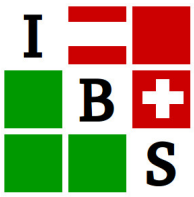
UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA
FOR EXPO 2015

JOINT MEETING
of the International Biometric Society (IBS)
Austro-Swiss and Italian Regions
15-19 June 2015

www.ibs-roes.org/iroes-2015/



JOINT MEETING of the International Biometric Society (IBS) Austro-Swiss and Italian Regions

Contents

Overview

Time		Session Title	Chairs	Session Room
Tuesday, 16 June 2015	09:00 - 09:30	Welcome Session	Held, Leonhard; Valsecchi, Maria Grazia	Aula Martini
	09:30 - 11:00	Does Size Really not Matter? Evaluation of Treatment Effects in Subgroups and Other Small Populations	Posch, Martin	Aula Martini
		Recent Advances in Health Impact Evaluation		U6-A10
	11:30 - 12:50	Flexible and Adaptive Designs	Senn, Stephen	Aula Martini
		Methods in Epidemiology	Decarli, Adriano	U6-A10
		Methods and Applications in Genetics/Omics (1)	Finos, Livio	U6-A11
	14:00 - 15:30	Statistical Methods for Omics Data Integration	Di Serio, Clelia	Aula Martini
		Innovative Study Designs and Analysis Issues in Epidemiology	Valsecchi, Maria Grazia	U6-A10
	16:00 - 17:20	Clinical Trials	Schmidli, Heinz	U6-A11
		Regression	Alfo', Marco	U6-A10
		Survival Analysis (1)	Ambrogi, Federico	Aula Martini
	18:00 - 19:30	Poster Reception		
Wednesday, 17 June 2015	09:00 - 10:30	Evidence Synthesis and the Use of Co-Data (CEN Invited Session)	Held, Leonhard; Friede, Tim	U6-A10
		Recent Advances in ROC Methodology (Invited Session of the Italian-Spanish-ERM Regions)	Antolini, Laura; Reiser, Benjamin	Aula Martini
	11:00 - 12:00	Statistical Methods in Infectious Disease	Frigessi, Arnoldo	U6-A11
		Methods and Applications in Genetics/Omics (2)	van Wieringen, Wessel N.	Aula Martini
		Analysis of Registries and Administrative Databases	Baccini, Michela	U6-A10
	12:15 - 13:30	General Assembly of IBS Italy	Valsecchi, Maria Grazia	U6-A10
		ROeS General Assembly	Held, Leonhard	U6-A11
Thursday, 18 June 2015	09:30 - 11:00	Robust Methods in Biostatistics	Farcomeni, Alessio	U6-A10
		Recent Developments to Tackle Time-Dependent Covariates in Clinical Research	Mittlboeck, Martina	Aula Martini
	11:30 - 12:50	Other Topics	Friede, Tim	U6-A11
		Survival Analysis (2)	Schemper, Michael	Aula Martini
		Young Statistician Session	Berghold, Andrea	U6-A10
	14:00 - 15:30	Statistical Methods in Environmental Sciences, Agriculture and Forestry	Moder, Karl; Castrignanò, Annamaria	U6-A10
		Adaptive Clinical Trials with Subpopulation Selection	Heinzmann, Dominik; Rufibach, Kaspar	Aula Martini
	16:00 - 17:00	Diagnostic Studies and Meta-Analysis	Copas, John Brian	U6-A11
		Missing Data Analysis	Rousson, Valentin	Aula Martini
		Studies in Veterinary, Agricultural and Environmental Research	Nothdurft, Arne	U6-A10
Friday, 19 June 2015	09:30 - 12:30	Talks	Decarli, Adriano	Napoleonic Hall
	14:00 - 16:00	Satellite Meeting - Round Table	Ulmer, Hanno	Napoleonic Hall

Session Details

Welcome Session

Tuesday, 16 June 2015
09:00 - 09:30
Room: Aula Martini
Session chair: Held, Leonhard; Valsecchi, Maria Grazia

Welcome from the Dean of the University Milano-Bicocca, Prof. Maria Cristina Messa

Does Size Really not Matter? Evaluation of Treatment Effects in Subgroups and Other Small Populations

Invited session
Tuesday, 16 June 2015
09:30 - 11:00
Room: Aula Martini
Session chair: Posch, Martin

Discussant: Willi Maurer

Friede, Tim : ***Identifying subgroups of treatment responders: A meta-analytic approach***
Author list: *Friede, Tim*

The identification of subgroups of patients responding particularly well to a treatment has become of increased interest in a move to personalized (or stratified) medicine. We start by introducing methods for the identification of subgroups within a single trial. Then we extend these ideas to the setting of several trials using a meta-analytic framework. We close by discussing how subgroup identification and confirmation can be integrated in clinical development plans.

Senn, Stephen : ***Depersonalising medicine***
Author list: *Senn, Stephen*

I argue that the quest for personalised medicine has the possible side-effect of damaging health-care. We may end up adding noise to the system.

The demand to prove efficacy in each and every subgroup may cause us to waste resources that would be better spent on looking for new treatments.

Statisticians need to go back to basics. As Deming showed, understanding the variation in the system is the duty of every manager. Those who do not know what is assignable and what is pure noise run the risk of intervening pointlessly in response to the latter. I suggest that regulators would do better to request demonstrations of larger mean effects rather than demanding proof of efficacy for everybody.

Rosenkranz, Gerd : ***Exploratory subgroup analyses in confirmatory trials***
Author list: *Rosenkranz, Gerd*

An upcoming "Guideline on the investigation of subgroups in confirmatory clinical trials" by the European Medicines Agency (EMA) states that "Investigation into the effects of treatment in well-defined subsets of the trial population is an integral part of clinical trial planning, analysis and inference that follows the inspection of the primary outcome of the trial." Though not explicitly mentioned in the title, the document is concerned with exploratory subgroup analyses, in particular it says that: "Most investigations will consider subgroups identified on the basis of a single factor. Subgroups defined on multiple factors (e.g. females aged >65) may be of interest on occasion but for simplicity, the descriptions in this document will make reference to a subgroup defined on a single factor (e.g. gender categorised as male and female), and this will suffice for most investigations."

The talk is proposing some methodology for subgroup analyses to address this request. The main idea is to fit a statistical model for each predefined factor including a term for the factor and it's interaction with treatment in the model. The interaction reflects the difference in effects of treatment for different factor levels. The model that fits best is determining the subgroup.

To account for selection bias, subgroup selection is repeated on datasets obtained by sampling with replacement from the original study data. The model selected most often in this process defines the subgroup and a bias corrected estimate of the interaction term can be obtained from the samples. The latter provides an indicator of relevance of the effect in a descriptive way.

The method can handle binary, (ordered) categorical and continuous data. Without the requirement to investigate subgroups, it could handle continuous factors without dichotomization and provide a predictive model for treatment effects based on covariates.

Recent Advances in Health Impact Evaluation

Invited session
Tuesday, 16 June 2015
09:30 - 11:00
Room: U6-A10
Session chair:

Discussant: Fabrizia Mealli

Mattei, Alessandra : ***Short term impact of PM10 exposure on mortality: A propensity score approach***
Author list: *Baccini, Michela; Mattei, Alessandra; Mealli, Fabrizia*

Exposure to air pollution is associated with short-term increase in mortality and the interest in this field has recently moved to health impact assessment. Impact is usually evaluated in terms of deaths attributable to air pollution levels exceeding pre-specified thresholds corresponding to different counterfactual scenarios. To this end, a two-step approach is typically implemented: first the exposure-response function is estimated through specification of a regression model on the daily number of events, which accounts for possible confounders, including seasonality, meteorological conditions and influenza epidemics; second the estimated exposure-response curve is combined with the observed number of events and the observed air pollution levels in order to evaluate the absolute excess of deaths under each specific scenario. In this work, we propose a new statistical causal approach based on propensity score matching methods. "Exposed" days (i.e., days with air pollution levels exceeding pre-specified thresholds) are matched with "unexposed" days having similar values of all observed confounders. Then the number of deaths attributable to air pollution is obtained comparing the number of deaths between matched days. We apply our approach to evaluate the short term impact of fine air-born particles (PM10) on mortality in the Italian city of Milan during the period 2003-2006. The results are compared to those obtained using the traditional approach (see Baccini et al. 2011), and advantages and disadvantages of the two methods are discussed.

Zell, Elizabeth : ***A potential outcomes approach to documenting the public health impact of the introduction of PCV13 for the prevention of invasive pneumococcal disease***
Author list: *Zell, Elizabeth*

Recommendations are often made to improve the health, education, well-being, etc. of individuals or communities. Estimating the impacts of such recommendations is important. Often costs are associated with their implementation and cost-benefit analyses should be considered. After a new recommendation has been implemented is NOT the time to "look for" baseline data from the pre-implementation period. Planning in advance is important to the evaluation process. Here we illustrate a potential outcomes approach to estimate IPD (Invasive Pneumococcal Disease) outcomes in the counterfactual world in which a new vaccine, PCV13, had not been introduced. This effort was possible because the CDC had baseline active, population-based IPD surveillance data prior to the introduction of PCV13. This approach allowed the assessment of disease reduction attributed to the new vaccine for those recommend for vaccination (direct effect) and for those not recommended for vaccination (indirect effect). Comparing the PCV13 rates in counterfactual world, estimated using time series models, to the rates that actually occurred estimates the number of IPD cases prevented due to the introduction of the vaccine.

Baccini, Michela : ***Managing different sources of uncertainty in long-term impact of heat on population health: the case of the city of Skopje***
Author list: *Baccini, Michela; Sanchez, Gerardo*

High ambient temperatures can affect population health in several ways, up to increasing mortality either directly or through exacerbation of pre-existing conditions. This evidence, coupled with the fact that climate change scenarios project rising temperatures and an increase in frequency and intensity of heat waves, confirms the need of developing analytical tools for predicting the future impact of heat on population health, so that appropriate mitigation or protection policies can be implemented. These methods should account for uncertainties related to the incomplete knowledge of the phenomenon, that usually arise from different sources and have different entities. This study attempts to describe the health effects of the change in heat exposure that the urban population of Skopje, the capital and largest city of the Republic

of Macedonia, could experience in two future time periods within the 21st century: the near future period 2026-2045 and the far future period 2081-2100. First, we focus on estimation of the heat-mortality relationship based on historical data, and on prediction of the future population size. We then calculate the impact of heat in terms of attributable deaths, by combining the previous results with the temperature projections arising from a climate model designed to analyze the urban heat island effect. We account for variability around the estimate of the heat-mortality relationship by using Monte Carlo simulations, while uncertainties related to population prediction and future weather conditions are treated by considering alternative population growth models and three different global climate models which exhibit respectively the lowest, the maximal and the median temperature for the far future period 2081-2100.

Flexible and Adaptive Designs

Contributed session
Tuesday, 16 June 2015
11:30 - 12:50
Room: Aula Martini
Session chair: Senn, Stephen

Rufibach, Kaspar : ***Evaluation of possible designs for a three-arm clinical trial: Comparing a closed-testing design to potential alternatives***
Author list: *Asikanius, Elina; Rufibach, Kaspar; Bahlo, Jasmin; Bieska, Gabriele; Burger, Hans-Ulrich*

To optimize resources, randomized clinical trials with multiple arms can be considered an attractive option to simultaneously testing various treatment regimens. Motivated by the success of a three-arm randomized clinical trial whose inference strategy was based on a closed testing procedure, we compare four different potential strategies to run a three arm clinical trial and quantify the differences in power and the time it takes until approval of the new drug for these designs. We discuss statistical and operational aspects of implementing a design using a closed testing procedure, and situations where it is of limited use. In conclusion, using a closed testing procedure is an innovative method for a multiple-arm clinical trial, may bring the new drug fastest to patients, the power loss due to the global test in a closed testing procedure is minimal, and - according to our assessment - was the ideal method to achieve the goals in the said trial in a time- and resource-efficient and innovative way.

