical data and their corresponding full models. *Computational Statistics and Data Analysis*, **54**, 585-597, 2010. R10121.

[245] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11012.

[246] Molenberghs, G., Beunckens, C., Jansen, I., Thijs, H., Verbeke, G. and M.G. Kenward, Donnes manquantes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta. (Eds.)* Paris: Editions Technip., 2010. R10180.

[247] Molenberghs, G., Book Review of Falagas, M.E. (2009). Meta-analysis in Infectious Diseases, An Issue of Infectious Disease Clinics (The Clinics: Internal Medicine), Published by WB Saunders, 240 p. *HIV Clinical Trials*, **11**, 121-123, 2010. R10139.

[248] Molenberghs, G., Burzykowski, T., Alonso, A., Assam, P., Tilahun, A., and M. Buyse, A unified framework for the evaluation of surrogate endpoints in clinical trials. *Statistical Methods in Medical Research*, **19**, 205-236, 2010. R10115.

[249] Molenberghs, G., Editorial on "Generalized estimating equations: notes on the choice of the working correlation matrix (A. Ziegler, M. Vens)". *Methods of Information in Medicine*, **49**, 419-420, 2010. R10138.

[250] Molenberghs, G., Incomplete data in clinical and epidemiological studies. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10133.

[251] Molenberghs, G., Incomplete data in clinical and epidemiological studies. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10129.

[252] Molenberghs, G., Linear Mixed Model. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10131.

[253] Molenberghs, G., Linear Mixed Model (version 3). *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10135.

[254] Molenberghs, G., Meta-analysis in Infectious Diseases, An Issue of Infectious Disease Clinics (The Clinics: Internal Medicine). *HIV Clinical Trials*, **11**, 121-123, 2010. R10244.

[255] Molenberghs, G., Repeated Measures. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10130.

[256] Molenberghs, G., Repeated Measures. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10134.

[257] Molenberghs, G., Sensitivity Analysis. *International Encyclopedia of Statistical Sciences*, 2010. R10132.

[258] Molenberghs, G., Sensitivity Analysis. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10136.

[259] Molenberghs, G., The applied statistical scientist in a high-profile academic environment. In: Data and context in statistical education. Towards an evidence-based society. *Proceedings of the 8th International Conference on Teaching Statistics, C. Reading (ed.)* July 12-16, Ljubljana, Slovenia (6p.), 2010. R10193.

[260] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, **21**, 187-206, 2011. R11013.

[261] Poleto, F.Z., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. In: Proceedings of the 25th International Workshop on Statistical Modeling. *Glasgow, United Kingdom, A. Bowman (ed.)* 445-448, 2010. R10192.

[262] Moons, P., Bovijn, L., Budts, W., Belmans, A. and M. Gewillig, Temporal trends in survival to adulthood among patients born with congenital heart disease from 1970-1992 in Belgium. *Circulation*, **122**, 2264-2272, 2010. R10210.

[263] Mutsvari, T., Lesaffre, E., Garcia-Zattera, M., Declerck, D. and L. Diya, Factors that influence data quality in caries experience detection: a multi-level modeling approach. *Caries Research*, **44** (5), 438-444, 2010. R10211.

[264] Pittayapat, P., Thevissen, P., Fieuws, S., Jacobs, R. and G. Willems, Forensic oral imaging quality of hand-held dental x-ray devices: comparison of two image receptors and two devices. *Forensic Science International*, **194**, 20-27, 2010. R10229.

[265] Pramana, S., Lin, D., Haldermans, P., Shkedy, Z., Verbeke, T., Ghlmann, H., De Bondt, A., Talloen, W., and L. Bijnens, ISoGene: an R package for analyzing dose-response studies in microarray experiments. *The R Journal*, **2/1**, 5-12, 2010. R10160.

[266] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011.

[267] Pryseley, A., Mintiens, K., Knapen, K., Van der Stede, Y., and G. Molenberghs, Estimating precision, repeatability, and reproducibility from Gaussian and non-Gaussian data: a mixed models approach. *Journal of Applied Statistics*, **37**, 1729-1739, 2010. R10108.

[268] Rizopoulos, D., Verbeke, G., and G. Molenberghs, Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival data. *Biometrics*, **66**, 20-29, 2010. R10123.

[269] Serroyen, J., Molenberghs, G., Aerts, M., Vloeberghs, E., De Deyn, P.P., and G. Verbeke, Flexible estimation of serial correlation in linear mixed models. *Journal of Applied Statistics*, **37**, 833-846, 2010. R10114.

[270] Serroyen, J., Bruckers, L., Rogiers, G., and G. Molenberghs, Characterizing persistent disturbing behavior using longitudinal and multivariate techniques. *Journal of Applied Statistics*, **37**, 341-355, 2010. R10120.

[271] Siebens, K., Miljoen, H., Fieuws, S., Drew, B., De Geest, S. and C. Vrints, Implementation of the guidelines for the management of patients with chest pain through a critical pathway approach improves length of stay and patient satisfaction but not anxiety. *Critical Pathways in Cardiology*, **9**(1), 30-34, 2010. R10225.

[272] Sotto, C., Beunckens, C., Molenberghs, G., and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2010. R10127.

[273] Sunaert, P., Bastiaens, H., Nobels, F., Feyen, L., Verbeke, G., Vermeire, E., De Maeseneer, J., Willems, S. and A. De Sutter, Effect evaluation of the introduction of a Chronic Care Model-based program for type 2 diabetes in Belgium. *BMC Health Services Research*, **10**, 2010. R10227.

[274] Thevissen, P., Fieuws, S. and G. Willems, Human dental age estimation using third molar developmental stages: does a Bayesian approach outperform regression models to discriminate between juveniles and adults?. *International Journal of Legal Medicine*, **124** (1), 35-42, 2010. R10224.

[275] Tilahun, A., Lin, D., Shkedy, Z., Geys, H., Alonso, A., Peeters, P. Talloen, W., Drinkenburg, W., Goehlmann, H., Gorden, E., Bijnens, L. and G. Molenberghs, Genomic biomarkers for depression feature-specific and joint biomarkers. *Statistics in Biopharmaceutical Research*, **2**, 3, 419-434, 2010. R10168.

[276] Troxel, A., Lipsitz, S., Fitzmaurice, G., Ibrahim, J., Sinha, D., and G. Molenberghs, A weighted combination of pseudo-likelihood estimators for longitudinal binary data subject to non-ignorable non-monotone missingness. *Statistics in Medicine*, **29**, 1511-1521, 2010. R10117.

[277] Tsonaka, R., Rizopoulos, D., Verbeke, G. and E. Lesaffre, Non-ignorable models for intermittently missing categorical longitudinal responses. *Biometrics*, **66** (3), 834-844, 2010. R10214.

[278] Vandekerckhove, J., Verheyen, S. and F. Tuerlinckx, A crossed random effects diffusion model for speeded semantic categorization decisions. *Acta Psychologica*, **133** (3), 269-282, 2010. R10069.

[279] Van den Heede, K., Simoens, S., Diya, L., Lesaffre, E., Vleugels, A. and W. Sermeus, Increasing nurse staffing levels in Belgian cardiac surgery centres: a cost-effective patient safety intervention?. *Journal of Advanced Nursing*, **66** (6), 1291-1296, 2010. R10239.

[280] Van den Heede, K., Aiken, L., Diya, L., Luk, B. and W. Sermeus, Predictive validity of the RN4CAST-instruments: a pilot Study. *Journal of Clinical Nursing*, **19**, European Nursing Congress Older Persons: the future of care. Rotterdam, october 4-7 2010 (pp. 121-121) Blackwell Scientific Publications, 2010. R10243.

[281] Van de Perre, V., Permentier, L., De Bie, S., Verbeke, G. and R. Geers, Effect of unloading, lairage, pig handling, stunning and season on pH of pork. *Meat Science*, **86**, 931-937, 2010. R10240.

[282] Van de Perre, V., Permentier, L., De Bie, S., Verbeke, G. R. Geers, Effect of transport, unloading, lairage, pig handling and stunning on meat quality. *Book of Abstracts of the 61st Annual Meeting of the EAAP*, **16** (1), Heraklion, 23-27 August, 287-287, Abstract No. 36/19, 2010. R10242.

[283] Vangeneugden, T., Molenberghs, G., Laenen, A., Geys, H., Beunckens, C., and C. Sotto, Marginal correlation in longitudinal binary data based on generalized linear mixed models. *Communications in Statistics, Theory Methods*, **39**, 2540-3557, 2010. R10109.

[284] Vanhaecht, K., Bellemans, J., De Witte, K., Diya, L., Lesaffre, E. and W. Sermeus, Does the organization of care processes affect outcomes in patients undergoing total joint replacement?. *Journal of Evaluation in Clinical Practice*, **16** (1), 121-128, 2010. R10222.

[285] Vannieuwenhuyze, J., Loosveldt, G. and G. Molenberghs, A method for evaluating mode effects in mixed mode surveys. *Public Opinion Quarterly*, **75** (5), 1027-1043, 2010. R10237.

[286] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2010. R10116.

[287] Verbeke, G., Molenberghs, G. and D. Rizopoulos, Random effects models for longitudinal data. In: Longitudinal Research with Latent Variables. *K. van Montfort, H. Oud, and Al Satorra (Eds.)* New York: Springer, pp. 37-96, 2010. R10179.

[288] Verbeke, G. and G. Molenberghs, Modles marginaux et modles mixtes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta (Eds.)* Paris: Editions Technip., 2010. R10181.

[289] Verhamme, P., Pakola, S., Jensen, T., Berggren, K., Sonesson, E., Saint-Remy, J., Balchen, T., Belmans, A., Cahillane, G., Stassen, J., Peerlinck, K., Glazer, S. and M. Jacquemin, Tolerability and Pharmacokinetics of TB-402 in Healthy Male Volunteers. *Clinical Therapeutics*, **32** (6), 1205-1220, 2010. R10230.

[290] Willems, G., Thevissen, P., Belmans, A. and H. Liversidge, Willems II. Non-gender-specific dental maturity scores. *Forensic Science International*, **201** (1-3), 84-85, 2010. R10212.

[291] Wouters, K., Cortinas Abrahantas, J., Molenberghs, G., Geys, H., Bijnens, L. Ahnaou, A. and W.H.I.M. Drinkenburg, Correction for model selection bias using a modified model averaging approach for supervised learning methods applied to EEG experiment. *Journal of Biopharmaceutical Statistics*, **20** (4), 768-786, 2010. R10157.

[292] Zhang Y., Todem D., Kim K., and E. Lesaffre, Bayesian latent variable models for spatially correlated tooth-level binary data in caries research. *Statistical Modelling*, **11**, 25-47, 2010. R10207.

3.3.3 Refereed publications (in press)

[293] Agbaje, J., Mutsvari, T., Lesaffre, E. and D. Declerck, Measurement, analysis and interpretation of examiner reliability in caries experience surveys: some methodological thoughts. *Clinical Oral Investigations*, 2010. RP10147.

[294] Alqerban, A., Jacobs, R., Fieuws, S. and G. Willems, Comparison of the accuracy of two cone beam computed tomography systems versus panoramic imaging for localization of impacted maxillary canines and detection of root resorption on maxillary lateral incisors. *European Journal of Orthodontics*, 2010. RP10038.

[295] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Information-theoretic approach to surrogate markers evaluation with failure time endpoints. *Lifetime Data Analysis*, 2010. RP10102.

[296] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple imputation based generalized estimating equations. *Journal of Biopharmaceutical Statistics*, 2010. RP10107.

[297] Breusegem, C., Spielberg, L., Van Ginderdeuren, R., Vandewalle, E., Renier, C., Van de Veire, S., Fieuws, S., Zeyen, T. and I. Stalmans, Preoperative topical non-steroidal anti-inflammatory drug or steroid and clinical outcomes after trabeculectomy. A randomized controlled trial. *Ophthalmology*, 2010. RP10037.

[298] Cortinas, J., Sotto, C., Molenberghs, G., Vromman, G. and B. Bierinckx, A comparison of various software tools for dealing with missing data via imputation. *Journal of Statistical Computation and Simulation*, 2010. RP10100.

[299] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward, Shared-parameter models and missingness at random. *Statistical Modelling*, 2010. RP10109.

[300] Desmet, W., Bogaert, J., Dubois, C., Adriaenssens, T., Pappas, C., Ganame, J., Dymarkowski, S., Belmans, A. and F. Van de Werf, High Dose Intracoronary Adenosine for Myocardial Salvage in Patients with Acute ST-segment Elevation Myocardial Infarction. *Circulation*, 2010. RP10139.

[301] Desmet, W., Bogaert, J., Dubois, C., Sinnaeve, P., Adriaenssens, T., Pappas, C., Ganame, J., Dymarkowski, S., Janssens, S., Belmans, A. and F. Van de Werf, High-dose intracoronary adenosine for myocardial salvage in patients with acute ST-segment elevation myocardial infarction. *European Heart Journal*, 2010. RP10145.

[302] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, Discovering transgenic events: using information from early screening trials for improving experimental design. *Biological and Environmental Statistics*, 2010. RP10142.

[303] Diya, L., Lesaffre, E., Van den Heede, K., Sermeus, W. and A. Vleugels, Relationship between nurse staffing and hospital mortality: A multi-center nursing unit level study using clustered discrete-time logistic model. *Statistics in Medicine*, 2010. RP10028.

[304] Goderis, G., Borgermans, L., Heyrman, J., Van Den Broeke, C., Carbonez, A., Mathieu, C., Verbeke, G. and R. Grol, Monitoring modifiable cardiovascular risk in type 2 diabetes care in general practices: The use of an aggregated z-score. *Medical Care*, 2010. RP10032.

[305] Hernndez, M., Sorsa, T., Obregn, F., Tervahartiala, T., Valenzuela, M-A, Pozo, P., Dutzan, N., Lesaffre., E., Molas M. and J. Gamonal, MMP-13 enhances MMP-9 activation during progression of chronic periodontitis. *Journal of Clinical Periodontology*, 2010. RP10027.

[306] Hompes, R., Fieuws, S., Aerts, R., Thijs, M., Penninckx, F. and B. Topal, Results of single-probe microwave ablation of metastatic liver cancer. *European Journal of Surgical Oncology*, 2010. RP10140.