Klinglmueller, Florian : ***Estimation after blinded sample size reassessment***
Author list: *Klinglmueller, Florian; Posch, Martin; Koenig, Franz; Miller, Frank*

We investigate the properties of point estimates and confidence intervals following blinded sample size reassessment. We show that the sample mean and variance may be biased. We provide upper bounds for the absolute bias and quantify it by simulation. We show that the coverage probabilities of confidence intervals may lie below their nominal level. If the first stage sample size is low or if general sample size rules are considered the bias can be substantial. We conclude that sample size adaptation rules should be prespecified and only reassessment rules with negligible impact on the properties of estimators should be used.

Scarale, Maria Giovanna : ***An intensive care unit trial on melatonin reinterpreted according to an urn model***
Author list: *Scarale, Maria Giovanna; Mistraletti, Giovanni; Mariani, Luigi; Ferraroni, Monica*

In randomized and controlled clinical trial there is an ethical imperative to provide the best available medical treatment according to updated guidelines. In recent years, adaptive designs have been proposed to achieve this goal even in testing new therapies, by the using of sequential accrual coupled with dynamically updated probability of treatment assignment.

This is the distinguishing feature of trials based on the Response Adaptive Randomization, in which the allocation probability to each of different investigated treatments is modified according to previous history of treatment assignments and observed responses.

Response adaptive trials typically rely on the urn model, originally formulated by Pòlya. Our work is focused on one of its available developments, namely the Randomly Reinforced Urn (RRU)[1], and demonstrates by simulation the superiority of this approach toward the classical urn model. The RRU-design is characterized by the ethically optimal property that the assignment probability to the superior treatment asymptotically converges to one. Moreover, the random reinforcement improves the convergence speed, thus assuring an economic advantage. From this viewpoint, the choice of the transformation function, capturing the reinforcement level, is crucial. This function can be expressed in various forms according to the specific utility function that is used: different utility functions U imply different properties for the RRU-design, in terms of convergence rate and skewness of allocations.

We will first present the clinical context in which a trial was designed to assess the efficacy of oral melatonin administration in Intensive Care Unit (ICU) patients in decreasing the need for sedation[2]. Through sleep-wake rhythm regularization induced by melatonin, the study hypothesis was to observe a reduction of administered sedative drugs made by blinded physicians in the experimental arm. The application of the

urn model in this trial could bring some benefits for patients and researchers, highlighting the possibility to have practical implications even in a more general perspective.

References

[1] Muliere P., Paganoni A.M. and Secchi P. (2006a), "A Randomly Reinforced Urn", Journal of Statistical Planning and Inference, 136, 1853-1874.

[2] Mistraletti G, et al. "Oral melatonin decreases need for sedatives and analgesics in critically ill", ESICM Congress, Berlin 2011.

Methods in Epidemiology

Contributed session
Tuesday, 16 June 2015
11:30 - 12:50
Room: U6-A10
Session chair: Decarli, Adriano

Rota, Matteo : ***Dichotomizing biomarkers for failure time outcome: from classification probabilities to predictive values based methods***
Author list: *Rota, Matteo; Antolini, Laura; Valsecchi, Maria Grazia*

The identification of cut-points for continuous biomarkers is often based on criteria defined on classification probabilities (sensitivity and specificity) [1] with reference to the development or not of disease up to a predefined time point t [2]. Since the classification probabilities are standardized with respect to the prevalence of disease, these methods lead to cut-points which lie near the median value of the biomarker when mimicking the case of a prevalence of about 50%. When the true prevalence of disease is far from this value, the predictive value of the obtained dichotomization may not be satisfactory. A change of perspective to methods directly based on predictive values instead of classification probabilities may be worth. In this context, flexible B-spline modeling in a Cox regression framework could offer great flexibility and may allow to identify optimal cut-points thorough optimal knots location [3], overcoming also the setting of the reference time t. Meaning and implication of these two methods are discussed through a motivating example on the cytokine receptor-like factor 2, as measured at diagnosis in 464 children with acute lymphoblastic leukemia, and its influence on the relapse rate [4]. While classification probabilities based methods led to the same unsatisfactory optimal cut-point, predictive value based methods led to an optimal cut-point lying on the boundary of the biomarker distribution.

REFERENCES

[1] Rota M, Antolini L. Finding the optimal cut-point for Gaussian and Gamma distributed biomarkers. Comput Stat Data Anal. 2014; 69:1-14.

[2] Rota M, Antolini L, Valsecchi MG. Optimal cut-point definition in biomarkers: the case of censored failure time outcome. BMC Med Res Methodol. [in press].

[3] Molinari N, Daurès JP, Durand JF. Regression splines for threshold selection in survival data analysis. Stat Med. 2001; 20:237-47.

[4] Palmi C, Vendramini E, Silvestri D et al. Poor prognosis for P2RY8-CRLF2 fusion but not for CRLF2 over-expression in children with intermediate risk B-cell precursor acute lymphoblastic leukemia. Leukemia 2012; 26:2245-53.

Mezzetti, Maura : ***Combining individual and aggregated data to investigate the role of socio-economic disparities on cancer burden in Italy***
Author list: *Mezzetti, Maura*

Quantifying socio-economic disparities and understanding the role of these factors in the occurrence of chronic diseases are growing public health topics. Several decades of research have documented ongoing social disparities in cancer incidence, mortality, and survival across Europe. Among Western countries, Italy provides an excellent opportunity for investigating variation in site-specific cancer occurrence over different areas. Italian regions are characterized by substantial differences in lifestyle and environmental factors, as well as by heterogeneous cancer incidence across regions. In addition, due to the regional organization of the national healthcare system, there exists a plurality of regional services, besides National health system provide equitable access to care. However how socio-economic differences affects cancer occurrence in Italy has not been investigated in depth, mainly due to lack of data (Materia et al., 2005).

The principal aim of this work is to take up this challenge, and to quantify the role of social disparities on different types of cancer in Italy. Specifically, a Bayesian hierarchical model is developed to combine three main sources of data: a cohort study with individual-level outcome and exposure data from 5 Italian provinces, cancer registry data providing aggregate-level information on disease frequencies, and a survey providing information on different risk factors distribution from a sample of individuals over different areas, but without any information on disease outcomes. The idea behind the model arises from Jackson at al

(2006), and the main novelty consists in using survey data to simulate information on individual-level risk factors for each area. The distributional assumptions in Jackson at al (2006) can be relaxed by adopting a fully Bayesian model.

The proposed methodology also allows the relationships between socio-economic variables and cancer risk to be controlled for spatial confounding.

References

Jackson, C., Best, N., and Richardson, S. (2006). Improving ecological inference using individual-level data. Statistics in Medicine, 25: 2136-2159.

Materia, E., Cacciani, L., Bugarini, G., and et al (2005). Income inequality and mortality in Italy. Eur J Public Health, 15(4): 411-417.

Di Maso, Matteo : ***A different class-assignment rule to build classification trees for ordinal outcomes***

Author list: *Di Maso, Matteo; Lugo, Alessandra*

Introduction

Classification and regression trees (CART) are binary recursive partitioning methods designed to construct prediction models for categorical (classification) or continuous (regression) variables from data. One of the key elements of classification trees is the assignment rule of each terminal node (leaf) to a class outcome.

Objectives

Evaluating the performance of the 'median trees', built with a novel approach to class assignment, compared to the 'modal trees', built with the majority rule, when an ordinal outcome (y) is assumed.

Materials and Methods

Modal trees were estimated using the modal class among the observations that fall into each leaf to assign y-classes, whereas median trees were estimated through the median class. According to the assignment rule adopted, the predicted power of the trees was evaluated by two different approaches: modal trees minimized the total number of errors; median trees minimized the sum of absolute distances between predicted class and observed class. Tree performances were evaluated through the gamma statistic, measuring the association between observed and predicted classes. Three real datasets with different number of y-levels (from four to six) were analyzed. Each dataset was divided into a training set, for building the trees, and a testing set, for evaluating prediction accuracy. A resampling of the testing set (n=30) was carried out to derive robust estimates. Binomial test and paired t-test were used to compare the significance of differences between tree performances.

Results

Median tree performances were significantly better than modal ones with five and six y-classes. Significant differences were not observed with four levels of the outcome. No matter of the number of y-classes, median trees showed a simpler structure (smaller number of leaves) than modal ones.

Conclusion

Median trees showed a better performance than modal trees with an increasing number of y-levels and generally provided a simpler structure which allows an easier interpretation of the patterns and connections among groups of interest.

Methods and Applications in Genetics/Omics (1)

Contributed session
Tuesday, 16 June 2015
11:30 - 12:50
Room: U6-A11
Session chair: Finos, Livio

Zebrowska, Magdalena : ***Different notions of false discovery rate under the logic regression model***

Author list: *Zebrowska, Magdalena; Frommlet, Florian*

Epistasis is a fundamental concept for understanding the genetic architecture underlying complex traits, as there is common agreement that the interaction between genes will typically play an important role. We consider Bateson's and Mendel's definition of "biological" epistasis, where interaction is extended to include changes in the phenotype (like height or disease status) which are caused by a Boolean combination of genotypes of several QTLs. The correct identification of this kind of interactions becomes extremely difficult when using a classical regression approach. In contrast the logic regression method proposed in [1] is specifically designed to identify such logic epistatic interactions. In our earlier work [2] we could demonstrate that if the phenotypic variation is caused by a Boolean combination of genotypes, then logic regression attains significantly larger power to detect these effects compared with the classical regression approach.

However, the evaluation of performance of logic regression models is not entirely straight forward. In simulation studies, where the underlying true model is known, the definition of true positive and false positive results is no longer without ambiguity. Due to the more complex form of predictors, which are logic expressions, specification of what is meant by false detection seems to become especially vague. Should logic combinations of variables from the true logic expression, different from the true combination, be treated as true or as false detections? And what if one gets a sub-expression of the true logic expression or an expression for which the true interaction is a sub-expression? Answers to these questions depend partly on the subjective opinion of the researcher. As a consequence also the notion of false discovery rate (FDR) must be reconsidered, which obviously depends on the exact specification of true and false positives.

[1] I. Ruczinski, C. Kooperberg, and M. LeBlanc "Logic regression.", J. Comput. Graphical Statist, 12(3):474–511, 2003.

[2] M. Malina, K. Ickstadt, H. Schwender, M. Posch, and M. Bogdan."Detection of epistatic effects with logic regression and a classical linear regression model."Statistical Applications in Genetics and Molecular Biology, 13(1):83–104, 2014.

Schmid, Volker J. : ***Statistics for nucleomics***

Author list: *Schmid, Volker J.*

A growing interest in the structural aspects of nuclear organization reflects the increasing awareness that understanding the functional nuclear organization in space and time cannot be gained solely from studies at the molecular level, but requires insight into the intimate structure-function relationships at higher levels of organization as well (and vice versa). Recent developements in superresolution microscopy allow to the nucleus as a global biological system in space and time in the context of its environment.

Imaging the nucleus in 3D/4D brings a lot of statistical challenges. We will discuss several challenges of 4D nucleomics, including image segmentation, for example to detect and classify the chromatin (that is, the DNA) in the nucleus, point localization detection of marked proteins, and colocalization of different proteins. We will also discuss the possibilities of using methods from spatial point processes, and the issues which arise in using those methods for three-dimensional images.

References:

Popken, J. et al.: Functional reprogramming of fibroblast nuclei in cloned bovine embryos is paralleled by major structural remodeling with both striking similarities and differences to nuclear phenotypes of embryos fertilized in vitro. Nucleus 5:6 (2014) 555-589.

Smeets, D.,et al.: Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. Epigenetics & Chromatin 7:8 (2014).

Markaki, Y. et al.: The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture. BioEssays 34:5 (2012) 412-426.

Seiler, D. et al.: Double-strand break-induced transcriptional silencing is associated with loss of trimethylation at H3K4. Chromosome Research 19:7 (2011) 883-899.

Edefonti, Valeria : ***Combinatorial mixtures of multiparameter distributions: an application to microarray data***

Author list: *Edefonti, Valeria; Parmigiani, Giovanni*

The term 'combinatorial mixtures' refers to a flexible class of models for inference on mixture distributions whose components have multidimensional parameters. The idea behind it is to allow each element of the component-specific parameter vector to be shared by a subset of other components.

We develop Bayesian inference and computational approaches for this class of mixture distributions with an unknown number of components. We define the structure for a general prior distribution – a mixture of prior distributions itself - where a positive probability is put on every possible combination of sharing patterns. This partial sharing allows for generality and flexibility in comparison with traditional approaches to mixture modeling, while still allowing to assign significant mass to models that are more parsimonious than the general 'no sharing' case.

We illustrate our combinatorial mixtures in an application based on the normal mixture model for any number of components. We introduce normal mixture models fpr univariate and bivariate data, which are amenable to Markov Chain Monte Carlo computing. In the light of combinatorial mixtures, we assume a decomposition of the variance-covariance matrix, which separates out standard deviations and correlations, and thus allows us to model those parameters separately. Moreover, to provide valid posterior estimates of the parameters, we introduce a novel solution to the well-known 'label switching' problem and we compare it with the existing ones.

This development was originally motivated by applications in molecular biology, where one deals with continuous measures, such as RNA levels, or protein levels, that vary across unknown biological subtypes. In some cases, subtypes are characterized by an increase in the level of the marker measured, while in others they are characterized by variability in otherwise tightly controlled processes, or by the presence of otherwise weak correlations. Also, several mechanisms can coexist. It may also allow to model an interesting phenomenon observed in microarray analysis when two variables have the same mean and variance but opposite correlations in diseased and normal samples. We use data on molecular classification of lung cancer from the web-based information supporting the published manuscript Garber et al. (2001).

Ahrens, Maike : ***Identification of patient subgroups in high-throughput omics data***

Author list: *Ahrens, Maike; Turewicz, Michael; Marcus, Katrin; Meyer, Helmut E.; Eisenacher, Martin; Rahnenführer, Jörg*

In personalized medicine, the identification of patient subgroups with specific gene or protein expression is an important goal. Different subgroups can indicate different molecular subtypes of a disease which might correlate with disease progression, prognosis or therapy response, and the subgroup-specific genes or proteins are potential drug targets. Using high-throughput molecular data, the aim is to characterize the patient subgroup by identifying both, the set of samples that shows a distinct expression pattern as well as the set of features that are affected.

We present a workflow for the identification of patient subgroups from two sample comparisons (e.g. healthy vs. diseased). First, a pre-filtering based on the univariate score FisherSum (FS) is applied to assesses subgroup-specific expression of the features. FS has been shown to outperform other previously presented methods like the outlier sum [1] or PAK (Profile Analysis using Kurtosis, [2]) in several settings [3]. Afterwards, the selected features are compared regarding the samples that form the induced subgroup. The latter step uses the OrderedList method [4] that was originally developed for the comparison of result lists from gene expression experiments. Our workflow FSOL is compared to a reference workflow based on classical biclustering in a comprehensive study using simulated and real world data.

The results of this exploratory approach may give hints to yet unknown mechanisms in pathologic processes and assist in the generation of new research hypotheses.

References

[1] Tibshirani R and Hastie T (2007). Biostatistics 8.

[2] Teschendorff AE et al. (2006). Bioinformatics 22.

[3] Ahrens M, Turewicz M, Casjens S, May C, Pesch B, Stephan C, Woitalla D, Gold R, Brüning T, Meyer HE, Eisenacher M, Rahnenführer J. (2013). PLOS ONE 8 (11).

[4] Yang X, Scheid S, Lottaz C (2008). OrderedList: Similarities of Ordered Gene Lists. R package version 1.38.0.

Statistical Methods for Omics Data Integration

Invited session
Tuesday, 16 June 2015
14:00 - 15:30
Room: Aula Martini
Session chair: Di Serio, Clelia

Discussant: Giuseppe Jurman

Frigessi, Arnoldo : ***Integrative genomics: a rapid overview and some new approaches***
Author list: *Frigessi, Arnoldo*

Recent and emerging -omics technologies enable measurement of molecular phenotypes at unprecedented levels of breadth and resolution. We can now view a specific molecular biological system from many perspectives, producing various types of data that capture different phases of the biological dynamics. But despite important successes, the fundamental complex dependencies between various molecular regulations are often not fully understood. A deeper understanding would help to make more precise predictions of disease progression and therapeutic efficacy.

A major outstanding aim is the development of methods that allow the joint analysis of multiple -omics components. In this lecture I will first give a systematic review of recent statistical methodology that exploits collectively multivariate -omics data sets. Then I will report on our own approaches for the unsupervised classification of cancer patients based on multiple layers of -omics data.

van Wieringen, Wessel N. : ***Ridge estimation of Gaussian graphical models from high-dimensional data***
Author list: *van Wieringen, Wessel N.*

Molecular biology aims to understand the molecular processes that occur in the cell. That is, which molecules present in the cell interact, and how is this coordinated? For many cellular process, it is unknown which genes play what role. A valuable source of information to uncover gene-gene interactions are (onco)genomics studies. Such studies comprise samples from few (n) individuals with, e.g., cancer of the same tissue. Each sample is interrogated molecularly and the expression levels of many (p) genes are measured simultaneously. From these high-dimensional omics data the gene-gene interaction network may be unravelled when the presence (absence) of a gene-gene interaction is operationalized as a conditional (in)dependency between the corresponding gene pair. Then, under the assumption of multivariate normality, the gene-gene interactions correspond to zeroes in the precision matrix (which are proportional to the partial correlations).

When dealing with high-dimensional data, the sample covariance matrix is singular and the sample precision matrix is not defined. But even if $p < n$ and p approaches n, the sample precision matrix yields inflated partial correlations. Both situations require a form of regularization to obtain a well-behaved estimate of the precision matrix, and consequently of the gene-gene interaction network.

We study ridge estimation of the precision matrix in the high-dimensional setting. We first review two archetypal ridge estimators and note that their penalties do not coincide with common quadratic ridge penalties. Subsequently, starting from a proper l2-penalty, analytic expressions are derived for two alternative ridge estimators of the precision matrix. The alternative estimators are compared to the archetypes with regard to eigenvalue shrinkage and risk. The alternatives are also compared to the graphical lasso within the context of graphical modeling. The comparisons may give reason to prefer the proposed alternative estimators.

Time allowing several extensions are discussed. For instance, the multi-group case which employs a fused ridge penalty. This penalty penalizes not only the absolute size of the precision elements but also their difference among the group precisions. Alternatively, ridge estimation of dynamic networks is considered.

Schimek, Michael G. : ***Statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists***
Author list: *Schimek, Michael G.*

Statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists

Michael G. Schimek

Medical University of Graz, Austria, Institute for Medical Informatics, Statistics and Documentation, Research Unit of Statistical Bioinformatics

High-throughput sequencing techniques are increasingly affordable and produce massive amounts of data. Together with other high-throughput technologies, such as microarrays, there are a wealth of resources in bioinformatics repositories. For more than a decade, researchers have contributed valuable experimental data to these databases. Despite different technologies and designs, many experiments share the same goal. For instance, the aims of RNA-seq studies often coincide with those of differential gene expression experiments based on microarrays. As such, it would be logical to utilize all available data. However, there is a lack of statistical methods for the integration of results from different sources. Although diverse technological platforms produce different raw data, one commonality for experiments with the same goal is that all the outcomes can be transformed into a platform-independent data format - rankings - for the same list of items (e.g. genes). In this talk we give an overview of recently proposed methods for inference [1] and stochastic aggregation [2] [3] of multiple omics ranked lists. In addition, we present the new TopKLists R package [4], and as a typical application, we integrate microRNA data of non-small cell lung cancer across different platforms and show numerical and graphical results.

References

[1] Hall, P. and Schimek, M. G. (2012). Moderate deviation-based inference for random degeneration in paired rank lists. Journal of American Statistical Association, 107, 661-672.

[2] Lin, S. and Ding, J. (2009). Integration of ranked lists via Cross Entropy Monte Carlo with applications to mRNA and microRNA studies. Biometrics, 65, 9-18.

[3] Lin, S. (2010) Space oriented rank-based data integration. Statistical Applications in Genetics and Molecular Biology, 9,1.

[4] Schimek, M. G. et al. (2015) TopKLists: a comprehensive R package for statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists. Statistical Applications in Genetics and Molecular Biology, to appear.

Innovative Study Designs and Analysis Issues in Epidemiology

Invited session
Tuesday, 16 June 2015
14:00 - 15:30
Room: U6-A10
Session chair: Valsecchi, Maria Grazia

Discussant: Antonella Zambon

Wolkewitz, Martin : ***Advanced nested case-control and case-cohort approaches in hospital epidemiology***
Author list: *Wolkewitz, Martin*

The occurrence and consequences of time-dependent adverse events (such as hospital-acquired infections, delirium or decubitus) are frequently of interest in hospital epidemiology. Multi-state modeling has become an established strategy to study risk factors for hospital-acquired infections (HAI) as well as the impact of HAI on mortality and/or extra length of hospital stay. It accounts for discharge and death as competing events for HAI and treats HAI as time-dependent exposure when analysing mortality and length of stay.

Often, it is required to use only a reduced data set, for instance due to expensive covariate information. Most studies in hospital epidemiology using subsampling are classical case-control studies which ignore competing events; case-cohort studies are rare. However, the sampling strategy has important consequences for the analysis and the results. Thus, adequate subsampling must account for the complexity of the data (competing events, time- dependent exposures) to avoid typical pitfalls such as informative censoring or time-dependent bias (1).

We propose extended sampling strategies for nested case-control and case-cohort designs which are able to approximate the quantities of interest in multi-state models from the full cohort (such as transition hazards and probabilities, subdistribution hazards and cumulative incidence functions).

A cohort study from two Spanish intensive care units is used to demonstrate the proposed methods (2). Cumulative hazards and incidence functions from the full cohort are approximated with artificially created nested case-control or case-cohort data.

References

(1) Schumacher M, Allignol A, Beyersmann J, Binder N, Wolkewitz M. Hospital-

acquired infections: appropriate statistical treatment is urgently needed. Int J Epidemiol. 2013;42:1502–1508.

(2) Wolkewitz M, Cooper BS, Palomar-Martinez M, Olaechea-Astigarraga P, Alvarez-Lerma F, Schumacher M. Nested case-control studies in cohorts with competing events. Epidemiology. 2014 Jan;25(1):122–125.

Reilly, Marie : ***Extensions to the analysis and interpretation of nested case-control data***
Author list: *Reilly, Marie; Delcoigne, Benedicte; Stoer, Nathalie; Czene, Kamila; Salim, Agus*

Background:

Until recently, one could not estimate absolute risk from nested case-control data, nor readily re-use the collected data to answer secondary research questions. Despite this weakness, case-control designs are widely used in genetic and molecular studies due to their cost-efficiency. The huge investment in the collection and measurement of biomarkers in these studies has generated interest in re-using the data to address further hypotheses. In this talk I will illustrate how data from a nested case-control design can be used to address a secondary research question and to estimate absolute risk.