[307] Jara, A., Hanson, T.E. and E. Lesaffre, Robustifying generalized linear mixed models using mixtures of multivariate Polya trees. *Journal of Computational Statistics and Graphics*, 2010. RP10026.

[308] Leonard, D.,Penninckx, F., Fieuws, S., Jouret-Mourin, A., Sempoux, C., Jehaes, C. and E. Van Eycken, MDFactors predicting the quality of total mesorectal excision for rectal cancer. *Annals of Surgery*, 2010. RP10033.

[309] Litire, S., Alonso, A., and G. Molenberghs, Correction. *Statistics in Medicine*, 29, 2010. RP10097.

[310] Litire, S., Alonso, A. and G. Molenberghs, Rejoinder to reader reaction: Type I and type II error random-effects misspecification in generalized linear mixed models. *Biometrics*, 2010. RP10099.

[311] Molas, M. and E. Lesaffre, Hurdle models for multilevel zero-inflated data via H-likelihood. *Statistics in Medicine*, 2010. RP10031.

[312] Molenberghs, G. and G. Verbeke, A note on a hierarchical interpretation for negative variance components. *Statistical Modelling*, 2010. RP10137.

[313] Molenberghs, G., Verbeke, G., Demtrio, C.G.B. and A. Vieira, A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, 2010. RP10098.

[314] Mwambi, H., Ramroop, S., Shkedy, Z. and G. Molenberghs, An approach to estimating the force of infection and the rate of recovery from Respiratory Syncytial Virus (RSV) data. *Statistical Methods in Medical Research*, 2010. RP10069.

[315] Poleto, F., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. *Test*, 2010. RP10105.

[316] Roca Pardias, J., Cadarso Surez, C.M., Pardo Vzquez, J.L., Leborn, V., Molenberghs, G., Faes, C. and C. Acua, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*, 2011. RP11005.

[317] Siebens, K., Miljoen, H., Fieuws, S., Drew, B., De Geest, S. and C. Vrints , Implementation of the guidelines for the management of patients with chest pain through critical pathway approach improves length of stay and patient satisfaction but not anxiety. *Critical Pathways in Cardiology*, 2010. RP10034.

[318] Sinha, S.K., Troxel, A.B., Lipsitz, S.R., Sinha, D., Fitzmaurice, G.M., Molenberghs, G. and J.G. Ibrahim, A bivariate pseudo-likelihood for incomplete longitudinal binary data with nonignorable non-monotone missingness. *Biometrics* 66, 2010. RP10104.

[319] Star, H., Thevissen, P., Jacobs, R., Fieuws,S., Solheim, T. and G. Willems, Human dental age estimation by calculation of pulp/tooth volume ratios yielded on clinically acquired cone beam computed tomography (CBCT) images of mono radicular teeth. *Journal of Forensic Sciences*, 2010. RP10035.

[320] Tomsin, K., Mesens, T., Molenberghs, G., Peeters, L. and W. Gyselaers, Time-interval between maternal electrocardiogram and venous Doppler waves in normal pregnancy and preeclampsia: a pilot study. *European Journal of Ultrasound*, 2010. RP10101.

[321] Vandenbroeck, S., De Geest, S., Dobbels, F., Fieuws, S., Stalmans, I. and Zeyen, The Prevalence and Correlates of Self-reported Non-adherence: The Belgian Compliance Study in Ophthalmology (BCSO). *Acta Ophthalmologica*, 2010. RP10144.

[322] VandeVoorde P., Sabbe, M., Rizopoulos, D., Tsonaka, S., De Jaeger, E., Lesaffre, E., Peters, M. and The Penta study group, Assessing the level of consciousness in children: a plea for the Glasgow Coma Motor Subscore. *Resuscitation*, 2010. RP10022.

[323] VandeVoorde P., Sabbe, M., Rizopoulos, D., Tsonaka, S., De Jaeger, E., Lesaffre, E., Peters, M. and The Penta study group, Paediatric trauma and trauma care in Flanders (Belgium). Methodology and first descriptive results of the Penta registry. *European Journal of Pediatrics*, 2010. RP10023.

[324] Vandekerckhove, J., Tuerlinckx, F. and M. Lee, Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 2010. RP10143.

[325] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demtrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, 2010. RP10096.

[326] Verbeke, G. and E. Lesaffre, Repeated Measurements. *Encyclopedia of Statistical Sciences*, New York:. Wiley Sons, 2010. RP10146.

[327] Vieira, A.M.C., Leandro, R.A., Demtrio, C.G.B. and G. Molenberghs, Double generalized linear model for tissue culture proportion data: A Bayesian perspective. *Journal of Applied Statistics*, 2010. RP10103.

3.4 Universiteit Gent, UG

3.4.1 Refereed publications (published)

[328] Abebe, G., Deribew, A., Apers, L., Woldemichael, K., Shiffa, J., Tesfaye, M., Abdissa, A., Deribie, F., Jira, C., Bezabih, M., Aseffa, A., Duchateau, L. and Colebunders, R., Knowledge, Health Seeking Behavior and Perceived Stigma towards Tuberculosis among Tuberculosis Suspects in a Rural Community in Southwest Ethiopia. *PLOS ONE*, **5**, Article Number e13339, 2010. R10096.

[329] Babanezhad, M., Vansteelandt, S. and E. Goetghebeur, Comparison of causal effect estimators under exposure misclassification. *Journal of Statistical Planning and Inference*, **140**, 1306-1319, 2010. R10021.

[330] Barberet, V., Pey, P., Duchateau, L., Combes, A., Daminet, S. and J.H. Saunders, Intra- and interobserver variability of ultrasonographic measurements of the adrenal glands in helathy beagles. *Vet. Radiol. Ultrasound*, **51**, 656-660, 2010. R10093.

[331] Baele, G., van de Peer, Y. and S. Vansteelandt, Using nonreversible context-dependent evolutionary models to study substitution patterns in primate non-coding sequences. *Journal of Molecular Evolution*, **71**, 34-50, 2010. R10087.

[332] Baele, G., van de Peer, Y. and S. Vansteelandt, Modelling the ancestral sequence distribution and model frequencies in contextdependent models for primate non-coding sequences. *BMC Evolutionary Biology*, **10**, Article 244, 2010. R10088.

[333] Bekaert, M., Benoit, D., Decruyenaere, J. and S. Vansteelandt, A note on statistical association and causality derived from epidemiological ICU data - Reply. *Intensive Care Medicine*, **36**, 550, 2010. R10019.

[334] Bekaert, M., Vansteelandt, S. and K. Mertens, Adjusting for time-varying confounding in the subdistribution analysis of a competing risk. *Lifetime Data Analysis*, **16**, 45-70, 2010. R10018.

[335] De Decker, S., Gielen, I. M. V. L., Duchateau, L., Polis, I., Van Bree, H. J. J. and L.M.L. Van Ham, Agreement and Repeatability of Linear Vertebral Body and Canal Measurements Using Computed Tomography (CT) and Low Field Magnetic Resonance Imaging (MRI). *Vet. Surg.*, **39**, 28-34, 2010. R10102.

[336] Deribew, A., Tesfaye, M., Hailmichael, Y., Apers, L., Abebe, G., Duchateau, L. and R. Colebunders, Common mental disorders in TB/HIV co-infected patients in Ethiopia. *BMC Infectious diseases*, **10**, Article Number 201, 2010. R10098.

[337] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, A principal. *International*, **6**, Article 20, 2010. R10095.

[338] Erkkola, R., Vervarcke S., Vansteelandt S., Rompotti P., De Keukeleire, D. and Heyerick, A., A randomized, double-blind, placebocontrolled, cross-over pilot study on the use of a standardized hop extract to alleviate menopausal discomforts. *Phytomedicine*, **17**, 389-396, 2010. R10085.

[339] Himpens, E., Oostra, A., Franki, I., Vansteelandt, S., Vanhaesebrouck, P. and C. Van den Broeck, Predictability of Cerebral Palsy in a high-risk NICU population. *Early Human Development*, **86**, 413-417, 2010. R10089.

[340] Hoffmann, T.J., Lange, C., Vansteelandt, S., Raby, B.A., Demeo, D.L., Silverman, E.K. and N.M. Laird, Parsing the Effects of Individual SNPs in Candidate Genes with Family Data. *Human Heredity*, **69**, 91-103, 2010. R10017.

[341] Khan, J.A., Van Aelst, S., and R.H. Zamar, Fast robust estimation of prediction error based on resampling. *Computational Statistics and Data Analysis*, **54** (12), 3121-3130, 2010. R10073.

[342] Lampo, E., Duchateau, L., Schepens, B., Van Poucke, M., Saelens, X., Erkens, T., Van Zeveren, A. and L.J. Peelman, Identification of polymorphisms in the ovine Shadow of prion protein (SPRN) gene and assessment of their effect on promoter activity and susceptibility for classical scrapie. *Anim. Genet.*, **41**, 169-178, 2010. R10100.

[343] Moerkerke, B., Vansteelandt, S. and C. Lange, A doubly-robust test for gene-environment interaction in family-based studies of affected offspring. *Biostatistics*, **11**, 213-225, 2010. R10084.

[344] Sjolander, A., Vansteelandt, S. and K. Humphreys, A principal stratification approach to assess the differences in prognosis between cancers caused by hormone replacement therapy and by other factors. *International Journal of Biostatistics*, **6**, Article 20, 2010. R10086.

[345] Smets, P.M.Y., Meyer, E., Maddens, B., Duchateau, L. and S. Daminet, Effect of sampling method and storage conditions on albumin, retinol-binding protein, and N-acetyl-beta-D-glucosaminidase concentrations in canine urine samples. *J. Vet. Diagn. Invest.*, **22**, 896-902, 2010. R10094.

[346] Stuyven, E., Verdonck, F., Van Hoek, I., Daminet, S., Duchateau, L., Remon, J. P., Goddeeris, B. M. and E. Cox, Oral Administration of beta-1,3/1,6-Glucan to Dogs Temporally Changes Total and Antigen-Specific IgA and IgM. *Clin. Vaccine Immunol.*, **17**, 281-285, 2010. R10101.

[347] Van Aelst, S., and G. Willems, Inference for robust canonical variate analysis. *Advances in Data Analysis and Classification*, **4** (2-3), 181-197, 2010. R10074.

[348] Van Aelst, S., Vandervieren, E., and G. Willems, Stahel-Donoho Estimators with Cellwise Weights. *Journal of Statistical Computation and Simulation*, **81** (1), 1-27, 2010. R10075.

[349] Van Aelst, S., and G. Willems, Robust principal component analysis based on pairwise correlation estimators. *COMPSTAT 2010: Proceedings in Computational Statistics* Y. Lechevallier and G. Saporta, Heidelberg: Physika-Verlag, 573-580, 2010. R10076.

[350] Van Aelst, S., and G. Willems, Selecting variables in two-group robust linear discriminant analysis. *COMPSTAT 2010: Proceedings in Computational Statistics* Y. Lechevallier and G. Saporta, Heidelberg: Physika-Verlag, 1677-1684, 2010. R10077.

[351] VanderWeele, T.J. and S. Vansteelandt, Odds ratios for mediation analysis for a dichotomous outcome. *American Journal of Epidemiology*, **172**, 1339-1348, 2010. R10090.

[352] VanderWeele, T.J. and S. Vansteelandt, VanderWeele and Vansteelandt Respond to 'Decomposing With a Lot of Supposing' and 'Mediation'. *American Journal of Epidemiology*, **172**, 1355-1356, 2010. R10091.

[353] VanderWeele, T.J., Vansteelandt, S. and J.M. Robins, Marginal structural models for sufficient cause interactions. *American Journal of Epidemiology*, **171**, 506-514, 2010. R10016.

[354] Vansteelandt, S., Estimation of controlled direct effects on a dichotomous outcome using logistic structural direct effect models. *Biometrika*, **97**, 921-934, 2010. R10092.

[355] Vansteelandt, S., Carpenter, J. and M.G. Kenward, Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology*, **6**, 37-48, 2010. R10079.

[356] Villarreal, I., Vranckx, K., Duchateau, L., Pasmans, F., Haesebrouck, F., Jensen, J.C., Nanjiani, I.A. and Maes, D., Early Mycoplasma hyopneumoniae infections in European suckling pigs in herds with respiratory problems: detection rate and risk factors. *Vet. Med.*, **55**, 318-324, 2010. R10097.

[357] Yewhalaw, D., Van Bortel, W., Denis, L., Coosemans, M., Duchateau, L. and N. Speybroeck, First Evidence of High Knockdown Resistance Frequency in Anopheles arabiensis (Diptera: Culicidae) from Ethiopia. *Am. J. Trop. Med. Hyg.*, **83**, 122-125, 2010. R10099.

3.4.2 Refereed publications (in press)

[358] Buyze, J., Van Rompaye, B. and E. Goetghebeur, Designing a sequentially randomized study with adherence enhancing interventions for diabetes patients. *Statistics in Medicine*, 2010. RP10041.

[359] De Bock, K.W., Coussement, K. and D. Van Den Poel, Ensemble classification based on generalized additive models. *Computational Statistics and Data Analysis*, 2010. RP10045.

[360] De Boeck, B., Thas, O. and J.P. Ottoy, Reweighted smooth tests of goodness of fit. *Communications in Statistics - Theory and Methods*, 2010. RP10043.

[361] De Boeck, B., Thas, O., Rayner, J.C.W. and D.J. Best, Smooth tests for the gamma distribution. *Journal of Statistical Computation and Simulation*, 2010. RP10044.

[362] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, Discovering transgenic events: using information from early screening trials for improving experimental design. *Biological and Environmental Statistics*, 2010. RP10142.

[363] Manoochehr, B., Vansteelandt, S. and E. Goetghebeur, Comparison of causal effect estimators under exposure misclassification. *Journal of Statistical Planning and Inference*, 2010. RP10040.

[364] Tolley, E., Harrison, P., Goetghebeur E. and al., Adherence and its measurement in phase 2/3 microbicide trials. *AIDS and Behavior*, 2010. RP10039.

[365] Van Aelst, S., Vandervieren, E. and G. Willems, Stahel-Donoho estimators with cellwise weights. *Journal of Statistical Computation and Simulation*, 2010. RP10046.