Methods:

The method uses a weighted likelihood approach to reflect the sampling of the available nested case-control data. The matching on time is broken and weighting by the inverse of the probability of inclusion is introduced to represent the person-time experience in the background cohort. Assuming a Cox proportional hazards model, data from a previous nested case-control study can be reused to estimate a hazard ratio from the weighted partial likelihood. To estimate absolute risk, we combine weighted partial likelihood with weighted baseline hazard estimates and compare this to the Langholz-Borgan method.

Results:

The relative efficiency of reusing data from a nested case–control study depends on the average follow-up time in the first study with respect to the outcome in the second, and the probability that individuals in the first study are included in either study. Where control data is available from a previous study, our method also enables the design and analysis of new studies that gather only incident cases. Application to population data and disease cohorts show considerable gains in efficiency compared to sampling controls specifically for the study at hand. For the estimation of absolute risk, our method showed excellent agreement with the "true" risk profile in simulation studies and in an application to a large population-based cohort.

Conclusions:

Using weighted likelihood, the potential of nested case-control designs can be extended to analysis of secondary outcomes and estimation of absolute risk. The ability to consider reusing data gathered for previous studies may be of particular value in genetic and molecular epidemiology, where nested case–control studies use expensive measurements from biological specimens. The ability to estimate absolute risk from nested case-control data has implications for studies that aim to develop or validate risk prediction models, offering an efficient alternative to the traditional full cohort study or the case-cohort design.

Rebora, Paola : ***Two-phase cohort designs with survival outcome: estimation of incidence in the presence of competing risks***

Author list: *Rebora, Paola; Antolini, Laura; Glidden, David V; Valsecchi, Maria Grazia*

In many studies, some information might not be available for the whole cohort, some covariates, or even the outcome, might be ascertained in selected subsamples. These studies are part of a broad category termed two-phase studies. Common examples include the nested case-control and the case-cohort designs. For two-phase studies, appropriate weighted survival estimates have been derived; however, no estimator of cumulative incidence accounting for competing events has been proposed. This is relevant in the presence of multiple types of events, where estimation of event type specific quantities are needed for evaluating outcome. We develop a non parametric estimator of the cumulative incidence function of events accounting for possible competing events. It handles a general sampling design by weights related to the sampling probabilities. The variance is derived from the influence function of the subdistribution hazard. The proposed method has shown good performance through simulations and is applied in two different clinical contexts.

Clinical Trials

Contributed session
Tuesday, 16 June 2015
16:00 - 17:20
Room: U6-A11
Session chair: Schmidli, Heinz

Seibold, Heidi : ***Model-based recursive partitioning for subgroup analyses***
Author list: *Seibold, Heidi; Zeileis, Achim; Hothorn, Torsten*

The identification of patient subgroups with differential treatment effects is the first step towards individualised treatments. A current draft guideline by the EMA discusses potentials and problems in subgroup analyses and formulated challenges to the development of appropriate statistical procedures for the data-driven identification of patient subgroups. We introduce model-based recursive partitioning as a procedure for the automated detection of patient subgroups that are identifiable by predictive factors. The method starts with a model for the overall treatment effect as defined for the primary analysis in the study protocol and uses measures for detecting parameter instabilities in this treatment effect. The procedure produces a segmented model with differential treatment parameters corresponding to each patient subgroup. The subgroups are linked to predictive factors by means of a decision tree. The method is applied to the search for subgroups of patients suffering from amyotrophic lateral sclerosis that differ with respect to their Riluzole treatment effect, the only currently approved drug for this disease.

Dunger-Baldauf, Cornelia : ***Assessment of responsiveness for instruments capturing patient-reported outcomes***
Author list: *Dunger-Baldauf, Cornelia*

Patient-reported outcomes (PROs) have become increasingly important for assessing the effectiveness of treatments in clinical trials. In some disease areas such as irritable bowel syndrome, primary efficacy is measured by PROs. Many new instruments which capture PROs have been developed in recent years. An important prerequisite for the use of an instrument in a clinical trial is its responsiveness, i.e. its ability to reflect clinically important treatment effects. A variety of measures of responsiveness have been proposed. These are usually defined as functions of mean changes within a group, or between groups, and standard deviation terms. Responsiveness indices are then calculated using PRO data which were collected with the instrument of interest. Many authors provide multiple indices of responsiveness, and conclude that the instrument is responsive if these point into the same direction.

We assess measures of responsiveness by means of modeling the underlying PRO data and calculating the expected responsiveness dependent on PRO distributional properties. Using this method, the expected performance of a measure for responsiveness can be contrasted with what would be expected based on the PRO distribution. For example, the better the instrument reflects treatment effects, the higher the responsiveness index should be. For some responsiveness measures this is not always the case.

This insight can be useful for the interpretation of responsiveness results from PRO data collected in clinical studies, and the discussion of responsiveness for a given instrument. It is proposed to use modeling more often to assess properties of health economic measures.

Hsu Schmitz, Shu-Fang : ***A simulation study to compare estimate of event-free rate at a given time and operating characteristics between different frequentist and Bayesian analysis approaches in oncology single-arm phase II clinical trials using a double criteria design***
Author list: *Hsu Schmitz, Shu-Fang*

Declaring trial success using a double-criteria design [1] requires satisfying both a clinical criterion that the point estimate is not lower (or higher) than a specified clinical threshold, and a statistical criterion that the false positive probability is not greater than a specified level. Event-free rate (e.g. progression-free survival rate) at a given time (t, e.g. 6 months) is a frequently used endpoint in oncology single-arm phase II clinical trials. If each patient is followed up for a sufficient time (>=t), the outcome will be binary with either "success" or "failure". If some patients are lost to follow-up early or have not been followed up for a sufficient time (<t), their endpoint information is incomplete. Such cases may be considered as "failure" in some analysis approaches or as censored observations in other approaches. Tumor assessment (e.g. CT

scan) is usually performed at specified intervals. Disease progression identified at a tumor assessment likely had occurred earlier at some time between the last and its previous assessments, i.e. interval censored. The conventional analysis takes the last tumor assessment date for disease progression. The conventional and the interval-censored analysis approaches have been compared in some simulation studies for randomized settings. The objective of this simulation study, focusing on the setting of oncology single-arm phase II clinical trials using a double-criteria design, is to investigate the differences in estimate of event-free rate at a given time and in operating characteristics between five different analysis approaches: 1) as for a binary variable; 2) conventional Kaplan-Meier method; 3) Kaplan-Meier method considering interval censoring; 4) conventional parametric Bayesian analysis; 5) parametric Bayesian analysis considering interval censoring. Datasets are simulated for different scenarios with respect to sample size, tumor assessment interval, accrual rate, true event-free rate, censoring rate, and data cutoff. Point estimate and confidence interval (or Bayesian credible interval) for the endpoint are obtained using 5 approaches. The results are compared with respect to each of the two specified design criteria and the overall trial success.

References:

1. Neuenschwander B, et al. (2011). A proof of concept phase II non-inferiority criterion. Statistics in Medicine 30(13):1618-27

Regression

Contributed session
Tuesday, 16 June 2015
16:00 - 17:20
Room: U6-A10
Session chair: Alfo', Marco

Chaouch, Aziz : ***Approximate conditional percentiles for non-linear mixed effects models with continuous responses***
Author list: *Chaouch, Aziz; Rousson, Valentin*

Predicted conditional percentiles for non-linear mixed effect models are commonly calculated using a large number of Monte Carlo (MC) simulations. In some applications (e.g. implementation on miniaturized systems and/or calculation of confidence intervals for predicted quantiles using the sampling distribution of parameter estimates), the computation time is critical and approximation methods are desirable.

We present a general methodology to approximate the conditional distribution (i.e. given covariates) of the continuous response in a non-linear mixed effects model at a given time point by considering a second order Taylor expansion of the model, deriving analytically its first four moments and matching them with the corresponding moments of the family of sinh-arcsinh distributions [1]. The sinh-arcsinh distribution is a flexible four parameters unimodal distribution that can accommodate different levels of skewness and kurtosis and contains the normal distribution as a particular case. Parameter estimation using the method of moments allows one to reconstruct any predicted quantile for a subgroup of individuals sharing the same covariates (i.e. a priori percentiles) or for an individual for whom observed data is available (i.e. a posteriori percentiles). This has some application in the area of Therapeutic Drug Concentration Monitoring where, for certain drugs with tight therapeutic margins and high inter-individual variability, the quantile associated with a measured drug concentration in blood plasma might dictate an adjustment of the dose or the dosing regimen.

The proposed methodology is applied to the calculation of predicted percentile curves for a population pharmacokinetic model of an antifungal agent [2] and its performance is compared with that of MC simulations.

The four parameters of the sinh-arcsinh distribution also provide a mean to summarize the information from a non-linear mixed effects model. This opens interesting perspectives to compare different models using e.g. meta-analysis approaches.

References:

[1] Jones, M.C. and Pewsey, A. (2009). Sinh-arcsinh distributions. Biometrika 96, 4, 761-780

[2] Pascual, A. et al (2012). Challenging recommended oral and intravenous voriconazole doses for improved efficacy and safety: population pharmacokinetics-based analysis of adult patients with invasive fungal infections. Clin Infect Dis, 55, 3, 381-390

Solari, Aldo : ***Variable selection uncertainty in regression models***
Author list: *Solari, Aldo; Goeman, Jelle J.*

Variable selection in regression models is a longstanding problem. It arises when one want to model the relationship between the response variable and a subset of potential explanatory variables, but there is uncertainty about which subset to use.

We propose to quantify this uncertainty by constructing a confidence set that contains a number of subsets that explain the response variable equally well. If the uncertainty is high, distinguishing between competing subsets is difficult, and the confidence set will contain a large number of equally interesting subsets; conversely, if it is low, few subsets will stand out.

The proposed approach makes use of hypothesis testing by excluding subsets that are inconsistent with the data, since it focuses on rejecting, not on accepting, the null hypothesis of adequateness of a subset. Our work builds upon the 70's literature on variable selection for linear models and extends to more general regression models into the framework of [1] to provide useful information on the importance of the explanatory variables or their combinations in influencing the response.

References:

[1] Goeman JJ, Solari A. (2011) Multiple Testing for Exploratory Research (with Discussion and Rejoinder). Statistical Science, 26:584-597

De Bin, Riccardo : ***New insights on resampling-based variable selection for multivariable regression in the presence of correlation between covariates***

Author list: *De Bin, Riccardo; Boulesteix, Anne-Laure; Sauerbrei, Willi*

In biomedical studies, many potentially important variables are usually considered in the analysis of a phenomenon of interest. To include only the most useful in a multivariable regression model, a variable selection procedure is used, often backward elimination or forward selection. These traditional approaches, however, are known to have several shortcomings: one of the most problematic is their "instability", in the sense that small changes in the data may cause severe differences in the set of the selected variables. To tackle this issue, Chen & George (1985) proposed a variable selection approach which consists in keeping in the model those variables which are most of the times selected when applying a variable selection procedure to several bootstrap samples. Sauerbrei & Schumacher (1992) extended this approach by considering the bivariate relationship of variable inclusion frequencies: it may happen that the variable selection procedure alternatively identifies as relevant only one of two correlated variables in the bootstrap samples, leading to a relatively small bootstrap inclusion frequency for both of them. As a consequence, both variables may be excluded from the final model, especially if sparsity is a desirable property. Sauerbrei & Schumacher (1992) suggested two procedures based on the 2 x 2 tables of inclusion frequencies. In a simulation study we investigate the two procedures and discuss extensions for the situations with more complex relationship of inclusion frequencies.

References:

Chen, C.-H. & George, S. L. (1985). The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. Statistics in Medicine 4, 39–46.