[366] Van Rompaye, B., Jaffar, S. and E. Goetghebeur, Design and testing for clinical trials faced with misclassified causes-of-death. *Biostatistics*, 2010. RP10042.

3.5 Universiteit Hasselt, UH

3.5.1 Technical reports

[367] Van Keilegom, I. and N. Veraverbeke, Statistical models and methods for dependence in insurance data, 2011. TR11003.

[368] Vannieuwenhuyze, J. and G. Molenberghs, A SAS macro to disentangle mode effects on proportions and the mean of a categorical variable in an extended mixed-mode dataset, 11 pp. Leuven: K.U.Leuven. Ceso, 2010. TR10065.

3.5.2 Refereed publications (published)

[369] Aerts, M., Faes, C., Hens, N. and G. Molenberghs, Handling missingness and informative cluster sizes when modelling prevalence and force of infection. *Proceedings ENAR-JSM 2010* 3491-3501, 2010. R10194.

[370] Aerts, M., Hens, N. and J.S. Simonoff, selection based on presmoothing. *Journal of Applied Statistics*, **37-9**, 1455, 2010. R10140.

[371] Alonso, A., Litire, S. and G. Molenberghs, Testing for misspecification in generalized linear mixed models. *Biostatistics*, **11**, 771-786, 2010. R10125.

[372] Alonso, A., Litire, S. and A. Laenen, A note on the indeterminacy of the random-effects distribution in hierarchical models. *American Statistician*, **64** (4), 318-324, 2010. R10235.

[373] Alonso, A., Laenen, A., Molenberghs, G., Geys, H. and T. Vangeneugden, Reliability of Single Administered Tests: A Unified Approach. *Biometrics*, **66** (4), 1061-1068, 2010. R10236.

[374] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Using earlier measures in a longitudinal sequence as potential surrogate for a later one. *Computational Statistics and Data Analysis*, **54**, 1342-1354, 2010. R10137.

[375] Biard, R., Loisel, S., Macci, C. and N. Veraverbeke, Asymptotic behavior of the finite-time expected time-integrated negative part of some risk processes and optimal reserve allocation. *Journal of Mathematical Analysis and Applications*, **367**, 535-549, 2010. R10103.

[376] Bilcke, J., van Hoek, A., Hens, N., Aerts, M., Edmunds, W. and P. Beutels, Cost-effectiveness of varicella and zoster vaccination in England and Wales: importance measures for correlated parameters. *Procedia: Social and Behavioral Sciences*, **2**, 7611-7612, 2010. R10141.

[377] Bollaerts, K., Messens, W., Aerts, M., Dewulf, J., Maes, D., Grijspeerdt, K. and Y. Van der Stede, Evaluation of scenarios for reducing human salmonellosis through household consumption of fresh minced pork meat. *Risk Analysis*, **30**, 853-865, 2010. R10152.

[378] Boone, I., Van der Stede, Y., Aerts, M., Mintiens, K. and G. Daube, Quantitative microbial risk assessment: methods and quality assurance. *Vlaams Diergeneeskundig Tijdschrift*, **79**, 367-380, 2010. R10188.

[379] Boone, I., Van der Stede, Y., Dewulf, J., Messens, W., Aerts, M. Daube, G., and K. Mintiens, NUSAP: a tool to evaluate the quality of assumptions in a quantitative microbial risk assessment model for Salmonella in the pork production chain. *Journal of Risk Research*, **13**, 337-352, 2010. R10154.

[380] Broekmans, T., Roelants, M., Alders, G., Feys, P. Thijs, H. and B. Op 't Einde, Exploring the effects of a 20-week whole-body vibration training programme on leg muscle performance and function in persons with multiple sclerosis. *Journal of Rehabilitation Medicine*, **42** (9), 866-872, 2010. R10178.

[381] Bruckers, L., Serroyen, J., Molenberghs, G., Slaets, H., and W. Goeyvaerts, Latent-class analysis of persistent disturbing behavior patients using longitudinal profiles. *Applied Statistics*, **59**, 495-512, 2010. R10119.

[382] Burzykowski, T., Carpenter, J., Coens, C., Evans, D., France, L. Kenward, M. and Lane, P., Matcham, J. Morgan, D., Philips, A., Roger, J., Sullivan, B., White, I., Yu, L.M., of the PSI missing data expert group, Missing data: discussion points from the PSI missing data expert Group. *Pharmaceutical Statistics*, **6**, 2010. R10164.

[383] Burzykowski, T. and M. Buyse, The correlation structure of longitudinal measurements of vision in patients with macular degeneration. *Pharmaceutical Statistics*, 2010. R10165.

[384] Cortinas Abrahantes, J. and T. Burzykowski, Simplified modeling strategies for surrogate validation with multivariate failure-times data. *Computational Statistics and Data Analysis*, **54**, 1457-1466, 2010. R10161.

[385] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G., and M.G. Kenward, A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes. *Biometrical Journal*, **52**, 111-125, 2010. R10124.

[386] Creemers, A., Hens, N., Aerts, M. Molenberghs, G. Verbeke, G. and M.G. Kenward, A sensitivity analysis for shared-parameter models. *The Biometrical Journal*, **52**, 111-125, 2010. R10142.

[387] Cuypers, K., Levin, O., Thijs, H., Swinnen, S.R. and R. Meesen, Long-term TENS treatment improves tactile sensitivity in MS patients. *Neurorehabilitation and Neural Repair*, **24** (5), 420-427, 2010. R10175.

[388] De Bock, K., Bruckers, L. and Y. Coune, Continuteit van zorgaanbod in een ambulant Centrum voor Geestelijke Gezondheidszorg. *Zorgwijzer*, **12**, 26-28, 2010. R10186.

[389] Dhooge, W., Den Hond, E., Koppen, G., Bruckers, L. Nelen, V., Van De Mieroop, E., Bilau, M., Croes, L., Bayens, W., Schoeters, G. and N. Van Larebeke, Internal exposure to pullutants and body size in Flemish adolescents and adults: associations and dose-response relationships. *Environment International*, **36**, 4, 330-337, 2010. R10173.

[390] Dhooge, W., Bruckers, L., Schoeters, G., Nelen, V., van de Mieroop, E. Koppen, G., Bilau, M., Schroijen, C., Keune, H., Baeyens, W. and N. van Larebeke, Internal exposure to pollutants and sexual maturation in Flemish adolescents. *Journal of Exposure Science Environmental Epidemiology*, 2010. R10174.

[391] Ekamper, P., Faes, C. and F. Van Poppel, Vrijers die van verre komen zijn te schromen. Geografische homogamie van huwelijkspartners, 1812-1922. *Mens en Maatschappij*, **85**, 380-404, 2010. R10185.

[392] Falter, B., Schepers, S., Vrielinck, L. Lambrichts, I., Thijs, H. and C. Politis, Occurence of bad splits during sagittal split osteotomy. *Oral Surgery Oral Medicine Oral Phathology Oral Radiology and Endodontology*, **110** (4), 430-435, 2010. R10177.

[393] Gaddah, A. and R. Braekers, Testing under the extended Koziol-Green model. Durante, F., Hardle, W., Jaworski, P., Rychlik, T. (Eds.) Workshop on Copula Theory and its applications, Lectures Notes in Statistics 198. *Proceedings Springer, Berlin/Heidelberg*, 2010. R10195.

[394] Gaddah, A. and R. Braekers, Testing under the Koziol-Green model, Copula theory and its applications. Proceedings of the Workshop held in Warsaw 25-26 September 2009, ed. Piotr Jaworski, Fabrizio Durante, Wolfgang Karl Hrdle and Tomasz Rychlik. *Lecture Notes in Statistics* **198**, 2, 279-288, Springer, 2010. R10202.

[395] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919-1932, 2011. R11010.

[396] Girma, M.A., Hens, N., Aets, M., Coenen, S., Uller, A., Adriaenssens, N., Beutels, P., Molenberghs, G. and H. Goossens, In Bowman A.C., editor. Time series and mixed models

to study the country-specic outpatient antibiotic use in Europe. Proceedings of the 25th Workshop on Statistical Modelling, July 5-9, 2010 Glasgow. *Scotland, U.K.*, 53-56, 2010. R10201.

[397] Goegebeur, Y., De Boeck, P. and G. Molenberghs, Person fit for test speededness: normal curvatures, likelihood ratio tests and empirical Bayes estimates. *Methodology: European journal of research methods for the behavioral and social sciences*, **6**, 3-16, 2010. R10064.

[398] Goegebeur, Y., De Boeck, P., Molenberghs, G., and M.G. Kenward, Person fit for test speededness: Normal curvature, likelihood ratio tests, and empirical Bayes estimates. *Methodology*, **6**, 3-16, 2010. R10125.

[399] Goeyvaerts, N. Hens, N. Aerts, M. and P. Beutels, Models structure analysis to estimate basic immunological processes and maternal risk for parvovirus B19. *Biostatistics*, 2010. R10143.

[400] Goeyvaerts N. Hens, N. Ogunjimi, B., Aerts, M., Shkedy, Z., Van Damme, P. and P. Beutels, Estimating transmission parameters and the basic reproduction number using social contact data and serological data on varicella zoster virus in Belgium. *Journal of the Royal Statistical Society C*, **59**, 255-277, 2010. R10144.

[401] Govarts, E., Den Hond, E., Schoeters, G., Bruckers, L. (2010). Determinants of serum PCBs in adolescents and adults: regression tree analysis and linear regression analysis. *Human and Ecological Risk Assessment*, **16**, 5, 1115-1132, 2010. R10171.

[402] Gyselaers, W., Molenberghs, G., Hanssens, M., and L. Peeters, Maternal renal interlobar vein impedance index is higher in early- than in late-onset preeclampsia. *Ultrasound in Obstetrics and Gynecology*, **36**, 69-75, 2010. R10113.

[403] Harries, R., Lawson, S. and L. Bruckers, Assessment of microcalcifications with limited number of high-precision macrobiopsies. *European Journal of Cancer Prevention*, **16**, 5, 374-378, 2010. R10172.

[404] Hens, N. and A. Wienke, On the correlated gamma frailty model for current status data. Proceedings of the 19th Conference on Computational Statistics, 22-27 August 2010. *Paris France*, 2010. R10197.

[405] Hens, N. and A. Wienke, On the correlated gamma frailty model for current status data. Proceedings of the Joint Statistical Meetings, 31 July - 5 August, 2010. *Vancouver, Canada*, 2010. R10198.

[406] Hens, N., On the correlated gamma frailty model for current status data. Proceedings of the joint statistical Meetings. *Vancouver, Canada (31 July - 5 August 2010)*, 2010. R10199.

[407] Hens, N., On the correlated gamma frailty model for current status data. Proceedings of the 19th Conference on Computational Statistics, 22-27 August 2010. *Paris, France*, 2010. R10200.

[408] Hens, N., Kvitkovicova, A. , Aerts, M. Hlubinka, D. and P. Beutels, Modelling distortions in seroprevalence data using change-point fractional polynomials. *Statistical Modelling*, **10**, 159-175, 2010. R10145.

[409] Hens, N., Van Ranst, M., Aerts, M., Robesyn, E., Van Damme, P. and P. Beutels, Estimating the effective reproduction number for pandemic influenza from notification data made publicly available in real time: a multi-country analysis for influenza A/H1N1v 2009. *Vaccine*, 2010. R10148.

[410] Hochreiter, S., Bodenhofer, U., Heusel, M., Mayr, A., Mitterecker, A., Kasim, A., Khamiakova, T., Van Sanden, S., Lin, D., Talloen, W., Bijnens, L., Goehlmann, H.W.H., Shkedy, Z. and D.A. Clevert, FABIA: factor analysis for bicluster acquisition. *Bioinformatics*, **26** (12), 1520-1527, 2010. R10169.

[411] Jacobs, T., Straetemans, R., Molenberghs, G., Bouwknecht, A., and L. Bijnens, A latent pharmacokinetic time profile to model dose-response survival data. *Journal of Biopharmaceutical Statistics*, **20**, 759-767, 2010. R10118.

[412] Jansen, I. and G. Molenberghs, Pattern-mixture models for categorical outcomes with non-monotone missingness. *Journal of Statistical Computation and Simulation*, **80**, 1279-1296, 2010. R10107.

[413] Lenders, K., Ogunjimi, B., Beutels, Ph., Hens, N., Van Damme, P., Berneman, Z.N., Van Tendeloo, V.F.I. and E.L.J.M. Smits, The effect of apoptotic cells on virus-specific immune responses detected using IFN-Gamma elispot. *Journal of Immunological Methods*, **31**, 357(1-2) 51-4, 2010. R10146.

[414] Leuridan, E., Hens, N., Hutse, V., Ieven, M. Aerts, M. and P. Van Damme, Early warning of variacella and rubella antibodies. *Vaccine*, 2010. R10150.

[415] Iozzi, F., Trusiano, F., Chinazzi, M., Billari, F.C., Zagheni, E., Merler, S., Ajelli, M. Del Fava, E. and P. Mandredi, Little Italy: an agent-based approach to the estimation of contact patterns-fitting predicted matrices to serological data. *PLOS Computational Biology*, 6(12): e1001021, 2010. R10159.

[416] Kasim, A., Lin, D., Van Sanden, S., Clevert, D., Bijnens, L., Goehlmann, W.H.H., Amaratunga, D., Hochreiter, S., Shkedy, Z., and W. Talloen, Informative or noninformative calls for gene expression in microarray experiments: a latent variable approach. *Statistical Application in Genetics and Molecular Biology*, **9**, Iss. 1, Article 4, 2010. R10166.

[417] Kasim, A., Shkedy, Z. Lin, D., Van Sanden, S., Cortinas Abrahantes, J., Goehlmann, H.W., Bijnens, L., Yekutieli, D., Camilleri, M., Aerssens, J. and W. Talloen, Translating gene signatures across tissues: a statistical View, A principal. *Proceedings of the 2010 International Conference on Bioinformatics Computational Biology* 298-301, Las Vegas, USA, 2010. R10196.

[418] Laenen, A. and A. Alonso, The Functional Living Index - Cancer: Estimating its Reliability based on Clinical Trial Data. *Quality of Life Research*, **19**, 103-109, 2010. R10232.

[419] Laenen, A., Alonso, A., Molenberghs, G., Vangeneugden, T. and C. Mallinckrodt, Coping with memory effect and serial correlation when estimating reliability in a longitudinal framework. *Applied Psychological Measurement*, **34** (4), 255-266, 2010. R10234.

[420] Leuridan, E., Hens, N., Hutse, V., Ieven, M., Aerts, M. and P. Van Damme, Early warning of maternal measles antibodies in the era of global measles vaccination. *The British Medical Journal*, 340:c1626, 2010. R10147.

[421] Lin, D., Shkedy, Z., Molenberghs, G., Talloen, W., Goehlmann, H.W.H. and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiment. *Online Journal of Bioinformatics*, **11** (1), 106-127, 2010. R10167.

[422] Lin, D., Shkedy, Z., Cortinas Abrahantes, J., Tilahun, A., Molenberghs, G., Talloen, W., Goehlmann, H., and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiments. *Online Journal of Bioinformatics*, **11**, 106-127, 2010. R10126.

[423] Little, R.J.A., D'Agostino, R., Dickersin, K., Emerson, S.S., Farrar, J.T., Frangakis, C., Hogan, J.W., Molenberghs, G., Murphy, S.A., Neaton, J.D., Rotnitzky, A., Scharfstein, D., Shih, W., Siegel, J.P. and H. Stern, The prevention and treatment of missing data in clinical trials. Panel on handling Missing Data in Clinical Trials. *Washington, D.C.: The National Academies Press*, 2010. R10205.

[424] Lousbergh, D., Buntinx, F., Op de Beeck, L., Rummens, J., Vanden Brande, J., Dhollander, D., Kellen, E., Hense, K., Faes, C., Bruckers, L., Cloes, E., Lathouwers, D. and E. Meekers, Tien jaar kanker in de provincie Limburg (1996-2005): De belangrijkste individuele lokalisaties. *Tijdschrift voor Geneeskunde*, **66**, 3-10, 2010. R10182.

[425] Lousbergh, D., Buntinx, F., Op de Beeck, L., Rummens, J., Vandebrande, J. Dhollander, D., Kellen, E., Hensen, K., Faes, C., Bruckers, L. Cloes, E., Lathouwers, D. and E. Meekers, Tien jaar kanker in de provincie Limburg (1996-2005): De belangrijkste individuele lokalisaties. *Tijdschrift voor Geneeskunde*, **66**, 11-18, 2010. R10183.

[426] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Design archetypes for phase II clinical trials in central nervous system disorders. *Drug Information Journal*, **44**, 421-430, 2010. R10110.

[427] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for using positive controls in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 431-441, 2010. R10111.

[428] Malhotra-Kumar, S., Van Heirstraeten, L., Lee, A., Cortinas Abrahantes, J., Lammens, C., Vanhommerig, E., Molenberghs, G., Aerts, M., Harbart, S., Goossens, H., on behalf of the MOSAR WP2 Study Team, Evaluation of molecular assays for rapid detection of

methicillin-resistant Staphylococcus aureus. *Journal of Clinical Microbiology*, **48**, 4598-4601, 2010. R10122.

[429] Malhorta-Kumar, S., Cortinas Abrahantas, J., Sabiiti, W., Lammens, C., Vercauteren, G., Ieven, M. Molenberghs, G., Aerts, M., Goossens, H., on behalf of the MOSAR XP2 Study Team, Evaluation of chromogenetic media for detection of methicillin-reisistant Staphylococcus anreaus. *Journal of Clinical Microbiology*, **48**, 1040-1046, 2010. R10153.

[430] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for comparing a test drug with standard of care in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 443-452, 2010. R10112.

[431] Mebis, J., Jansens, H., Minalu, G., Molenberghs, G., Schroyens, W.A., Gadisseur, A.P., Van de Velde, A., Vrelust, I., Goossens, H., and Z.N. Berneman, Long-term epidemiology of bacterial susceptibility profiles in adults suffering from febrile neutropenia with hematologic malignancy after antibiotic change. *Infection and Drug Resistance*, **3**, 53-61, 2010. R10128.

[432] Meesen, R., Dendale, P., Cuypers, K., Berger, J., Hermand, A., Thijs, H. and O. Levin, Neuromuscular electrical stimulation as a possible mean to prevent muscle tissue wasting in artificially ventilated and sedated patients at the intensive care unit: a pilot study. *Neuromodulation*, **13** (4), 315-321, 2010. R10176.

[433] Molenberghs, G. and M.G. Kenward, Semi-parametric marginal models for hierarchical data and their corresponding full models. *Computational Statistics and Data Analysis*, **54**, 585-597, 2010. R10121.

[434] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11012.

[435] Molenberghs, G., Beunckens, C., Jansen, I., Thijs, H., Verbeke, G. and M.G. Kenward, Donnes manquantes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta. (Eds.)* Paris: Editions Technip., 2010. R10180.

[436] Molenberghs, G., Burzykowski, T., Alonso, A., Assam, P., Tilahun, A., and M. Buyse, A unified framework for the evaluation of surrogate endpoints in clinical trials. *Statistical Methods in Medical Research*, **19**, 205-236, 2010. R10115.

[437] Molenberghs, G., Book Review of Falagas, M.E. (2009). Meta-analysis in Infectious Diseases, An Issue of Infectious Disease Clinics (The Clinics: Internal Medicine), Published by WB Saunders, 240 p. *HIV Clinical Trials*, **11**, 121-123, 2010. R10139.

[438] Molenberghs, G., Editorial on "Generalized estimating equations: notes on the choice of the working correlation matrix (A. Ziegler, M. Vens)". *Methods of Information in Medicine*, **49**, 419-420, 2010. R10138.

[439] Molenberghs, G., Incomplete data in clinical and epidemiological studies. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10133.

[440] Molenberghs, G., Incomplete data in clinical and epidemiological studiesl. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10129.

[441] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, 21, 187-206, 2011. R11013.

[442] Molenberghs, G., Linear Mixed Model (version 3). *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10135.

[443] Molenberghs, G., Linear Mixed Model. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10131.

[444] Molenberghs, G., Meta-analysis in Infectious Diseases, An Issue of Infectious Disease Clinics (The Clinics: Internal Medicine). *HIV Clinical Trials*, **11**, 121-123, 2010. R10244.

[445] Molenberghs, G., Repeated Measures. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10134.

[446] Molenberghs, G., Repeated Measures. *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010. R10130.

[447] Molenberghs, G., Sensitivity Analysis. *International Encyclopedia of Statistical Sciences*, 2010. R10132.

[448] Molenberghs, G., Sensitivity Analysis. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*, 2010. R10136.

[449] Molenberghs, G., The applied statistical scientist in a high-profile academic environment. In: Data and context in statistical education. Towards an evidence-based society. *Proceedings of the 8th International Conference on Teaching Statistics, C. Reading (ed.)* July 12-16, Ljubljana, Slovenia (6p.), 2010. R10193.

[450] Poleto, F.Z., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. In: Proceedings of the 25th International Workshop on Statistical Modeling. *Glasgow, United Kingdom, A. Bowman (ed.)* 445-448, 2010. R10192.

[451] Pramana, S., Lin, D., Haldermans, P., Shkedy, Z., Verbeke, T., Ghlmann, H., De Bondt, A., Talloen, W., and L. Bijnens, ISoGene: an R package for analyzing dose-response studies in microarray experiments. *The R Journal*, **2/1**, 5-12, 2010. R10160.

[452] Pryseley, A., Mintiens, K., Knapen, K., Van der Stede, Y., and G. Molenberghs, Estimating precision, repeatability, and reproducibility from Gaussian and non-Gaussian data: a mixed models approach. *Journal of Applied Statistics*, **37**, 1729-1739, 2010. R10108.

[453] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011.

[454] Rizopoulos, D., Verbeke, G., and G. Molenberghs, Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival data. *Biometrics*, **66**, 20-29, 2010. R10123.

[455] Sanchez, L., Cortinas Abrahantas, J., Pelaez, O., Gutierrez, H., Concepcion, D. and P. Van der Stuyft, Breteau index threshold levels indicating risk for dengue transmission in areas with low aedes infestation. *Tropical Medicine and International Health*, **15** (2), 173-175, 2010. R10156.

[456] Scientific Report of the Working Group on Left Censored Data on the Management of left-censored data in dietary exposure assessment of chemical substances1. *The EFSA Journal*, **8** (3), 1557, 1-96, 2010. R10189.

[457] Scientific Report on the Analysis of the baseline survey on the prevalence of Campylobacter in broiler batches and of Campylobacter and Salmonella prevalence estimates. *The EFSA Journal*, **8** (3), 1503, 1-99 , 2010. R10190.

[458] Scientific Report on the Analysis of the baseline survey on the prevalence of Campylobacter in broiler batches and of Campylobacter and Salmonella on broiler carcasses, in the EU, Part B: Analysis of factors associated with Campylobacter colonization of broiler batches and with Campylobacter contamination of broiler carcasses; and investigation of the culture method diagnostic characteristics used to analyse broiler carcass s amples. *The EFSA Journal*, **8** (8), 1522, 1-132, (Marc Aerts), 2010. R10191.

[459] Serroyen, J., Bruckers, L., Rogiers, G., and G. Molenberghs, Characterizing persistent disturbing behavior using longitudinal and multivariate techniques. *Journal of Applied Statistics*, **37**, 341-355, 2010. R10120.

[460] Serroyen, J., Molenberghs, G., Aerts, M., Vloeberghs, E., De Deyn, P.P., and G. Verbeke, Flexible estimation of serial correlation in linear mixed models. *Journal of Applied Statistics*, **37**, 833-846, 2010. R10114.

[461] Sotto, C., Beunckens, C., Molenberghs, G., and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2010. R10127.

[462] Standaert, A., Van Holderbeke, M., Carnelis, C., Van Deun, J., Van Campenhout, K., Van Gestel, G., Wildemeersch, D., Nelen, V., Bruckers, L. and R. Torfs, Modeling human exposure to cadmium and arsenic in a zinc smelter area-A Belgian case study. *Epidemiology*, **20**, 553-554, 2010. R10204.

[463] Talloen, W., Hochreiter, S., Bijnens, L., Kasim, A., Shkedy, Z., Amaratunga, D. and H.W.H. Goehlmann, Filtering data from high-throughput experiments based in measurement reliability. *Proceedings of the National Academy of Science USA*, **107**, 46, E173-E174, 2010. R10170.

[464] Theeten H., Hutse, V., Hens, N. Yavuz Y., Hoppenbrouwers, K., Beutels, P. Vranckx, R. and P. Van Damme, Are we hitting ummunity targets? The 2006 age specific seroprevalence of measles, mumps, rubella, diphtheria and tetanus in Belgium. *Epidemiology and Infection*, 1-11, 2010. R10149.

[465] Theeten, H., Hens, N., Aerts, M. Vandermeulen, C., Roelants, M., Hoppenbrouwers, K., Van Damme, P. and P. Beutels, Caregivers' willigness to pay to reduce the number of vaccine injections in infants. *Pediatric Infectious Disease Journal*, **23**, 61-63, 2010. R10151.

[466] The GASTRIC (Global Advanced/Adjuvant Stomach Tumor Research through International Collaboration) Group, Benifit of adjuvant chemotherapy for respectable gastric cancer: A meta-analysis of individual patient data from 17 randomized trials. *Journal of the American Medical Association*, **303**, 1729-1737, 2010. R10162.

[467] Tilahun, A., Lin, D., Shkedy, Z., Geys, H., Alonso, A., Peeters, P. Talloen, W., Drinkenburg, W., Goehlmann, H., Gorden, E., Bijnens, L. and G. Molenberghs, Genomic biomarkers for depression feature-specific and joint biomarkers. *Statistics in Biopharmaceutical Research*, **2**, 3, 419-434, 2010. R10168.

[468] Troxel, A., Lipsitz, S., Fitzmaurice, G., Ibrahim, J., Sinha, D., and G. Molenberghs, A weighted combination of pseudo-likelihood estimators for longitudinal binary data subject to non-ignorable non-monotone missingness. *Statistics in Medicine*, **29**, 1511-1521, 2010. R10117.

[469] Vanden Boer, L., Bronselaer, J., Declercq, A., Demaerschalck, M. and G. Molenberghs, Het belang van omgevingsfactoren voor beleidsgericht ouderenonderzoek. *Tijdschrift voor Welzijnswerk*, **34**, 5-11, 2010. R10187.

[470] Vanden Boer, L., Bronselaer, J., Declercq, A., Demaerschalck, M. and G. Molenberghs, Het belang van omgevingsfactoren voor beleidsgericht ouderenonderzoek. *Tijdschrift voor Welzijnswerk*, **34**, 5-11, 2010. R10187.

[471] Van De Putte, G., Vlasselaer, J., Teshome, B., Gaddah, A., Burzykowski, T., Schobbens, J.C., Benijts, G. and E.T.M. De Jonge, Outcome of per protocol best-evidence based routine breast cancer care in a large regional hospital in Belgium: the importance of a prospective database in quality assurance. *Facts, Views and Vision in Obstetrics and Gynaecology*, **2**, 119-124, 2010. R10163.

[472] Vankrunkelsven, P., Kellen, E., Lousbergh, D., Cloes, E., Op de Beeck, L., Faes, C., Bruckers, L., Mertens, R. and F. Buntinx, De verkoop van producten voor hormonale substiteitherapie: evolutie van borstkankerincidentie tussen 1992 en 2008. *Tijdschrift voor Geneeskunde*, **66**, 25-30, 2010. R10184.

[473] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2010. R10116.

[474] Verbeke, G. and G. Molenberghs, Modles marginaux et modles mixtes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta (Eds.)* Paris: Editions Technip., 2010. R10181.

[475] Verbeke, G., Molenberghs, G. and D. Rizopoulos, Random effects models for longitudinal data. In: Longitudinal Research with Latent Variables. *K. van Montfort, H. Oud, and Al Satorra (Eds.)* New York: Springer, pp. 37-96, 2010. R10179.

[476] Van Wilder, Ph., Zhu, Q. and N. Veraverbeke, Identifying and rewarding pharmaceutical innovation: application in an EU member state. *Drug Information Journal*, **44**, 317-323, 2010. R10104.