Sauerbrei, W. & Schumacher, M. (1992). A bootstrap resampling procedure for model building: application to the Cox regression model. Statistics in Medicine 11, 2093–2109.

Heinze, Georg : ***Variable selection with the change-in-estimate criterion: a bluff package?***

Author list: *Heinze, Georg; Dunkler, Daniela*

Typical statistical modeling situations in prognostic or etiologic research often involve a large number of potential explanatory variables. Selecting a suitable subset in an objective and practical manner is a non-trivial task, and is usually based on significance.

Alternatively, the change-in-estimate criterion postulates that an adjustment variable should not be removed from a statistical model if its removal causes a relevant change in the estimate of another important variable in the model. We provide an approximation to the change-in-estimate criterion, from which a simple significance test for the change-in-estimate can be derived. Further investigation shows that using a significance-based threshold for the change-in-estimate criterion reduces to a simple significance-based selection of variables, as if the change-in-estimate criterion is not considered at all.

We discuss practical situations where the change-in-estimate criterion may still be useful to consider, in particular if it is suitably standardized. A comprehensive algorithm, ‚Augmented Backward Elimination', combines backward elimination of non-significant effects with the standardized change-in-estimate criterion and is available as a SAS macro at www.tinyurl.com/abe-sas.

Survival Analysis (1)

Contributed session
Tuesday, 16 June 2015
16:00 - 17:20
Room: Aula Martini
Session chair: Ambrogi, Federico

Bernasconi, Davide Paolo : ***Flexible regression models in survival analysis for dynamic prediction with non-reversible time-varying treatment***
Author list: *Bernasconi, Davide Paolo; Antolini, Laura; Valsecchi, Maria Grazia*

In survival studies it is often of interest to compare the effect of treatments on the survival experience. In some situations it may happen that the treatment changes during the follow-up: a typical example is chemotherapy vs stem-cell transplantation in Acute Lymphoblastic Leukemia (ALL) where patients are treated initially with chemotherapy and they can receive stem-cell transplant after some time. In this context one may be interested in: (i) assessing the impact of the waiting time to transplant on the hazard of failure; (ii) developing profile-specific (dynamic) predictions.

In this work, we point out that a regression model with time measured in the original scale (i.e. from remission), where treatment switch is accounted through left truncation, does not allow to properly address issue (i). In addition, in the presence of a time-varying effect of the time-varying treatment variable, the extended Cox model is not useful to achieve goal (ii).

Concerning (i) we show that the coefficient of waiting time obtained fitting a regression model only on transplanted patients, measuring time on the clock back scale (i.e. since transplant), provides an accurate estimate of the real impact of the waiting on the hazard of failure. About (ii) we review and discuss two recently proposed flexible regression models, namely the "Hanley-Miettinen model" [1] and the "Landmark regression model" [2], that can overcome the limits of the Cox model to develop profile-specific (dynamic) predictions in complex situations. We also discuss the use of these models using different time scales, arguing that measuring time since transplant (clock-back scale) for the transplanted patients provides better predictions than adopting the original time scale for all patients.

To support our conclusions we compared the traditional and the proposed methods both through a simulation protocol and with an application to real data on pediatric ALL.

References

[1] Hanley JA, Miettinen OS. Fitting smooth-in-time prognostic risk functions via logistic regression. International Journal of Biostatistics, (2009).

[2] Van Houwelingen HC, Putter H. Dynamic prediction in clinical survival analysis. Chapman & Hall, (2012).

Orenti, Annalisa : ***Estimating relapse free survival as a net probability: regression models and graphical representation***
Author list: *Orenti, Annalisa; Biganzoli, Elia; Boracchi, Patrizia*

In several experimental or observational clinical studies, the evaluation of the effect of a therapy and the impact of prognostic factors is based on relapse-free survival and the suited regression models. Relapse free survival is a net survival and needs to be interpreted as the survival probability that would be observed if all patients experienced relapse sooner or later.

Death without evidence of relapse prevents the subsequent observation of relapse, acting in a competing risks framework.

Relapse free survival is often estimated by standard regression models after censoring times to death. The association between relapse and death is thus ignored. However to better estimate relapse free survival a bivariate distribution of times to events needs to be considered, for example by means of copula models, with ad hoc estimating procedures. We concentrate here on the copula graphic estimator, for which a pertinent regression model has been developed (Lo and Wilke). The advantage of this approach is based on the relationship between net survival, overall survival and cause specific hazard. Regression models can be fitted for the latter quantity by standard statistical methods and the estimates can be

used to compute net survival through a copula structure. Parametric models are preferred. To avoid the constraint of parametric distribution, we propose piecewise regression models. A consistent estimate of the association parameter for the copula model can be obtained by considering the semi-competing risks framework, because death can be observed after relapse. The drawback of the copula graphic regression model is that no direct parametric estimation of the regression coefficient for the covariates is available. To obtain an overall view of the association between covariate levels and net relapse free survival we propose a multivariate visualisation approach through Multiple Correspondence Analysis.

This approach has been applied to two case series of patients with breast cancer and extremity soft tissue sarcoma respectively, in order to compare the results obtained by piecewise exponential model on cause specific hazard and net relapse free survival computed through copula graphic estimator.

Fritz, Josef : ***Statistical approaches in mediation analysis: a comparison of methods for survival data***

Author list: *Fritz, Josef; Ulmer, Hanno*

In clinical research and epidemiology, there is increasing interest to quantify effects of a given treatment or risk exposure into different causal pathways via the so called mediation analysis. Traditionally, an approach introduced in 1986 by Baron and Kenny was used to decompose the total effect of a given treatment or exposure into a direct and an indirect effect through one or more mediators. Recently, Lange et al. have presented a new framework for mediation analysis based on marginal structural models that is also applicable in the case of survival data. It has been shown that the Baron and Kenny method works well in the special case of linear models without interactions, but is mathematically inconsistent otherwise.

We applied both approaches, Baron and Kenny versus Lange et al., on two examples in the field of cardiovascular epidemiology. In these applications, the relationship between (1) sex and (2) body-mass index with coronary heart disease mediated by cardiovascular risk factors was investigated. In both examples, we found substantial differences between the two methods. We present and discuss both methods with regard to their theoretical features and their applicability in a large dataset.

Strohmaier, Susanne : ***A simple to implement algorithm for natural direct and indirect effects in survival studies with a repeatedly measured mediator***

Author list: *Strohmaier, Susanne; Rosenkranz, Nicolai; Wetterslev, Jørn; Lange, Theis*

Important questions within the fields of social sciences, epidemiology as well as clinical trial research involve the challenge of decomposing the effect of an interventions into direct and indirect effects working through a defined mediator, thereby aiming for a better understanding of the underlying mechanisms.

For the case of a single and multiple mediators measured at a single point in time, researchers have established theoretical properties (e.g. Pearl (2012)[1]) and developed practical tools for the analysis of a broad class of mediator and outcome models (e.g. Lange et al. (2012, 2014)[2,3]) by employing the counterfactual framework. However, data structures are often more complex than the described scenarios.

We present an extension to the procedure by Lange et al. to the setting of a time-to-event outcome and a repeatedly measured mediator, where the number of measurements is determined by survival time. We suggest an estimation algorithm, that allows for direct parametrisation of direct and indirect natural effects and is easy to implement using standard software.

The proposed method enables us to analyse the mediating role of KDIGO (a measure of severity of kidney impairment) on mortality in the Scandinavian Starch for Severe Sepsis/Septic Shock trial (6S) comparing two substances for fluid resuscitation among patients with severe sepsis admitted to intensive care units.

References:

[1] Pearl, J. (2012). The causal mediation formula - A guide to the assessment of pathways and mechanisms. Prevention Science, (13):426-436.

[2] Lange, T., Vansteelandt, S., and Bekaert, M. (2012). A simple unified approach for estimating natural direct and indirect effects. American Journal of Epidemiology, 176(3):190-195.

[3] Lange, T., Rasmussen, M., and Thygesen, L. C. (2014). Assessing natural direct and indirect effects through multiple pathways. American Journal of Epidemiology, 179(4):513-518.

Poster Reception

Contributed session
Tuesday, 16 June 2015
18:00 - 19:30
Room:
Session chair:

Gu, Fei : ***Raw data maximum likelihood estimation for principal component analysis and two types of common principal component model: A state space approach***
Author list: *Gu, Fei; Yung, Yiu-Fai*

Raw data maximum likelihood (ML) method implemented by a state space approach was proposed to principal component analysis (PCA) and two types of common principal component (CPC) models, namely the CPC model for independent random vectors [1-2] and the CPC model for dependent random vectors [3-4]. The state space approach uses the raw data ML method that provides both the point estimates and standard error (SE) estimates for all parameters under the normality assumption. Moreover, the raw data ML method is advantageous to the existing Wishart-likelihood approach for the CPC model for dependent random vectors, because the SE estimates for all parameters are provided in a single step, unlike the procedure developed by Neuenschwander and Flury [5] that has to be repeatedly applied only for functionally independent parameters. Detailed specification of a state space model to PCA and the two types of CPC model are provided, and morphometric examples are used to illustrate the proposed state space approach. A small simulation study is also conducted to assess the accuracy of the SE estimates from the state space approach under normality and nonnormality conditions. It is found that the SE estimates obtained from the state space approach are robust to the violation of normality assumption, whereas the results from Neuenschwander and Flury's procedure are sensitive to nonnormality. Future research is discussed at the end.

REFERENCES

[1] Flury, B. N. (1984). Common principal components in k groups. Journal of the American Statistical Association, 79, 892-898.

[2] Flury, B. N. (1988). Common principal components and related multivariate models. New York, NY: Wiley.

[3] Flury, B. D., & Neuenschwander, B. E. (1995). Principal component models for patterned covariance matrices with application to canonical correlation analysis of several sets of variables. In W. J. Krzanowski (Ed.), Recent advances in descriptive multivariate analysis (pp. 90-112). New York, NY: Oxford University Press.

[4] Klingenberg, C. P., Neuenschwander, B. E., & Flury, B. D. (1996). Ontogeny and individual variation: Analysis of patterned covariance matrices with common principal components. Systematic Biology, 45, 135-150.

[5] Neuenschwander, B. E., & Flury, B. D. (1995). Common canonical variates. Biometrika, 82, 553-560.

Greco, Teresa : ***The attractiveness of network meta-analysis***
Author list: *Greco, Teresa*

INTRODUCTION

Network meta-analysis provides a global estimate of comparative treatment effectiveness combining both direct and indirect evidence [1]. This method of simultaneously comparing all available healthcare interventions is very attractive to clinicians because it can respond to their major concern, namely which treatment is the best.

We investigated the dissemination of network meta-analyses in the clinical literature of discussing the principal areas of application and their general descriptive characteristics.

METHODS

We performed a systematic and narrative review (last updated on April 15, 2014) in order to assess the scholarly diffusion of network meta-analyses. The following data were collected: author identification, year and journal of publication, PubMed index, number of treatments and studies included, characteristics of

network configuration, nature of primary outcome, clinical indication, type of intervention investigated and medical area.

RESULTS

Since 2003 there has been an exponential increase in the number of published network meta-analyses. Out of 340 articles included according to our selection criteria, encompassing 248 treatment networks, cardiovascular (25 of 71, 35%) was the most prevalent topics, with an average of 5 treatments (1st quartile-3rd quartile: 3-8; minimum-maximum: 2-120) being compared stemming from an average of 10 controlled trials (1st quartile-3rd quartile: 21-43; minimum-maximum: 2-267). The types of intervention most frequently analyzed were: coronary stents (8 of 71, 11.3%), antihypertensive drugs (7 of 71, 9.9%) and bronchodilator drugs (5 of 71, 7.0%).

DISCUSSION

The specific pattern of uptake of network meta-analysis may be due to the established hierarchy in world-wide causes of morbidity and mortality. Indeed, the World Health Organization (WHO) reported that the most common causes of death were ischemic heart disease and stroke. Accordingly, it is not surprising to see such attention to cardiopulmonary topics among reviewers, journals, and readers.

In conclusion, network meta-analyses are becoming increasingly attractive as they offer a comprehensive framework for decision-making. Whether they will also contribute to improvements in patient outlook remains to be proven.

REFERENCES

1. Greco T, Landoni G, Biondi-Zoccai G, D'Ascenzo F, Zangrillo A. A Bayesian network meta-analysis for binary outcome: how to do it. Stat Methods Med Res 2013 [Epub ahead of print].

Hayoz, Stefanie : ***Effect of one-patient clusters on power in cluster-randomized trials***
Author list: *Hayoz, Stefanie; Klingbiel, Dirk*

Aims: The aim of this research was to assess the effect of one-patient clusters on the power in cluster-randomized trials with small cluster sizes.

Methods: The assumptions for the simulations were based on the trial SAKK 95/06 where 40 physicians (clusters) with equal patient numbers of 4 (160 patients in total) were planned. A linear mixed model with physician as random effect should yield a power of 81.9% with alpha 5%. In reality the cluster sizes varied between 1 and 10, with 22.6% one-patient clusters. Simulations were performed with 10000 repetitions per scenario.

Results: With 40 physicians, cluster sizes between 1 and 10 and 5.0%-25.0% one-patient clusters, the achieved power varied between 79.6% and 80.7%. If the number of patients per cluster was restricted to a maximum of 6 and the number of physicians increased to 41–55 with 4.9%–36.4% one-patient clusters, the achieved power varied between 81.8% and 82.6%. Resampling from the cluster sizes of SAKK 95/06 with 7%–43% one-patient clusters yielded a power between 78.1% and 79.3%. When the one-patient clusters were excluded from the analysis the achieved power was 0.3%–0.9% lower. The proportion of models with numerical problems was never higher than 0.2%.

Conclusion: In all considered scenarios the decrease in power in comparison to the theoretical model with equal cluster sizes was negligible. When the number of patients per cluster was restricted and instead the number of physicians was increased there was no relevant loss in power even with a high proportion of one-patient clusters.

De Capitani, Lucio : ***Reproducibility probability estimation for common nonparametric testing***
Author list: *De Martini, Daniele; De Capitani, Lucio*

The evaluation of the reproducibility of clinical trial results is a very important and hot topic [1]. When experimental data are evaluated by statistical tests, an analytical technique can be adopted to evaluate the reproducibility of test results. Indeed, the reproducibility of results can be viewed as the reproducibility of statistical significance. Due to the random nature of significance, reproducibility can be evaluated by computing the Reproducibility Probability (RP), i.e. the probability of obtaining "a statistically significant result" in a new, identical and independent trial.

The RP is unknown, since it depends on the true effect size of the drug under study. Nevertheless, the RP can be estimated using trial data, providing useful information on the stability of the test outcome.

Moreover, the RP estimate not only evaluates how much a significant result is reproducible, but it can also play the role of test statistic. In detail, there exists a general threshold for the statistical significance based on the RP estimate, that is 1/2. The decision rule is, therefore: "the null hypothesis is rejected when it is estimated that there is more inclination to reject it than to accept it". This technique, which is called RP-testing, holds for a wide family of parametric tests [2], and it has been extended to a commonly used nonparametric test [3].

In this talk, RP-estimation and RP-testing are studied for some nonparametric tests (Sign Test, Wilcoxon Signed Rank Test, Kendall test). In detail, the performance of several RP-estimators for the aforementioned nonparametric tests are evaluated: their MSE is computed first, in order to find the most efficient estimator; then, the RP-based decision rule given by RP-estimators is studied. When RP-estimators satisfy a certain condition, the RP-based decision rules exactly replicates the original nonparametric test; otherwise, very small probabilities of disagreement are observed.

REFERENCES

1. Collins, FS, Tabak, LA. NIH plans to enhance reproducibiliy. NATURE Vol.505(7485), 612-613, 2014.

2. De Martini D. Reproducibility probability estimation for testing statistical hypotheses. STAT.&PROB.LETT. 78(9), 1056-1061, 2008.

3. De Capitani L, De Martini D. Reproducibility probability estimation and testing for the Wilcoxon rank-sum test. J.STAT.COMPUT.&SIMUL. 85(3), 468-493, 2015.

Riedl, Regina : ***The evaluation of the Austrian's Disease Management Program in patients with diabetes mellitus type 2 - An application of propensity score matching***

Author list: *Riedl, Regina; Berghold, Andrea*

One of the Disease Management Programs (DMPs) in Austria is called "Therapie aktiv – Diabetes im Griff" which is a systematic treatment program for patients with diabetes mellitus type 2. The implementation of this program started in 2007 and the participation is optional and free of charge for general practitioners and patients.

Based on routine health insurance data for type 2 diabetic patients, we evaluated the DMP concerning patient- relevant outcomes (mortality, myocardial infarction and stroke) and costs. We conducted a retrospective study using a control group design, whereat the DMP-group consists of participants enrolled in the program during 2008 and 2009 (n=7181). A comparable control group, consisting of patients with no participation in the DMP up to 2013, was constructed by using a propensity score (PS) matching approach. The PS, defined as conditional probability of participating in the DMP given a set of baseline covariates, was estimated via logistic regression including patient's characteristics, form of antidiabetic therapy, several prescriptions, the number of hospital admissions and days, main discharge diagnoses and costs. The matching was performed stratified for baseline year and state, by using a nearest-neighbour-matching without replacement choosing 3 PS-matched controls in sequential order.

Before matching large imbalances were observed for age, the form of antidiabetic therapy, total costs and hospital days, indicating that DMP participants tended to be younger and healthier compared to non-participants. After matching no substantial imbalances were observed for all our measured baseline characteristics. Over a follow-up period of four years, we observed a significantly lower mortality rate and a reduction in total costs for the DMP participants in comparison to the control group. We will discuss strengths and limitations of the study.

Ofner-Kopeinig, Petra : ***The use of a generalized measure of imbalance to allow the use of the big stick randomization procedure for more than two treatment groups and unequal allocation rates: a simulation study***

Author list: *Ofner-Kopeinig, Petra; Errath, Maximilian; Berghold, Andrea*

Most randomization procedures use the absolute difference of the number of the allocated treatments as a measure of imbalance to determine the allocation probability in the next randomization step. In the randomization process, using the Big Stick method (BSD) by Soares and Wu (1), treatments are allocated at random until the absolute difference reaches a tolerance of "a". Then the underrepresented treatment group is allocated with a probability of 1.

In the case of more than two treatments or unequal allocation rates any urn model and also permuted blocks randomization can be adapted very simple. In both cases the allocation probability for the next randomization depends on the composition of the urn.

For BSD one would need a measure of imbalance to determine allocation probabilities for the next randomization step. Here using absolute differences is not possible anymore. For this case we propose a measure of imbalance that takes into account the expected frequencies of the treatments at each randomization step. Using this generalized measure of imbalance one can expand BSD to more than two treatment groups and unequal allocation rates. This measure also works for the Biased Coin method proposed by Efron (2).

To evaluate BSD with this generalized measure of imbalance a simulation study was undertaken. Using the simulation tool of the "Randomizer for Clinical Trials" each simulation was performed 10000 times for an accuracy of 1 % and a confidence interval of 95 %. Treatment imbalances at different stages of the study were calculated to show the balance behaviour. To estimate the predictability the probability of correct guessing and the probability of deterministic allocation were calculated. For BSD tolerances of 6, 10 and 14 were used and compared with other randomization methods.

(1) Soares Jf, Wu CF. Some restricted randomization rules in sequential designs. Commun Stat Theory Methods 1983; 12:2017-34

(2) Efron B, Forcing a sequential experiment to be balanced, Biometrika 57: 403-417, 1971

Dunkler, Daniela : ***Confidence intervals for fractional polynomials***

Author list: *Dunkler, Daniela; Gregorich, Mariella; Heinze, Georg*

Fractional polynomials can be used to model nonlinear associations of continuous risk factors with an outcome of interest (FPs, Royston & Altman, 1994). Specifically, one or two elements of a polynomial transformation of the original variable are selected such that in subsequent regression analysis an optimal fit is obtained. When constructing confidence intervals (CI) for the expected outcome at different values of the risk factor or for contrasts between different risk factor values this selection is usually ignored. This can lead to undercoverage if the selection is not stable, or if FPs are not flexible enough to capture a specific underlying shape of non-linearity. Reliable CIs can be obtained by the bootstrap, but this requires a large number of models to be evaluated. Therefore, we consider some alternative methods for CI estimation: model-based CIs with three degrees of freedom (MB3DF) to account for selection of two powers and Bayesian model averaging (BMA) of several evaluated FP models.

In a logistic regression setting we compared these approaches with simple, standard model-based CIs and bootstrapped CIs in a simulation study, assuming various nonlinear associations between a continuous risk factor and a binary outcome.

Our simulations revealed that with a true linear association BMA proves satisfactorily, and MB3DF may overcover. With nonlinear associations that can be modeled with FPs, both methods improve over simple model-based CIs without increasing the computational demand. In particular, we recommend the implementation of Bayesian model averaging CIs in software for FP estimation to stimulate their practical use.

Decarli, Adriano : ***The evaluation of the activities and decisions in health systems: a new methodological approach***

Author list: *Rosato, Valentina; Andreano, Anita; Russo, Antonio; Decarli, Adriano*

On behalf of the OSSERVA Working Group (OSServatori Epidemiologici e Registri tumori per la VAlutazione in sanità)

When evaluating if the care delivered to oncologic patients is adherent to evidence-based guidelines, it is necessary to consider multiple aspects of the diagnostic, therapeutic and follow-up pathways. Quality indicators have been developed from guidelines to measure the various aspects of the clinical pathway. However, single indicators do not give a complete view of the cure paths and their adherence to the guidelines. Moreover, it is difficult to evaluate several indicators simultaneously. Thus, different approaches have been proposed to summarise these indicators in few and more easily interpretable components of the decisional process, but they either oversimplify the problem or are difficult to interpret. In fact, the evaluation of the whole care pathway of an oncologic disease is complex, involving numerous elements of the decisional process that are mutually influenced. For these reasons, other statistical methods are needed.

The Partial Least Squares Path Modeling (PLS-PM) involves two type of models: the 'measurement' takes into account the relations between observable indicators and the corresponding non measurable latent variable while the 'structural' model explores the relations between the latent variables. PLS-PM is

characterized by an iterative procedure of linear regressions taking into account the relations in both the 'measurement' and the 'structural' model. The result is the estimate of a set of weights which are used to calculate the latent variables as linear combinations of their associated indicators. This method allows the reduction of dimensionality of several indicators into a smaller number of components, taking into account the relationships between these components.

The cohort under study is based on the incident breast cancers occurring between 2007 and 2009 in six local health authorities in Lombardy. The aim of our study is to evaluate the appropriateness of the procedures performed in each step of the diagnostic and therapeutic breast cancer pathway summarizing each phase through a single measure estimated using the PLS-PM. Our study represents an unique example of PLS-PM application in such an evaluation of the adherence to the care pathway of an oncologic disease.

Eusebi, Paolo : ***Bivariate model with non parametric random effects for diagnostic meta-analyses***

Author list: *Eusebi, Paolo*

The rapid growth of evidence-based medicine has led to a dramatic increase in attention to evidence-based diagnosis by meta-analysis of diagnostic test accuracy studies.

Several types of statistical methods are currently available for such studies. One of these methods is the Bivariate Model (BM; Reisma et al. 2005), which involves a simultaneous analysis of the sensitivity and specificity from a set of studies.