[477] Vangeneugden, T., Molenberghs, G., Laenen, A., Geys, H., Beunckens, C., and C. Sotto, Marginal correlation in longitudinal binary data based on generalized linear mixed models. *Communications in Statistics, Theory Methods*, **39**, 2540-3557, 2010. R10109.

[478] Vannieuwenhuyze, J., Loosveldt, G. and G. Molenberghs, A method for evaluating mode effects in mixed mode surveys. *Public Opinion Quarterly*, **75** (5), 1027-1043, 2010. R10237.

[479] Veraverbeke, N., Nonparametric estimation of residual quantiles in a conditional Koziol-Green model with dependent censoring. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: a Festschrift in honor of Prof. Jana Jureckov. IMS Collections*, **7**, 245-253, 2010. R10105.

[480] Wouters, K., Cortinas Abrahantas, J., Molenberghs, G., Geys, H., Bijnens, L. Ahnaou, A. and W.H.I.M. Drinkenburg, Correction for model selection bias using a modified model averaging approach for supervised learning methods applied to EEG experiment. *Journal of Biopharmaceutical Statistics*, **20** (4), 768-786, 2010. R10157.

[481] Zhu, Q., Valkenborg, D. and T. Burzykowski, A Markov-chain-based heteroscedastic regression model for the analysis of high-resolution enzymatically 18O-labeled mass spectra. *Journal of Proteome Research*, **9** (5), 2669-2677, 2010. R10158.

3.5.3 Refereed publications (in press)

[482] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Information-theoretic approach to surrogate markers evaluation with failure time endpoints. *Lifetime Data Analysis*, 2010. RP10102.

[483] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple imputation based generalized estimating equations. *Journal of Biopharmaceutical Statistics*, 2010. RP10107.

[484] Creemers, A., Aerts, M., Hens, N., Shkedy, Z., Desmet, F. and P. Beutels, Revealing age-specific past and future unrelated costs of pneumococcal infections by flexible generalized estimation equations. *Journal of Applied Statistics*, 2010. RP10108.

[485] Cortinas, J., Sotto, C., Molenberghs, G., Vromman, G. and B. Bierinckx, A comparison of various software tools for dealing with missing data via imputation. *Journal of Statistical Computation and Simulation*, 2010. RP10100.

[486] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward, Shared-parameter models and missingness at random. *Statistical Modelling*, 2010. RP10109.

[487] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, 2010. RP10094.

[488] Goeyvaerts, N., Hens, N., Ogunjimi, B., Aerts, M., Shkedy, Z., Van Damme, P. and P. Beutels, Estimating infectious disease parameters from data on social contacts and serological status. *Journal of the Royal Statistical Society - Series C*, 2010. RP10048.

[489] Jacobs, N., Clays, E., De Bacquer, D., De Backer, G., Dendale, P., Thijs, H., de Bourdeaudhuij, I. and N. Claes, Effect of a tailored behavior change program on a composite lifestyle change score: a randomized controlled trial. *Submitted to Health Education Research*, 2010. RP10117.

[490] Jacobs, N., De Bourdeaudhuij, I., Thijs, H., Dendale, P. and N. Claes, Effect of a cardiovascular prevention program on health behavior and BMI in highly educated adults: a randomized controlled trial. *Patient Education and Couseling*, 2010. RP10095.

[491] Leuridan, E., Hens, N., Peeters, N., De Witte, L. and P. Van Damme, Effect of an adult pertussis booster dose on maternal antibody titers in young infants. *PIDJ*, In Press 2011. RP11001.

[492] Levin, O., Cuypers, K., Netz, T., Thijs, H. Nutting, B., Helsen, W. and R. Meesen, Age-related differences in human corticospinal excitability during simple reaction time. *Neuroscience Letters*, uncorrelated proof, available online 6 October 2010) 2010. RP10106.

[493] Lin, D., Tilahun, A., Cortinas Abrahantes, J., Shkhedy, Z., Molenberghs, G., Talloen, W., Goehlmann, H.W.H. and L. Bijnens, Comparison of methods for the selection of genomic biomarkers. *International Journal of Data Mining and Bioinformatics*, 2010. RP10116.

[494] Litire, S., Alonso, A., and G. Molenberghs, Correction. *Statistics in Medicine*,29, 2010. RP10097.

[495] Litire, S., Alonso, A. and G. Molenberghs, Rejoinder to reader reaction: Type I and type II error random-effects misspecification in generalized linear mixed models. *Biometrics*, 2010. RP10099.

[496] Luyten, L., Marais, C., Hens, N., De Schrijver, K. and P. Beutels, Imputing QALYs from single time point health state descriptions on the EQ-5D and the SF-6D: a comparison of methods for hepatitis A patients. *Value in Health*, 2010. RP10110.

[497] Molenberghs, G., Verbeke, G., Demtrio, C.G.B. and A. Vieira, A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, 2010. RP10098.

[498] Mwambi, H., Ramroop, S., Shkedy, Z. and G. Molenberghs, An approach to estimating the force of infection and the rate of recovery from Respiratory Syncytial Virus (RSV) data. *Statistical Methods in Medical Research*, 2010. RP10069.

[499] Ogunjimi, B., Smits, E., Hens, N., Hens, A., Lenders, K., Ieven, M., Van Tendeloo, V., Van Damme, P. and P. Beutels, Exploring the impact of exposure to primary varicella in children on varicella-zoster virus immunity of parents. *Viral Immunology*, 2010. RP10111.

[500] Poleto, F., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. *Test*, 2010. RP10105.

[501] Pramana, S., Lin, D., Haldermans, P., Shkedy, Z. and T. Verbeke, Analysis of dose-response studies in microarray experiments using the R IsoGene package. *Journal of Statistical Software*, 2010. RP10115.

[502] Pramana, S., Shkedy, Z., Ghlmann, H., Talloen, W., De Bondft, A., Straetemans, R., Lin, D. and J. Pinheiro, Model-based parametric approaches, in Lin, D., Shhedy, Z., Yekutieli, D., Amaratunga, D., and Bijnens, L. editors. Modeling dose-response microarray data in early drug development experiments using R. Springer, 2011. RP11002.

[503] Pramana, S., Haldermans, P. and T. Verbeke, A user friendly interface for analyzing dose-response studies in microarray experiments: the IsoGene-GUI package. In Lin, D., Shkedy, Z., Yekutieli, D., Amaratunga, D. and Bijnens, L. editors. Modeling dose-response microarray data in early drug development experiment using R. Springer, 2011. RP11003.

[504] Roca Pardias, J., Cadarso Surez, C.M., Pardo Vzquez, J.L., Leborn, V., Molenberghs, G., Faes, C. and C. Acua, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*, 2011. RP11005.

[505] Sinha, S.K., Troxel, A.B., Lipsitz, S.R., Sinha, D., Fitzmaurice, G.M., Molenberghs, G. and J.G. Ibrahim, A bivariate pseudo-likelihood for incomplete longitudinal binary data with nonignorable non-monotone missingness. *Biometrics* 66, 2010. RP10104.

[506] Tomsin, K., Mesens, T., Molenberghs, G., Peeters, L. and W. Gyselaers, Time-interval between maternal electrocardiogram and venous Doppler waves in normal pregnancy and pre-eclampsia: a pilot study. *European Journal of Ultrasound*, 2010. RP10101.

[507] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demtrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, 2010. RP10096.

[508] Van Sanden, S., Shkedy, Z., Burzykowski, T., Ghlmann, H.W., Talloen, W. and L. Bijnens, Genomic biomarkers for a binary clinical outcome in early drug development microarray experiments. *Journal of Biopharmaceutical Statistics*, 2010. RP10113.

[509] Van Sanden, S. and T. Burzykowski, Evaluation of Laplace distribution-based ANOVA models applied to microarray data. *Journal of Applied Statistics* , 2010. RP10114.

[510] Veraverbeke, N., Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scadinavian Journal of Statistics*, 2010. RP10093.

[511] Vieira, A.M.C., Leandro, R.A., Demtrio, C.G.B. and G. Molenberghs, Double generalized linear model for tissue culture proportion data: A Bayesian perspective. *Journal of Applied Statistics*, 2010. RP10103.

3.6 Université Joseph Fourier, UJF–LMC–IMAG

3.6.1 Technical reports

[512] Antoniadis, A., Brossat, X, Cugliari, J. and J.M. Poggi, Multivariate Functional Time Series Clustering, 2011. TR11008.

[513] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11007.

[514] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in additive models using P-splines, 2010. TR10064.

[515] Antoniadis, A., Helbert, C., Prieur, C. and L. Viry, Spatio-temporal prediction for West Africa monsoon, 2010. TR10062.

[516] Boshnakov, G. and S. Lambert-Lacroix, Periodic Levinson-Durbin algorithm for entropy maximisation, 2010. TR10063.

[517] Lambert-Lacroix, S. and L. Zwald, Robust regression through the Huber's criterion and adaptive lasso penalty, 2011. TR11005.

[518] Lambert-Lacroix, S. and F. Letué, Partial Least Squares and Cox model with application to gene expression, 2011. TR11006.

3.6.2 Refereed publications (published)

[519] Antoniadis, A., Bigot, J. and S. Lambert-Lacroix. Peaks detection and alignment for MASS spectrometry data. *Journal de la Socit Franaise de Statistique*, **151**, 1, 17-37, 2010. R10267.

[520] Antoniadis, A., Brossat, X., Cugliari, J. and J.M. Poggi. Clustering Functional data with wavelets. *e-book of COMPSTAT 2010*, 697-704, 2010. R10269.

[521] Antoniadis, A., Comments on L1-Penalization for Mixture Regression Models. *Test*, **19**, 2, 257-258, 2010. R10266.

[522] Antoniadis, A., Letu, F. and P. Fryzlewicz, The Dantzig selector for regression with right censored data. *Scandinavian Journal of Statistics*, **37**, 4, 531-552, 2010. R10274.

[523] Amato, U., Antoniadis, A., Samarov, A. and A. Tsybakov. Noisy Independent Factor Analysis Model for Density Estimation and Classification, *Electronic Journal of Statistics*, **4**, 706-736, 2010. R10268.

[524] Aseervatham, S., Antoniadis, A., Burlet, M., Denneulin, Y. and E. Gaussier, A sparse version of the ridge logistic regression for large-scale text categorization. *Pattern Recognition Letters*. **32**, 2, 101-106, 2010. R11014.

[525] Brouste, A, Istas, J. and S. Lambert-Lacroix, On simulation of manifold indexed fractional Gaussian fields. *Journal of Statistical Software*, **4**, 36, 1-14, 2010. R10270.

[526] Cao, R., Antoniadis, Anestis, Garca-Magarios, M. and W. Gonzlez-Manteiga, Lasso Logistic Regression, GSoft and the Cyclic Coordinate Descent Algorithm. *Statistical Applications in Genetics and Molecular Biology*,**9**, 1, 30, 2010. R10257.

[527] Garcia, N., Galves, A. and C. Prieur, Perfect simulation of the $\bar{d}$-minimal coupling between ordered pairs of binary chains of infinite order. *Journal of Statistical Physics*, **141**, 669–682. R10271.

[528] Guillotin-Plantard, N. and C. Prieur, Central limit theorem for sampled sums of dependent random variables. *ESAIM P&S*. **14**, 299-314, 2010. R10273.

[529] Guillotin-Plantard, N. and C. Prieur, Limit theorem for random walk in weakly dependent random scenery. *Annales de l'Institut Henri Poincar P&S*, **46**, 4, 1178-1194, 2010. R10272.

3.6.3 Refereed publications (in press)

[530] Antoniadis, A., Gijbels, I. and M. Nikolova, Penalized likelihood regression for generalized linear models with nonquadratic penalties. *Annals of the Institute of Mathematical Statistics*, 2010. RP11006.

3.7 Erasmus Medical Center, EMC

3.7.1 Refereed publications (published)

[531] Bloemberg, T. G., Gerretzen, J., Wouters, H.J.P., Gloerich, J., van Dael, M., Wessels, H.J.C.T., van den Heuvel, L.P., Eilers, P.H.C., Buydens, L.M.C. and R. Wehrens, Improved parametric time warping for proteomics. *Chemometrics And Intelligent Laboratory Systems*, **104**, 65-74, 2010. R10245.

[532] Dejardin, D., Lesaffre, E. and G. Verbeke, Joint modelling of progression-free survival and death in advanced cancer clinical trials. *Statistics in Medicine*, **29**, 1724-1734, 2010. R10228.

[533] Diya, L., Lesaffre, E., Van den Heede, K., Sermeus, W. and A. Vleugels, Establishing the relationship between nurse staffing and hospital mortality using a clustered discrete-time logistic model. *Statistics in medicine*, **29** (7-8), 778-785, 2010. R10208.

[534] Garcia-Zattera, M., Mutsvari, T., Jara, A., Declerck, D. and E. Lesaffre, Correcting for misclassification for a monotone disease process with an application in dental research. *Statistics in Medicine*, **29** (30), 3103-3117, 2010. R10238.

[535] Ghidey, W., Lesaffre E. and G. Verbeke, A comparison of methods for estimating the random effects distribution of a linear mixed model. *Statistical Methods in Medical Research*, **19**, 575-600, 2010. R10209.

[536] Hartendorp, M.O., Van der Stigchel, S., Burnett, H.G., Jellema, T, Eilers, P.H.C. and A. Postma, Categorical perception of morphed objects using a free-naming experiment. *Visual Cognition*, **18**, 9, 1320-1347, 2010. R10246.

[537] Komarek, A., Hansen B.E., Kuiper E.M., van Buuren H.R. and E. Lesaffre, Discriminant analysis using a multivariate linear mixed model with a normal mixture in the random effects distribution. *Statistics in Medicine*, **29**, 3267-3283, 2010. R10206.

[538] Marx, B.D., Eilers, P.H.C., Gampe, J. and R. Rau, Bilinear modulation models for seasonal tables of counts. *Statistics and Computing*, **20**, 2, 191-202, 2010. R10248.

[539] Mutsvari, T., Lesaffre, E., Garcia-Zattera, M., Declerck, D. and L. Diya, Factors that influence data quality in caries experience detection: a multi-level modeling approach. *Caries Research*, **44** (5), 438-444, 2010. R10211.