The BM assumes that random effects are normally distributed.

The BM was extended (Eusebi et al 2014) adding a discrete latent variable for identifying clusters of studies to gain additional insight into the accuracy of a test. This Latent Class BM (LC-BM) provides for clusters estimates of sensitivity. Schlattaman and others followed a similar approach (2014) replacing the random effects of the BM with a discrete latent variable, also providing cluster-specific estimates of sensitivity and specificity.

We proposed a Nonparametric Bivariate Model (NP-BM) extension of the standard BM, by using a non-parametric distribution for the random effects, allowing for a reduction of the number of estimated parameters. The two discretized random effects can be seen as an unconstrained alternative to the mixture formulation of Schlattmann et al. (2015).

We have applied the model to a meta-analysis of a tumor marker for bladder cancer (Glas et al. 2003).

Our proposed approach can lead in some situations to less biased estimates and more reliable confidence intervals for sensitivities and specificities.

References

Eusebi et al. Latent class bivariate model for the meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol. 2014.

Glas AS, Roos D, Deutekom M, Zwinderman AH, Bossuyt PM, Kurth KH

Tumor markers in the diagnosis of primary bladder cancer. A systematic review.

J Urol. 2003

Reitsma et al. Bivariate analysis of sensitivity and specificity pro- duces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005.

Schlattmann P, Verba M, Dewey M, Walther M. Mixture models in diagnostic meta-analyses–clustering summary receiver operating characteristic curves accounted for heterogeneity and correlation. J Clin Epidemiol. 2015.

Ferré, Chiara : ***Spatial mixed effect models to investigate relationships between trees and soil in an alluvial plain***

Author list: *Ferré, Chiara; Castrignanò, Annamaria; Comolli, Roberto*

The standard statistical approach, mostly based on regression methods, has been extensively applied to explore the relationships between soil properties and plant variables. However, linear regression with uncorrelated errors is subject to assumptions which are not always plausible. The object of the study was to investigate the relationships between tree growth and soil characteristics by comparing two regression approaches: a linear model with uncorrelated error (non-spatial model) and a mixed effect model including

spatial correlation structure of the residuals (spatial model). The study area (20 hectares) is located in an abandoned meander of the Oglio river (southern Lombardy, Italy), with young soils of alluvial origin (Calcaric Fluvisols). During 2002 a tree plant for wood production was realized (oak, hornbeam, ash, alder, and walnut; poplar only in the first part of the growth cycle).

In 2004 the soil was sampled at 121 points, according to a regular grid within the shallow soil horizon (Ap). On the collected soil samples pH in H2O and KCl, texture, total carbonates, soil organic C, available P, exchangeable K and depth to groundwater were determined.

In 2014 diameter at breast height of all the trees was measured.

A linear mixed effect model with correlated errors of tree diameter was estimated using soil properties as covariates and "species" as a nominal variable. The method to obtain parameter estimates was REML.

To evaluate soil-plant relationships, the measured diameter was migrated to the nearest soil sample up to a maximum distance of 8 m. Diameter was submitted to Gaussian anamorphosis to linearize the relationship with predictors.

For both models total carbonates, coarse sand , fine silt and species were highly significant, however the estimates of the coefficients, their errors and F statics were different. In particular the F-values to test the null hypothesis were greater in non-spatial model. Therefore, while the general conclusions for this particular data set do not change, it is easy to see that they could be greatly affected in other data sets.

Lucà, Federica : ***Discrimination of alluvial fan depositional processes through a statistical approach (Calabria, S. Italy)***
Author list: *Lucà, Federica; Robustelli, Gaetano*

At the outlet of steep catchments, depositional processes ranging from streamflow to debris flow are usually responsible for the construction of alluvial fans. Apart from geological and tectonic factors controlling sediment availability in the basin, several authors underlined the existence of morphometric controls on depositional process incidence on fans.

This work, carried out on the western coast of northern Calabria, aims to analyze the relationships between alluvial fans and relative drainage basin morphology and to identify, through a statistical approach, which of the variables can best differentiate sedimentary processes. In the study area, Holocene fans were mapped at the mouth of steep V-shaped valleys dissecting the Coastal Range where weathered crystalline–metamorphic and subordinate carbonatic rocks are overlaid by thick and discontinuous pedoregolithic covers.

For each fan and its basin, morphometric variables describing the relief and the drainage were derived from a digital terrain model constructed from 1:10000 topographic maps. Percentage of lithological units cropping out in the catchments was also taken into account. Through geological field surveys and historical documents, on the basis of the dominant sedimentary response observed at the catchment outlet, alluvial fans were classified into two groups (i) those that are known to produce debris-flows and (ii) those that do not show any stratigraphic or morphological evidence of such processes.

Within each group, morphometric relationships between alluvial fans and their basins were analyzed and, after splitting the samples classified as Type I or Type II into a training set and a validation set, a stepwise logistic regression was used to construct a model for predicting alluvial fan dominant depositional processes.

The analysis showed statistically significant differences between debris-flow and fluvial fan morphometric features supporting the hypothesis that alluvial fan morphology reflects the processes occurring in the drainage basin. In addition, logistic regression gave satisfactory results, with 82 and 86 per cent of the fans correctly classified, in the training set and calibration set respectively.

Alunni Fegatelli, Danilo : ***Flexible behavioral capture-recapture modelling***
Author list: *Alunni Fegatelli, Danilo; Tardella, Luca*

In the context of discrete-time closed capture-recapture analysis, alternative model frameworks have been proposed to handle different features of the behavioural effect to capture. We propose a new strategy for building and fitting flexible parametric capture-recapture models to address a better understanding of behavioural patterns. Within a natural transition model framework, we rely upon the capture probabilities conditioned on each possible partial capture history. This is a natural way of keeping track of the sequential behavioural changes.

The main idea is to suitably summarize any partial capture history with the aim of modelling and understanding the accumulation of memory effects possibly vanishing or reinforcing during the course of the capture sessions. We get new models based on these summary statistics used as explanotary variables in a logistic model framework. Linear and non-linear logistic regression based on a suitable behavioural summary are illustrated. In particular we show how, when the non linear regression is based on a step function it turns out to partition conditional probability parameters into equivalence classes, possibly recovering known standard behavioural or temporal models and discovering new meaningful ones. We show how one can easily implement the unconditional likelihood approach within a generalized linear model framework recycling standard GLM routines. Finally, we illustrate the potential of our proposal with the analysis of a well known dataset and a simulation study.

Pramov, Aleksandar : ***Confidence intervals for a partially identified parameter with bounds estimated by the minimum and the maximum of two correlated and normally distributed statistics***

Author list: *Pramov, Aleksandar; Bonetti, Marco*

We study the problem of constructing a Confidence Interval (CI) for a partially identified parameter. The lower and upper bounds of its identification region are given by the minimum and the maximum of two functions of the distributional parameters.

The motivation behind this work comes from a problem of comparing the relative accuracy of two binary diagnostic tests in presence of noncompliance with a subsequent nonignorably missing disease indicator. As a consequence, some parameters of the observed data distribution are many-to-one functions of the original parameters.

Hence, relevant measures of relative accuracy, such as the relative true positive rate or relative false positive rate, are not point-identified from the observed data distribution. However, they might be partially identified [1]. The bounds of the identification region are estimated by the min and the max of two asymptotically jointly normal estimators.

Existing methods can deal with the construction of CIs for such partially identified parameters or their identification regions, but often require (joint) asymptotic normality of the statistics [2]. However, when the quantity of interest is bounded by the minimum and maximum of two correlated normal variables, the resulting estimators for the bounds are generally not normal.

We give an overview of some existing methods for partial identification and propose extensions to deal with the case of using such a particular type of non-differentiable functions as bounds. We aim to provide analytic results which deliver asymptotically the desired coverage for the CI pointwise. Additionally, we compare our approach with some bootstrap alternatives for construction of CIs for a partially identified parameter.

[1] Manski, Charles F. “Partial identification of probability distributions.” Springer Science & Business Media, 2003.

[2] Imbens, Guido W., and Charles F. Manski. "Confidence intervals for partially identified parameters." Econometrica 72.6 (2004): 1845-1857.

Gasparini, Alessandro : ***A weighted form of the estimator for the cause-specific cumulative incidence function can be computationally more efficient using R***

Author list: *Gasparini, Alessandro*

Introduction

The standard cumulative incidence function estimator can be written in many alternative forms [1]: one is a weighted empirical cumulative distribution function. As a consequence, it is possible to analyze competing risks data using weighted versions of standard survival analysis methods. We want to investigate an eventual computational advantage using the weighted estimator.

Methods

Using R, we generated competing risks data (with different sample sizes up to 10000 individuals). Consequently, we performed two simulations. First, we compared the time necessary to estimate the ordinary Fine & Gray model ("crr" function, "cmprsk" package) versus the procedure using the weighted estimator ("crprep" function,"mstate" package) and the regular Cox model. We ran each procedure 100 times at each sample size, and averaged the results. Then, we assumed the need to estimate a competing risks model

many times, as in multiple imputation, for instance. Hence, we ran "crr" 10 to 100 times and compared the total time with the overall time needed to compute weights once and estimate Cox models. R code available on Github [2].

Results

We found less computational burden running the standard procedure compared to the weighted one. The latter consistently required more time to compute the results: the actual time difference increased with larger sample sizes, but the increment was less that proportional. It required 5.5 times the time for a small sample size, and decreased to 1.5 times for a sample size of 10000. On the contrary, the weighted procedure required overall less time when many models needed to be estimated. The time advantage was slight with small n and sample size, and dramatic (up to a 90% reduction) for a sample size of 10000 and 100 models.

Discussion

The most commonly used R function to estimate a Fine & Gray model is usually more computationally efficient than the weighted procedure. On the contrary, to estimate more than once a model the weighted procedure should be preferred.

References

[1] R. B. Geskus, "Cause-specific cumulative incidence estimation and the Fine and Gray model under both left truncation and right censoring", Biometrics, 67, 39–49 (2011).

[2] https://github.com/ellessenne/crr-vs-crprep

Castrignanò, Annamaria : ***Use of mixed effect models to investigate the relationships of hyperspectral and fluorescence data with plant variables***

Author list: *Stellacci, Anna Maria; Stelluti, Matteo; Losciale, Pasquale; Rossi, Roberta; De Benedetto, Daniela; Tarricone, Luigi; D'Andrea, Laura; Tomaiuolo, Matteo; Campi, Pasquale; Giglio, Luisa; Fornaro, Francesco; Vitti, Carolina; Vonella, Vittorio; Castrignanò, Annamaria*

Proximal sensors, such as hyperspectral and fluorescence devices, provide important information about plant status as they enable field investigation in real time and at very fine spatial and temporal scales. Statistical analysis is fundamental to extract critical features and assess relationships between proximal indices and plant variables. In modelling such relationships, spatial autocorrelation of residuals should be taken into account as it can affect estimates of model coefficients and inference from statistical models. The aim of this study was to compare spatial and non-spatial (OLS) models for quantifying predictive relationships between proximal data (hyperspectral and fluorescence) and plant variables (N.uptake and LAI).

Proximal and biometric data were collected in a field located in Foggia (southern Italy), at anthesis stage of durum wheat in 2014, on 104 georeferenced locations. Fluorescence data were collected using a portable fluorimeter (Multiplex). Measured variables (fluorescence emitted in yellow, red and far-red induced by UV, blue, green and red light) were synthetized by factor analysis and the first two extracted components (F1, F2) were used in this study. Hyperspectral signatures were recorded using a sensor operating in the 325-1075nm region (Fieldspec, ASD) and were collected only in 52 locations, at about 50cm above the canopy. Data were processed by computing: discrete vegetation indices; multivariate indices using the whole spectrum; indices based on the shape of reflectance curve.