[540] Perperoglou, A. and P.H.C. Eilers, Penalized regression with individual deviance effects. *Computational Statistics*, **25**, 9, 2, 341-361, 2010. R10247.

[541] Rizopoulos, D., Verbeke, G., and G. Molenberghs, Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival data. *Biometrics*, **66**, 20-29, 2010. R10123.

[542] Tsonaka, R., Rizopoulos, D., Verbeke, G. and E. Lesaffre, Non-ignorable models for intermittently missing categorical longitudinal responses. *Biometrics*, **66 (3),** , 834-844, 2010. R10214.

[543] Van den Heede, K., Simoens, S., Diya, L., Lesaffre, E., Vleugels, A. and W. Sermeus, Increasing nurse staffing levels in Belgian cardiac surgery centres: a cost-effective patient safety intervention?. *Journal of Advanced Nursing*, **66** (6), 1291-1296, 2010. R10239.

[544] van Eijk, R., Eilers, P.H.C., Natte R., Cleton-Jansen, A.M., Morreau, H., van Wezel, T. and J. Oosting, MLPAinter for MLPA interpretation: an integrated approach for the analysis, visualisation and data management of Multiplex Ligation-dependent Probe Amplification. *BMC Bioinformatics*, **11**, 67, 2010. R10032.

[545] Vanhaecht, K., Bellemans, J., De Witte, K., Diya, L., Lesaffre, E. and W. Sermeus, Does the organization of care processes affect outcomes in patients undergoing total joint replacement?. *Journal of Evaluation in Clinical Practice*, **16** (1), 121-128, 2010. R10222.

[546] Verbeke, G., Molenberghs, G. and D. Rizopoulos, Random effects models for longitudinal data. In: Longitudinal Research with Latent Variables. *K. van Montfort, H. Oud, and Al Satorra (Eds.)* New York: Springer, pp. 37-96, 2010. R10179.

[547] Zhang Y., Todem D., Kim K., and E. Lesaffre, Bayesian latent variable models for spatially correlated tooth-level binary data in caries research. *Statistical Modelling*, **11**, 25-47, 2010. R10207.

3.7.2 Refereed publications (in press)

[548] Agbaje, J., Mutsvari, T., Lesaffre, E. and D. Declerck, Measurement, analysis and interpretation of examiner reliability in caries experience surveys: some methodological thoughts. *Clinical Oral Investigations*, 2010. RP10147.

[549] Verbeke, G. and E. Lesaffre, Repeated Measurements. *Encyclopedia of Statistical Sciences*, New York:. Wiley Sons, 2010. RP10146.

3.8 Universidad de Santiago de Compostela, USC

3.8.1 Technical reports

[550] Bianco, A., Boente, G., Gonzalez-Manteiga, W. and A. Perez Gonzalez, Asymptotic behaviour of robust estimators in partially linear models with missing responses: The effect of estimating the missing probability on the simplified marginal estimators, 2010. TR10061.

[551] Meira-Machado, L., Roca-Pardinas, J., Van Keilegom, I. and C. Cadarso-Suarez, Estimation of transition probabilities in a non-Markov model with successive survival times, 2010. TR10057.

[552] Monsalve-Cobis, Abelardo E., Gonzalez-Manteiga,W. and M. Febrero-Bande, Goodness of Fit Test for Interest Rate Models: an approach based on Empirical Process, 2010. TR10060.

3.8.2 Refereed publications (published)

[553] Cao, R., Antoniadis, Anestis, Garca-Magarios, M. and W. Gonzlez-Manteiga, Lasso Logistic Regression, GSoft and the Cyclic Coordinate Descent Algorithm. *Statistical Applications in Genetics and Molecular Biology*,**9**, 1, 30, 2010. R10257.

[554] Crujeiras, R. and I. Van Keilegom, Least squares estimation of nonlinear spatial trends. *Computational Statistics and Data Analysis*, **54**, 452-465, 2010. R10004.

[555] Crujeiras, R., Comments on: A general science-based framework for nonlinear spatio-temporal dynamical models. *TEST*, **19**, 456-458, 2010. R10265.

[556] Crujeiras, R., Fernndez-Casal, R. and W. Gonzlez-Manteiga, Goodness-of-fit tests for the spatial spectral density. *Stochastic Environmental Research and Risk Assesment*, **24**, 67-79, 2010. R10014.

[557] Crujeiras, R., Fernndez-Casal, R., Gonzlez-Manteiga, W., Goodness-of-fit tests for the spatial spectral density. *Stochastic Environmental Research and Risk Assesment*, **24**, 67-79, 2010. R10250.

[558] Crujeiras, R., Fernndez-Casal, R. and W. Gonzlez-Manteiga, Nonparametric test for separability of spatio-temporal processes. *Environmetrics*, **21**, 382-399, 2010. R10251.

[559] Crujeiras, R., Moreira, C. and M.J. Bento, Mapping childhood cancer incidence at northern region of Portugal. *Advances and Applications in Statistical Sciences*, **3**, 177-194, 2010. R10264.

[560] de Ua lvarez, J., Rodrguez-Casal, A. and H.Y. Liang, Nonlinear wavelet estimator of the regression function under left truncated dependent data. *Journal of Nonparametric Statistics*, **22**, 319-344, 2010. R10261.

[561] Faraldo Roca, P., Crujeiras, R., Lombarda, M.J., Naya Fernndez, S., Matas, J.M., Carreira V. and B.M. Fernndez de Castro, *Estudo da insercin laboral dos titulados no Sistema Universitario de Galicia 2006-2007*. Axencia para a Calidade do Sistema Universitario de Galicia (ACSUG), 2010. R10048.

[562] Febrero-Bande, M., Galeano, P. and W. Gonzlez-Manteiga, Measures of influence for the functional linear model with scalar response. *Journal of Multivariate Analysis*, **101**, 327-339, 2010. R10015.

[563] Fernndez-Casal, R. and R. Crujeiras, Spatial dependence estimation using FFT of biased covariances. *Journal of Statistical Planning and Inference*, **140**, 2653-2668, 2010. R10263.

[564] Gonzlez-Manteiga, W. and A. Martnez Calvo, *Bootstrap Calibration in Functional Linear Regression Models with Applications*. Proceedings of COMPSTAT'2010, 199-207. Lechevallier, Yves; Saporta, Gilbert (Eds.). Springer Physica-Verlag Heidelberg, 2010. R10203.

[565] Gonzlez-Manteiga, W. and A. Martnez Calvo, Comments on: Dynamic relations for sparsely sampled Gaussian processes. *TEST*, **19**, 43-45, 2010. R10258.

[566] Gonzlez-Manteiga, W., Mateu, J., Crujeiras, R. and E. Porcu, On the second order properties of the multidimensional peridogram for regularly spaced data. *Theory of Probability and its Applications. SIAM*, **53**, 349-356, 2009. R09215.

[567] Gonzlez-Manteiga, W., Ordoez, C., Taboada Castro, J. and J.M. Matas, Partially linear support vector machines applied to the prediction of mine slope movements. *Mathematical and Computer Modelling*, **51**, 206-215, 2010. R10005.

[568] Muoz Bars, Jose Ignacio, Mara Sol Rodrguez-Calvo, Surez-Pearanda Jos M, Duarte N Vieira, Cadarso Surez, Carmen Mara, Febrero-Bande, M., PMICALC: An R code-based software for estimating postmortem interval (PMI) compatible with Windows, Mac and Linux operating systems. *Forensic Science International*, **194**, (1), 49-52, 2010. R10254.

[569] Pateiro-Lpez, B. and A. Rodrguez-Casal, Generalizing the Convex Hull of a Sample: The R Package alphahull. *Journal of Statistical Software*, **22**, (5), 1-28, 2010. R10260.

[570] Reboredo, J.C., Gonzlez-Manteiga, W., Febrero-Bande, M. and J.M. Matas, Boosting GARCH and neural networks for the prediction of heteroskedastic time series. *Mathematical and Computer Modelling*, **51**, 256-271, 2010. R10003.

[571] Reboredo, J.C., Gonzlez-Manteiga, W. and G. Aneiros Prez, Semiparametric regression with a FARIMA-GARCH error process: Theory and application. *Advances and Applications in Statistical Sciences*, **3**, 83-112, 2010. R10253.

[572] Reboredo, J.C., Gonzlez-Manteiga, W., Febrero-Bande, M. and J.M. Matas, Boosting GARCH and neural networks for the prediction of heteroskedastic time series. *Mathematical and Computer Modelling*, **51**, 256-271, 2010. R10255.

[573] Rodrguez-Casal, A. and J.E. Chacn Durn, A note on the universal consitency of the kernel distribution function estimator. *Statistics and Probability Letters*, **80**, 1414-1419, 2010. R10262.

[574] Sperlich, S. and M.J. Lombarda, Local polynomial inference for small area statistics: estimation, validation and prediction. *Journal of Nonparametric Statistics*, **22**, (5), 633-648, 2010. R10259.

3.8.3 Refereed publications (in press)

[575] Berrendero, J.R., Cuevas, A., Pateiro-Lpez, B., A multivariate uniformity test for the case of unknown support. *Statistics and Computing*, 2010. RP10124.

[576] Bianco, A., Boente, G., Gonzlez-Manteiga, W. and A. Prez Gonzlez, Asymptotic behaviour of robust estimators in partially linear models with missing responses: The effect of estimating the missing probability on the simplified marginal estimators. *TEST*, 2010. RP10128.

[577] Bianco, A., Boente, G., Gonzlez-Manteiga, W. and A. Prez Gonzlez, Estimation of the marginal location under a partially linear model with missing responses. *Computational Statistics and Data Analysis*, 2010. RP10054.

[578] Cadarso Surez, Carmen Mara, Hervada Vidal, Xurxo, Rodrguez Muos, Miguel ngel, Seoane Pillado, Teresa, Free Software in the Medical Investigation. *Medicina Clnica*, 2010. RP10123.

[579] Chernick, M.R., Gonzlez-Manteiga, W., Crujeiras, R. and B.E. Barrios, Bootstrap methods. *International Encyclopedia of Statistical Science.*, 2010. RP10053.

[580] Crujeiras, R. and R. Fernndez-Casal, On the estimation of the spectral density for continuous spatial processes. *Statistics*, 2010. RP10050.

[581] Cuesta-Albertos, JA, Febrero-Bande, M., A simple multiway ANOVA for functional data. *TEST*, 2010. RP10120.

[582] Ferraty F, Gonzlez-Manteiga, W., Martnez Calvo, Adela, Vieu, Philippe. Presmoothing in functional linear regression. *Statistica Sinica*, 2010. RP10122.

[583] F Gude, K.T., Meira-Machado, L. and C.M. Cadarso Surez, Flexible hazard ratio curves for continuous predictors in multi-state models: An application to breast cancer. *Statistical Modelling*, 2010. RP10131.

[584] Galeano, P. and D. Pea, *An Unified Approach to Model Selection, Discrimination, Goodness of Fit and Outliers in Time Series*. Advances in Mathematical and Statistical Modeling, 2010. RP10001.

[585] Gonzlez-Manteiga, W., Lombarda, M. J., Molina Perala, I., Morales Gonzlez, D. and L. Santamara, Small area estimation under Fay-Herriot models with nonparametric estimation of heteroscedasticity. *Statistical Modelling*, 2010. RP10051.

[586] González-Manteiga, W. and C.A. Martnez, Comments on : Dynamic relations for sparsely sampled Gaussian processes. *TEST*, 2010. RP10056.

[587] Gonzlez-Manteiga, W. and R. Crujeiras, *A general view of the goodness-of-fit tests for statistical models*. Modern Mathematical Tools and Techniques in Capturing Complexity, 2010. RP10003.

[588] González-Manteiga, W., Pardo Fernández, J.C. and I. Van Keilegom, ROC curves in nonparametric location-scale regression models. *Scandinavian Journal of Statistics*, 2010. RP10009.

[589] Martnez Calvo, A. and W. Gonzlez-Manteiga, Bootstrap in functional linear regression. *Journal of Statistical Planning and Inference*, **141**, 453-461, 2010. RP10130.

[590] Moreira, C., de Ua lvarez, Jacobo, Crujeiras, R., DTDA, an R package to analyze randomly truncated data. *Journal of Statistical Software*, 2010. RP10118.

[591] MX Rodrguez-lvarez, Roca Pardias, Javier, Cadarso Surez, Carmen Mara, ROC curve and covariates: extending the induced methodology to the non-parametric framework. *Statistics and Computing*, 2010. RP10132.

[592] MX Rodrguez-lvarez, Garcia-Tahoces, Pablo, Cadarso Surez, Carmen Mara, Lado, M. Jos, Comparative study of ROC regression techniques. Applications for the computer-aided diagnostic system in breast cancer detection. *Computational Statistis and Data Analysis*, 2010. RP10126.

[593] Prez Gonzlez, Ana, Vilar Fernndez, Juan Manuel, Gonzlez-Manteiga, W., Nonparametric variance function estimation with missing data. *Journal of Multivariate Analysis*, **101**, (5), 1123-1142, 2010. RP10127.

[594] Reyes Santas F, MX Rodrguez-lvarez, Cadarso Surez, Carmen Mara, Estimating hospital production functions through flexible. *Mathematical and Computer Modelling*, 2010. RP10121.

[595] Roca Pardias, J., Cadarso Surez, C.M., Pardo Vzquez, J.L., Leborn, V., Molenberghs, G., Faes, C. and C. Acua, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*, 2011. RP11005.

[596] Ruiz-Medina, M.D., Crujeiras, R., Minimum contrast parameter estimation for fractal random fields based on the wavelet periodogram. *Communications in Statistics: Theory and Methods*, 2010. RP10119.

[597] Vicente Lustres-Prez, MX Rodrguez-lvarez, Maria P. Pata, Fernndez Pulpeiro E., Cadarso Surez, Carmen Mara, The application of Receiver Operating Characteristic (ROC) methodology in biological studies of marine resources: sex determination of Paracentrotus lividus (Lamarck, 1816). *SORT*, 2010. RP10129.

3.9 London School of Hygiene and Tropical Medicine, LSHTM

3.9.1 Technical reports

[598] Carpenter, J.R., Roger, J.H. and M.G. Kenward, Relevant, accessible sensitivity analyses using multiple imputation, 2010. TR10003.

3.9.2 Refereed publications (published)

[599] Burzykowski, T., Carpenter, J., Coens, C., Evans, D., France, L. Kenward, M. and Lane, P., Matcham, J. Morgan, D., Philips, A., Roger, J., Sullivan, B., White, I., Yu, L.M., of the PSI missing data expert group, Missing data: discussion points from the PSI missing data expert Group. *Pharmaceutical Statistics*, **6**, 2010. R10164.

[600] Carpenter, J.R. and I. Plewis, Coming to terms with non-response in longitudinal studies. *SAGE handbook of Methodological Innovation*, M. Williams and P. Vogt (eds.), London, SAGE, 2010. R10006.

[601] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G., and M.G. Kenward, A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes. *Biometrical Journal*, **52**, 111-125, 2010. R10124.

[602] Molenberghs, G. and M.G. Kenward, Semi-parametric marginal models for hierarchical data and their corresponding full models. *Computational Statistics and Data Analysis*, **54**, 585-597, 2010. R10121.

[603] Molenberghs, G., Beunckens, C., Jansen, I., Thijs, H., Verbeke, G. and M.G. Kenward, Donnes manquantes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta. (Eds.)* Paris: Editions Technip., 2010. R10180.

[604] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, 21, 187-206, 2011. R11013.

[605] Nur, U., Shack, L.G., Rachet, B., Carpenter, J.R. and M.P. Coleman, Modelling relative survival in the presence of incomplete data: a tutorial. *International Journal of Epidemiology*, **39**, 118-128, 2010. R10007.

[606] Skene, S. and M.G. Kenward, The analysis of very small samples of repeated measurements. I: An adjusted sandwich estimator. *Statistics in Medicine*, **29**, 2825-2837, 2010. R10276.

[607] Skene, S. and M.G. Kenward, The analysis of very small samples of repeated measurements II: A modified Box correction. *Statistics in Medicine*, **29**, 2838-2856, 2010. R10277.

[608] Sotto, C., Beunckens, C., Molenberghs, G., and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2010. R10127.

[609] Spratt, M., Carpenter, J.C., Sterne, J.A.C., Carlin, J.B., Heron, J., Henderson, J. and K. Tilling, Strategies for Multiple Imputation in Longitudinal Studies. *Journal of Epidemiology*, **172**, 478-487, 2010. R10278.

[610] Vansteelandt, S., Carpenter, J.R. and M.G. Kenward, Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology*, **6**, 37-48, 2010. R10079.

[611] White, I.R., Daniel, R. and P. Royston, Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables, Modelling. *Computational Statistics and Data Analysis*, **54**, 2267-2275, 2010. R10275.

3.9.3 Refereed publications (in press)

[612] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple imputation based generalized estimating equations. *Journal of Biopharmaceutical Statistics*, 2010. RP10107.

[613] Carpenter, J.R., Goldstein, H. and M.G. Kenward, REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. *Journal of Statistical Software*, 2010. RP10070.

[614] Carpenter, J.C., Rücker, G. and G. Schwarzer, Sensitivity of Meta-analysis to Selection Bias: A Multiple Imputation Approach. *Biometrics*, 2010. RP10133.

[615] Daniel, R.M., Kenward, M.G., Cousens, S. and B. de Stavola, Using directed acyclic graphs to guide analysis in missing data problems. *Statistical Methods in Medical Research*, 2010. RP10134.

[616] Kenward, M.G. and G. Rosenkranz, Joint modelling of outcome, observation time and missingness. *Journal of Biopharmaceutical Statistics*, 2010. RP11004.

[617] Lee, Y., Mangseok, N. and M.G. Kenward, Joint modelling of outcome, Robust estimation of dropout models using hierarchical likelihood. *Journal of Statistical Computation and Simulation*, 2010. RP10135.

[618] Rücker, G., Schwarzer, G., Carpenter, J.R., Binder, H. and Schumacher, M., Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistic*, 2010. RP10136.

[619] Schwarzer, G. Carpenter, J.R. and G. Rucker, Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis. *Journal of Clinical Epidemiology*, 2010. RP10059.

[620] Youngjo, L., Manegseok, N. and M.G. Kenward, Robust estimation of dropout models using hierarchical likelihood. *Journal of Statistical Computation and Simulation*, 2010. RP10060.

3.10 List of joint publications

3.10.1 Technical reports

[621] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in additive models using P-splines, 2010. TR10064. (KUL-1, UJF)

[622] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11007.

[623] Meira-Machado, L., Roca-Pardinas, J., Van Keilegom, I. and C. Cadarso-Suarez, Estimation of transition probabilities in a non-Markov model with successive survival times, 2010. TR10057. (UCL, USC)

[624] Omelka, M., Veraverbeke, N. and I. Gijbels, Bootstrapping the conditional copula, 2010. TR10072. (UH, KUL-1)

[625] Van Keilegom, I. and N. Veraverbeke, Statistical models and methods for dependence in insurance data, 2011. TR11003. (UCL, UH)

3.10.2 Refereed publications (published)

[626] Aerts, M., Faes, C., Hens, N. and G. Molenberghs, Handling missingness and informative cluster sizes when modelling prevalence and force of infection. *Joint Statistical Meetings 2010 Proceedings*, 3491-3501, 2010. R10194. (KUL-2, UH)

[627] Alonso, A., Litire, S. and A. Laenen, A Note on the indeterminacy of the random-effects distribution in hierarchical models. *American Statistician*, **64** (4), 318-324, 2010. R10235. (KUL-2, UH)

[628] Alonso, A., Litire, S. and G. Molenberghs, Testing for misspecification in generalized linear mixed models. *Biostatistics*, **11**, 771-786, 2010. R10125. (KUL-2, UH)

[629] Alonso, A., Laenen, A., Molenberghs, G., Geys, H. and T. Vangeneugden, Reliability of Single Administered Tests: A Unified Approach. *Biometrics*, **66** (4), 1061-1068, 2010. R10236. (KUL-2, UH)

[630] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Using earlier measures in a longitudinal sequence as potential surrogate for a later one. *Computational Statistics and Data Analysis*, **54**, 1342-1354, 2010. R10137. (KUL-2, UH)

[631] Bloemberg, T. G., Gerretzen, J., Wouters, H.J.P., Gloerich, J., van Dael, M., Wessels, H.J.C.T., van den Heuvel, L.P., Eilers, P.H.C., Buydens, L.M.C. and R. Wehrens, Improved parametric time warping for proteomics. *Chemometrics And Intelligent Laboratory Systems*, **104**, 65-74, 2010. R10245. (UCL, EMC)

[632] Bruckers, L., Serroyen, J., Molenberghs, G., Slaets, H., and W. Goeyvaerts, Latent-class analysis of persistent disturbing behavior patients using longitudinal profiles. *Applied Statistics*, **59**, 495-512, 2010. R10119. (KUL-2, UH)

[633] Burzykowski, T., Carpenter, J., Coens, C., Evans, D., France, L. Kenward, M. and Lane, P., Matcham, J. Morgan, D., Philips, A., Roger, J., Sullivan, B., White, I., Yu, L.M., of the PSI missing data expert group, Missing data: discussion points from the PSI missing data expert Group. *Pharmaceutical Statistics*, **6**, 2010. R10164. (LSHTM, UH)

[634] Cao, R., Antoniadis, Anestis, Garca-Magarios, M. and W. Gonzlez-Manteiga, Lasso Logistic Regression, GSoft and the Cyclic Coordinate Descent Algorithm. *Statistical Applications in Genetics and Molecular Biology,***9**, 1, 30, 2010. R10257. (UJF, USC)

[635] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G., and M.G. Kenward, A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes. *Biometrical Journal*, **52**, 111-125, 2010. R10124. (KUL-2, UH, LSHTM)

[636] Creemers, A., Hens, N., Aerts, M. Molenberghs, G. Verbeke, G. and M.G. Kenward, A sensitivity analysis for shared-parameter models. *The Biometrical Journal*, **52**, 111-125, 2010. R10142. (KUL-2, UH, LSHTM)

[637] Crujeiras, R. and I. Van Keilegom, Least squares estimation of nonlinear spatial trends. *Computational Statistics and Data Analysis*, **54**, 452-465, 2010. R10004. (UCL, USC)

[638] Diya, L., Lesaffre, E., Van den Heede, K., Sermeus, W. and A. Vleugels, Establishing the relationship between nurse staffing and hospital mortality using a clustered discrete-time logistic model. *Statistics in medicine*, **29** (7-8), 778-785, 2010. R10208. (EMC, KUL-2)

[639] Dejardin, D., Lesaffre, E. and G. Verbeke, Joint modelling of progression-free survival and death in advanced cancer clinical trials. *Statistics in Medicine*, **29**, 1724-1734, 2010. R10228. (EMC, KUL-2)

[640] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, A principal. *International*, **6**, Article 20, 2010. R10095. (UG, KUL-2)

[641] Gaddah, A. and R. Braekers, Testing under the extended Koziol-Green model. Durante, F., Hardle, W., Jaworski, P., Rychlik, T. (Eds.) Workshop on Copula Theory and its applications, Lectures Notes in Statistics 198. *Proceedings Springer, Berlin/Heidelberg*, 2010. R10195. (UCL, UH)

[642] Gaddah, A. and R. Braekers, Testing under the Koziol-Green model, Copula theory and its applications. Proceedings of the Workshop held in Warsaw 25-26 September 2009, ed. Piotr Jaworski, Fabrizio Durante, Wolfgang Karl Hrdle and Tomasz Rychlik. *Lecture Notes in Statistics* **198**, 2, 279-288, Springer, 2010. R10202. (UCL, UH)

[643] Garcia-Zattera, M., Mutsvari, T., Jara, A., Declerck, D. and E. Lesaffre, Correcting for misclassification for a monotone disease process with an application in dental research. *Statistics in Medicine*, **29** (30), 3103-3117, 2010. R10238. (EMC, KUL-2)

[644] Ghidey W., Lesaffre E. and G. Verbeke, A comparison of methods for estimating the random effects distribution of a linear mixed model. *Statistical Methods in Medical Research*, **19**, 575-600, 2010. R10209. (EMC, KUL-2)

[645] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919-1932, 2011. R11010. (UH, KUL-1)

[646] Girma, M.A., Hens, N., Aets, M., Coenen, S., Uller, A., Adriaenssens, N., Beutels, P., Molenberghs, G. and H. Goossens, In Bowman A.C., editor. Time series and mixed models to study the country-specic outpatient antibiotic use in Europe. Proceedings of the 25th Workshop on Statistical Modelling, July 5-9, 2010 Glasgow. *Scotland, U.K.*, 53-56, 2010. R10201. (UH, KUL-2)

[647] Goegebeur, Y., De Boeck, P. and G. Molenberghs, Person fit for test speededness: normal curvatures, likelihood ratio tests and empirical Bayes estimates. *Methodology: European journal of research methods for the behavioral and social sciences*, **6**, 3-16, 2010. R10064. (UH, KUL-1, KUL-2)

[648] Gyselaers, W., Molenberghs, G., Hanssens, M., and L. Peeters, Maternal renal interlobar vein impedance index is higher in early- than in late-onset preeclampsia. *Ultrasound in Obstetrics and Gynecology*, **36**, 69-75, 2010. R10113. (UH, KUL-2)

[649] Hens, N., Van Ranst, M., Aerts, M., Robesyn, E., Van Damme, P. and P. Beutels, Estimating the effective reproduction number for pandemic influenza from notification data made publicly available in real time: a multi-country analysis for influenza A/H1N1v 2009. *Vaccine*, 2010. R10148. (UH, KUL-2)

[650] Jacobs, T., Straetemans, R., Molenberghs, G., Bouwknecht, A., and L. Bijnens, A latent pharmacokinetic time profile to model dose-response survival data. *Journal of Biopharmaceutical Statistics*, **20**, 759-767, 2010. R10118. (UH, KUL-2)

[651] Jansen, I. and G. Molenberghs, Pattern-mixture models for categorical outcomes with non-monotone missingness. *Journal of Statistical Computation and Simulation*, **80**, 1279-1296, 2010. R10107. (UH, KUL-2)

[652] Komarek A., Hansen B.E., Kuiper E.M., van Buuren H.R. and E. Lesaffre, Discriminant analysis using a multivariate linear mixed model with a normal mixture in the random effects distribution. *Statistics in Medicine*, **29**, 3267-3283, 2010. R10206. (EMC, KUL-2)

[653] Laenen, A. and A. Alonso, The Functional Living Index - Cancer: Estimating its Reliability based on Clinical Trial Data. *Quality of Life Research*, **19**, 103-109, 2010. R10232. (UH, KUL-2)

[654] Laenen, A., Alonso, A., Molenberghs, G., Vangeneugden, T. and C. Mallinckrodt, Coping with memory effect and serial correlation when estimating reliability in a longitudinal framework. *Applied Psychological Measurement*, **34** (4), 255-266, 2010. R10234. (UH, KUL-2)

[655] Lin, D., Shkedy, Z., Cortinas Abrahantes, J., Tilahun, A., Molenberghs, G., Talloen, W., Goehlmann, H., and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiments. *Online Journal of Bioinformatics*, **11**, 106-127, 2010. R10126. (UH, KUL-2)

[656] Lin, D., Shkedy, Z., Molenberghs, G., Talloen, W., Goehlmann, H.W.H. and L. Bijnens, Selection and evaluation of gene-specific biomarkers in pre-clinical and clinical microarray experiment. *Online Journal of Bioinformatics*, **11** (1), 106-127, 2010. R10167. (UH, KUL-2)

[657] Little, R.J.A., D'Agostino, R., Dickersin, K., Emerson, S.S., Farrar, J.T., Frangakis, C., Hogan, J.W., Molenberghs, G., Murphy, S.A., Neaton, J.D., Rotnitzky, A., Scharfstein, D., Shih, W., Siegel, J.P. and H. Stern, The prevention and treatment of missing data in clinical trials. Panel on handling Missing Data in Clinical Trials. *Washington, D.C.: The National Academies Press*, 2010. R10205. (UH, KUL-2)

[658] Malhorta-Kumar, S., Cortinas Abrahantas, J., Sabiiti, W., Lammens, C., Vercauteren, G., Ieven, M. Molenberghs, G., Aerts, M., Goossens, H., on behalf of the MOSAR XP2 Study Team, Evaluation of chromogenetic media for detection of methicillin-reisistant Staphylococcus anreaus. *Journal of Clinical Microbiology*, **48**, 1040-1046, 2010. R10153. (UH, KUL-2)

[659] Malhotra-Kumar, S., Van Heirstraeten, L., Lee, A., Cortinas Abrahantes, J., Lammens, C., Vanhommerig, E., Molenberghs, G., Aerts, M., Harbart, S., Goossens, H., on behalf of the MOSAR WP2 Study Team, Evaluation of molecular assays for rapid detection of methicillin-resistant Staphylococcus aureus. *Journal of Clinical Microbiology*, **48**, 4598-4601, 2010. R10122. (UH, KUL-2)

[660] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Design archetypes for phase II clinical trials in central nervous system disorders. *Drug Information Journal*, **44**, 421-430, 2010. R10110. (UH, KUL-2)

[661] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for using positive controls in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 431-441, 2010. R10111. (UH, KUL-2)

[662] Mallinckrodt, C.H., Detke, M.J., Prucka, W.R., Ruberg, S.J., and G. Molenberghs, Considerations for comparing a test drug with standard of care in phase II clinical trials of central nervous system disorders. *Drug Information Journal*, **44**, 443-452, 2010. R10112. (UH, KUL-2)

[663] Mebis, J., Jansens, H., Minalu, G., Molenberghs, G., Schroyens, W.A., Gadisseur, A.P., Van de Velde, A., Vrelust, I., Goossens, H., and Z.N. Berneman, Long-term epidemiology of bacterial susceptibility profiles in adults suffering from febrile neutropenia with hematologic malignancy after antibiotic change. *Infection and Drug Resistance*, **3**, 53-61, 2010. R10128. (UH, KUL-2)

[664] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value

distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11012. (UH, KUL-2)

[665] Molenberghs, G. and M.G. Kenward, Semi-parametric marginal models for hierarchical data and their corresponding full models. *Computational Statistics and Data Analysis*, **54**, 585-597, 2010. R10121. (UH, KUL-2, LSHTM)

[666] Molenberghs, G., Beunckens, C., Jansen, I., Thijs, H., Verbeke, G. and M.G. Kenward, Donnes manquantes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta. (Eds.)* Paris: Editions Technip., 2010. R10180. (UH, KUL-2, LSHTM)

[667] Molenberghs, G., Burzykowski, T., Alonso, A., Assam, P., Tilahun, A., and M. Buyse, A unified framework for the evaluation of surrogate endpoints in clinical trials. *Statistical Methods in Medical Research*, **19**, 205-236, 2010. R10115. (UH, KUL-2)

[668] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, 21, 187-206, 2011. R11013. (UH, KUL-2, LSHTM)

[669] Mutsvari, T., Lesaffre, E., Garcia-Zattera, M., Declerck, D. and L. Diya, Factors that influence data quality in caries experience detection: a multi-level modeling approach. *Caries Research*, **44** (5), 438-444, 2010. R10211. (EMC, KUL-2)

[670] Poleto, F.Z., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. In: Proceedings of the 25th International Workshop on Statistical Modeling. *Glasgow, United Kingdom, A. Bowman (ed.)* 445-448, 2010. R10192. (UH, KUL-2)

[671] Pramana, S., Lin, D., Haldermans, P., Shkedy, Z., Verbeke, T., Ghlmann, H., De Bondt, A., Talloen, W., and L. Bijnens, ISoGene: an R package for analyzing dose-response studies in microarray experiments. *The R Journal*, **2/1**, 5-12, 2010. R10160. (UH, KUL-2)

[672] Pryseley, A., Mintiens, K., Knapen, K., Van der Stede, Y., and G. Molenberghs, Estimating precision, repeatability, and reproducibility from Gaussian and non-Gaussian data: a mixed models approach. *Journal of Applied Statistics*, **37**, 1729-1739, 2010. R10108. (UH, KUL-2)

[673] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011. (UH, KUL-2)

[674] Rizopoulos, D., Verbeke, G., and G. Molenberghs, Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival data. *Biometrics*, **66**, 20-29, 2010. R10123. (UH, KUL-2)

[675] Serroyen, J., Bruckers, L., Rogiers, G., and G. Molenberghs, Characterizing persistent disturbing behavior using longitudinal and multivariate techniques. *Journal of Applied Statistics*, **37**, 341-355, 2010. R10120. (UH, KUL-2)

[676] Serroyen, J., Molenberghs, G., Aerts, M., Vloeberghs, E., De Deyn, P.P., and G. Verbeke, Flexible estimation of serial correlation in linear mixed models. *Journal of Applied Statistics*, **37**, 833-846, 2010. R10114. (UH, KUL-2)

[677] Sotto, C., Beunckens, C., Molenberghs, G., and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2010. R10127. (UH, KUL-2, LSHTM)

[678] Tilahun, A., Lin, D., Shkedy, Z., Geys, H., Alonso, A., Peeters, P. Talloen, W., Drinkenburg, W., Goehlmann, H., Gorden, E., Bijnens, L. and G. Molenberghs, Genomic biomarkers for depression feature-specific and joint biomarkers. *Statistics in Biopharmaceutical Research*, **2**, 3, 419-434, 2010. R10168. (UH, KUL-2)

[679] Troxel, A., Lipsitz, S., Fitzmaurice, G., Ibrahim, J., Sinha, D., and G. Molenberghs, A weighted combination of pseudo-likelihood estimators for longitudinal binary data subject to non-ignorable non-monotone missingness. *Statistics in Medicine*, **29**, 1511-1521, 2010. R10117. (UH, KUL-2)

[680] Tsonaka, R., Rizopoulos, D., Verbeke, G. and E. Lesaffre, Non-ignorable models for intermittently missing categorical longitudinal responses. *Biometrics*, **66 (3),** , 834-844, 2010. R10214. (EMC, KUL-2)

[681] Vandekerckhove, J., Verheyen, S. and F. Tuerlinckx, A crossed random effects diffusion model for speeded semantic categorization decisions. *Acta Psychologica*, **133**, 269-282, 2010. R10069. (KUL-1, KUL-2)

[682] Vanden Boer, L., Bronselaer, J., Declercq, A., Demaerschalck, M. and G. Molenberghs, Het belang van omgevingsfactoren voor beleidsgericht ouderenonderzoek. *Tijdschrift voor Welzijnswerk*, **34**, 5-11, 2010. R10187. (UH, KUL-2)

[683] Van den Heede, K., Simoens, S., Diya, L., Lesaffre, E., Vleugels, A. and W. Sermeus, Increasing nurse staffing levels in Belgian cardiac surgery centres: a cost-effective patient safety intervention?. *Journal of Advanced Nursing*, **66** (6), 1291-1296, 2010. R10239. (EMC, KUL-2)

[684] Van De Putte, G., Vlasselaer, J., Teshome, B., Gaddah, A., Burzykowski, T., Schobbens, J.C., Benijts, G. and E.T.M. De Jonge, Outcome of per protocol best-evidence based routine breast cancer care in a large regional hospital in Belgium: the importance of a prospective database in quality assurance. *Facts, Views and Vision in Obstetrics and Gynaecology*, **2**, 119-124, 2010. R10163. (UH, UCL)

[685] Vangeneugden, T., Molenberghs, G., Laenen, A., Geys, H., Beunckens, C., and C. Sotto, Marginal correlation in longitudinal binary data based on generalized linear mixed models. *Communications in Statistics, Theory Methods*, **39**, 2540-3557, 2010. R10109. (UH, KUL-2)

[686] Vanhaecht, K., Bellemans, J., De Witte, K., Diya, L., Lesaffre, E. and W. Sermeus, Does the organization of care processes affect outcomes in patients undergoing total joint replacement?. *Journal of Evaluation in Clinical Practice*, **16** (1), 121-128, 2010. R10222. (EMC, KUL-2)

[687] Vannieuwenhuyze, J., Loosveldt, G. and G. Molenberghs, A method for evaluating mode effects in mixed mode surveys. *Public Opinion Quarterly*, **75** (5), 1027-1043, 2010. R10237. (UH, KUL-2)

[688] Vansteelandt, S., Carpenter, J.R. and M.G. Kenward, Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology*, **6**, 37-48, 2010. R10079. (UG, LSHTM)

[689] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2010. R10116. (UH, KUL-2)

[690] Verbeke, G. and G. Molenberghs, Modles marginaux et modles mixtes. In: Analyse statistique des donnes longitudinales. *J.J. Droesbeke and G. Saporta (Eds.)* Paris: Editions Technip., 2010. R10181. (UH, KUL-2)

[691] Verbeke, G., Molenberghs, G. and D. Rizopoulos, Random effects models for longitudinal data. In: Longitudinal Research with Latent Variables. *K. van Montfort, H. Oud, and Al Satorra (Eds.)* New York: Springer, pp. 37-96, 2010. R10179. (UH, KUL-2)

[692] Wouters, K., Cortinas Abrahantas, J., Molenberghs, G., Geys, H., Bijnens, L. Ahnaou, A. and W.H.I.M. Drinkenburg, Correction for model selection bias using a modified model averaging approach for supervised learning methods applied to EEG experiment. *Journal of Biopharmaceutical Statistics*, **20** (4), 768-786, 2010. R10157. (UH, KUL-2)

[693] Zhang Y., Todem D., Kim K., and E. Lesaffre, Bayesian latent variable models for spatially correlated tooth-level binary data in caries research. *Statistical Modelling*, **11**, 25-47, 2010. R10207. (EMC, KUL-2)

3.10.3 Refereed publications (in press)

[694] Agbaje, J., Mutsvari, T., Lesaffre, E. and D. Declerck, Measurement, analysis and interpretation of examiner reliability in caries experience surveys: some methodological thoughts. *Clinical Oral Investigations*, 2010. RP10147. (EMC, KUL-2)

[695] Antoniadis, A., Gijbels, I. and M. Nikolova, Penalized likelihood regression for generalized linear models with nonquadratic penalties. *The Annals of the Institute of Statistical Mathematics*, RP11006. (KUL-1, UJF)

[696] Assam, P., Tilahun, A., Alonso, A. and G. Molenberghs, Information-theoretic approach to surrogate markers evaluation with failure time endpoints. *Lifetime Data Analysis*, 2010. RP10102. (UH, KUL-2)

[697] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple imputation based generalized estimating equations. *Journal of Biopharmaceutical Statistics*, 2010. RP10107. (UH, KUL-2, LSHTM)

[698] Cortinas, J., Sotto, C., Molenberghs, G., Vromman, G. and B. Bierinckx, A comparison of various software tools for dealing with missing data via imputation. *Journal of Statistical Computation and Simulation*, 2010. RP10100. (UH, KUL-2)

[699] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward, Shared-parameter models and missingness at random. *Statistical Modelling*, 2010. RP10109. (UH, KUL-2, LSHTM)

[700] De Wolf, J., Duchateau, L., Verbeke, G. and E. Schrevens, Discovering transgenic events: using information from early screening trials for improving experimental design. *Biological and Environmental Statistics*, 2010. RP10142. (UG, KUL-2)

[701] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, 2010. RP10094. (UH, UCL)

[702] González-Manteiga, W., Pardo Fernández, J.C. and I. Van Keilegom, ROC curves in nonparametric location-scale regression models. *Scandinavian Journal of Statistics*, 2010. RP10009. (UCL, USC)

[703] Litire, S., Alonso, A. and G. Molenberghs, Correction. *Statistics in Medicine*,29, 2010. RP10097. (UH, KUL-2)

[704] Litire, S., Alonso, A. and G. Molenberghs, Rejoinder to reader reaction: Type I and type II error random-effects misspecification in generalized linear mixed models. *Biometrics*, 2010. RP10099. (UH, KUL-2)

[705] Molenberghs, G., Verbeke, G., Demtrio, C.G.B. and A. Vieira, A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, 2010. RP10098. (UH, KUL-2)

[706] Mwambi, H., Ramroop, S., Shkedy, Z. and G. Molenberghs, An approach to estimating the force of infection and the rate of recovery from Respiratory Syncytial Virus (RSV) data. *Statistical Methods in Medical Research*, 2010. RP10069. (UH, KUL-2)

[707] Poleto, F., Molenberghs, G., Paulino, C.D. and J.M. Singer, Sensitivity analysis for incomplete continuous data. *Test*, 2010. RP10105. (UH, KUL-2)

[708] Roca Pardias, J., Cadarso Surez, C.M., Pardo Vzquez, J.L., Leborn, V., Molenberghs, G., Faes, C. and C. Acua, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*, 2011. RP11005. (UH, KUL-2, USC)

[709] Sinha, S.K., Troxel, A.B., Lipsitz, S.R., Sinha, D., Fitzmaurice, G.M., Molenberghs, G. and J.G. Ibrahim, A bivariate pseudo-likelihood for incomplete longitudinal binary data with nonignorable non-monotone missingness. *Biometrics* 66, 2010. RP10104. (UH, KUL-2)

[710] Tomsin, K., Mesens, T., Molenberghs, G., Peeters, L. and W. Gyselaers, Time-interval between maternal electrocardiogram and venous Doppler waves in normal pregnancy and preeclampsia: a pilot study. *European Journal of Ultrasound*, 2010. RP10101. (UH, KUL-2)

[711] Vandekerckhove, J., Tuerlinckx, F. and M. Lee, Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 2010. RP10143. (KUL-1, KUL-2)

[712] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demtrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, 2010. RP10096. (UH, KUL-2)

[713] Veraverbeke, N., Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scadinavian Journal of Statistics*, 2010. RP10093. (UH, KUL-1)

[714] Verbeke, G. and E. Lesaffre, Repeated Measurements. *Encyclopedia of Statistical Sciences*, New York:. Wiley Sons, 2010. RP10146. (EMC, KUL-2)

[715] Vieira, A.M.C., Leandro, R.A., Demtrio, C.G.B. and G. Molenberghs, Double generalized linear model for tissue culture proportion data: A Bayesian perspective. *Journal of Applied Statistics*, 2010. RP10103. (UH, KUL-2)