Mixed-effects models with correlated errors were estimated between plant variables and sensing data. Non-spatial models, including same fixed effects of spatial models but assuming independence of errors, were also computed. Residual autocorrelation was evaluated through Moran test; spatial and non-spatial models were compared using log likelihood ratio.

Moran test and log ratio showed that residuals were not spatially correlated for regressions between hyperspectral indices and plant variables; this result might be attributed to the too coarse sampling scale. Significant spatial dependence was instead observed for fluorescence variables, particularly for the relationship between F1 and N.uptake (Moran test, P=0.0016; log ratio, P=0.055). Spatial models taking into account residual autocorrelation can be very helpful in agricultural and environmental research as they allow correct estimation of confidence intervals and hypothesis testing.

Wei, Wei : ***Aggregating epidemiological evidence through multivariate syndromic surveillance: methodological challenges with a particular focus on animal health data***

Author list: *Wei, Wei; Vial, Flavie; Berezowski, John; Held, Leonhard*

The question of how to aggregate animal health information derived from multiple data streams that vary in their specificity, scale, and behavior is not trivial. Since there is often different information contained in observations from different data sources, outbreak detection systems should be multivariate by nature. Outbreak detection in a multivariate animal health context should be viewed as a probabilistic prediction problem.

Animal health data have many characteristic features. They may originate from various sources (e.g. slaughterhouses, farmers, veterinarians) and may be collected over several species in diverse environments (e.g. wild animals, companion animals and livestock). Such data often exist in mixed forms: standardized ratios, proportions, counts of adverse events, categorical data and continuous numbers. The complexity of animal production systems dictates the concomitant monitoring of many time series; and makes the investigation of statistical signals imperative and at the same time difficult and resource intensive. Multivariate surveillance methods that can work across multiple data streams to increase both sensitivity and specificity are much needed.

After starting with an overview of current methodologies in multivariate outbreak detection, we favour, in this study, "outbreak prediction" based on suitably selected models as a significant step toward a model-based outbreak detection system. This alternative paradigm accommodates for historical outbreak and bases outbreak detection on the posterior distribution of suitable model parameters or on the predictive distribution. Evaluation of the one-step-ahead predictions can be used to assess if the predictions are well calibrated. In contrast, the quality of traditional outbreak detection methods is usually assessed by simulation based on artificially inserted outbreaks. While both outbreak detection and outbreak prediction have strengths and weaknesses for univariate time series, the benefits of a model-based predictive approach become overwhelming in multivariate surveillance. The challenges of multivariate surveillance, such as different data time lags, different frequencies of observations (e.g. weekly and daily) are also discussed.

Summers, Jennifer Ann : ***Characterising the provision of SeHCAT services in the United Kingdom: A multi-centre prospective survey***

Author list: *Summers, Jennifer Ann; Coker, Bola; McMillan, Viktoria; Ofuya, Mercy; Keevil, Stephen; Lewis, Cornelius; Peacock, Janet; Reid, Fiona*

Background

Bile Acid Malabsorption (BAM) is a cause of many diarrhoeal conditions. However, robust data on the prevalence of primary and secondary BAM does not exist in the United Kingdom (UK). A clinical diagnosis of BAM can be confirmed using SeHCAT (tauroselcholic [75selenium] acid), a radiolabelled synthetic bile acid. SeHCAT retention levels are assessed at two time points, and used to determine a BAM diagnosis, leading to potential treatment with bile acid sequestrants (BAS).

Methods

A prospective survey was conducted to characterise day-to-day practice associated with the clinical indications for referring patients for a SeHCAT test in the UK. This resulted in a dataset with more than 200 variables capturing centre and patient-level information. Eligible data from 38 centres and 1,036 patients were entered into a validated management system. Patients in the survey had a mean age of 50 years (range 6 -89 years), were predominantly female (65%) and the majority were listed as being white (77%).

Results

Clinicians recorded the suspected BAM type of patients before the SeHCAT scan. Type 1 (BAM secondary to ileal resection or ileal inflammation) was the smallest group (14%). For this subset of patients, 83% had Crohn's disease and 64% had an ileal resection.

The mean SeHCAT retention score for patients was 19%. However, this differed with suspected BAM type: Type 1 9%, Type 2 21%, Type 3 22%. A BAS prescription was not given to 27% of patients with a centre-defined 'abnormal' SeHCAT retention result.

The SeHCAT protocol varied between centres, for example, there was no standardised patient positioning, and a wide range of additional and alternative treatments post-SeHCAT were documented, regardless of whether or not a patient received a BAS prescription.

Conclusion

The survey has characterized the variability in centre provision of SeHCAT, patient clinical history, post-SeHCAT care pathway and results of the SeHCAT test. One of the most notable findings is the clear disparity in provision of both SeHCAT and BAS prescription across practices and amongst patients themselves.

Hayoz, Stefanie : ***Comparison of design options for phase IB clinical trials in oncology: simulation results.***

Author list: *Bigler, Martin; Hayoz, Stefanie; Xyrafas, Alexandros; Klingbiel, Dirk*

Background

Phase I trials play an essential role in the development of new drugs in cancer research. Many designs and many comparisons of these designs with simulations exist. To the best of our knowledge, no simulation results exist when there are only few dose levels to compare, as it is often the case in phase IB trials, where for example toxicity data in other indications are available.

Methods

We compared the 3+3 design, the continual reassessment method (CRM), the modified toxicity probability interval method (mTPI) and the rolling-6 design in a dose-finding study with only 2–4 dose levels of a drug. We explored the properties of the designs in several scenarios.

Results

The 3+3 design generally performed poorly.

When the prior probabilities for toxicity were close to the true values, CRM was the best method in terms of selecting the correct dose as maximum tolerated dose (MTD) and number of patients treated at the MTD. However, it suffered from misspecification of the priors.

In all scenarios the rolling-6 design required the shortest time to complete the trial.

The mTPI method was—depending on the design parameters—slightly better than the 3+3 and rolling-6 designs as it selected the correct dose more often and treated more patients at the MTD.

Conclusion

CRM has the best performance in terms of different metrics for phase IB trials in oncology given that the priors are well specified. As investigators seem to feel uncomfortable with this complicated method, there may be alternatives like mTPI to consider.

Summers, Jennifer Ann : ***Research protocol for a multi-centre study to investigate the diagnostic accuracy of the SeHCAT test in measuring bile acid malabsorption***

Author list: *Reid, Fiona; Peacock, Janet; Coker, Bola; McMillan, Viktoria; Lewis, Cornelius; Keevil, Stephen; Sherwood, Roy; Vivian, Gill; Logan, Robert; Summers, Jennifer*

Background

Bile acid malabsorption (BAM) is one possible explanation for chronic diarrhoea. BAM may be idiopathic, the result of surgery (ileal resection, cholecystectomy), or associated with other conditions such as irritable bowel syndrome or Crohn's disease. No 'gold standard' exists for clinical diagnosis of BAM, but response to treatment with a bile acid sequestrant (BAS) is often accepted as confirmation. The SeHCAT (tauroselcholic [selenium-75] acid) test delivers a radiolabelled synthetic bile acid and provides a diagnostic test for BAM, but its performance against 'trial of treatment' is unknown. Also, fibroblast growth factor 19 (FGF-19) and 7-alpha-hydroxy-4-cholesten-3-one (7
alphaC4) offer potential new biomarkers of BAM. This protocol describes a multi-centre study to evaluate the diagnostic accuracy of SeHCAT and two biomarkers in predicting BAM as assessed by 'trial of treatment'.

Methods

All eligible patients attending a gastrointestinal appointment will be invited to participate. On attending for the SeHCAT test, blood and faecal samples will be collected for analysis of FGF-19 by Enzyme-Linked Immunosorbent Assay, and for 7
alphaC4 and fractionated bile acids by liquid chromatography mass spectrometry. A capsule containing radiolabelled SeHCAT will be administered orally and a scan performed to measure SeHCAT activity. Patients will return on day seven to undergo a second scan to measure percentage SeHCAT retention. The

test result will be concealed from clinicians and patients. BAS will be dispensed to all patients, with a follow-up gastroenterologist appointment at two weeks for clinical assessment of treatment response and adherence.

The diagnostic accuracy of the SeHCAT test and biomarkers will be analysed at different thresholds using sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratios and area under the curve. Multivariable logistic and linear regression models will be used, respectively, to assess the association between continuous SeHCAT retention levels and presence, or clinical severity, of BAM after adjustment for confounders.

Challenges

Some forms of BAS treatment are unpleasant due to the texture and taste of the resin powder, which may negatively affect recruitment and treatment adherence. 'Trial of treatment' is not as 'golden' a standard as would be ideal, and itself warrants further study.

Lugo, Alessandra : ***Prevalence and determinants of overweight and obesity in Europe***

Author list: *Lugo, Alessandra; La Vecchia, Carlo; Decarli, Adriano; Gallus, Silvano*

Obesity is one of the leading causes of preventable morbidity and mortality from cardiovascular diseases, diabetes, cancer, and several other chronic diseases. Its prevalence nearly doubled worldwide over the last 3 decades and substantially increased also in several European countries over the past 2 decades. In Europe, only few population-based studies have been conducted on obesity in different countries at the same time using homogeneous methodologies.

The aim of this study was to provide updated information on prevalence of overweight and obesity in Europe, using data from a representative pan-European survey, conducted using a uniform protocol and comparable methods. The survey was conducted in 2010 in 16 European countries (Albania, Austria, Bulgaria, Czech Republic, Croatia, England, Finland, France, Hungary, Ireland, Italy, Latvia, Poland, Romania, Spain and Sweden), on a total of 14,685 adults aged >= 18 years providing information on self-reported height and weight.

We found that almost half of the interviewed European adults were overweight or obese (47.6% overall, 54.5% in men, 40.8% in women), and 12.8% (14.0% in men, 11.5% in women) were obese. Obesity prevalence was lower in western/southern Europe (11.1%), including Italy (7.6%), and in central/eastern Europe (12.4%), including Austria (10.5%), than in northern European countries (18.0%). The highest prevalence of obese participants was observed in Croatia (21.5%) and in England (20.1%). Prevalence of obesity significantly increased with age, and decreased with level of education. As compared to never smokers, obesity was less frequent in current smokers and more frequent in male, but not female, ex-smokers.

Although a favourable obesity pattern has been observed in those European countries as compared to USA, obesity rates vary greatly among countries and subgroups of population. Obesity was lower in Mediterranean countries, particularly in Italy and France. This could be explained by the dietary habits of these countries (adherent to the Mediterranean diet), which consume a higher amount of fruit, vegetables, fish, and a lower amount of meat, milk, sugar and soft drinks, as compared to northern and central/eastern European countries. Intervention to control obesity in Europe should focus on adults with lower socio-economic status and male ex-smokers.

Ciniselli, Chiara Maura : ***Influence of different level of haemolysis on microRNAs for the early diagnosis of colorectal cancer***

Author list: *Ciniselli, Chiara Maura; Pizzamiglio, Sara; Bottelli, Stefano; Zanutto, Susanna; Belfiore, Antonino; Gariboldi, Manuela; Verderio, Paolo*

Introduction: microRNAs (miRNAs) are small non-coding RNAs involved in the development of various cancers. From our discovery study we identified a set of plasma circulating miRNAs associated with colorectal cancer lesions (CRC). However, significant variation in miRNA expression can be related to haemolysis (rupturing of erythrocytes) occurring during sampling and handling procedures [1,2].

Materials and methods: we set-up an in-vitro controlled haemolysis experiment to investigate if miRNAs expression is influenced by different level of haemolysis. Red blood cells (%RBCs) were used to artificially introduce haemolysis in a heamolysis-free plasma sample (S12) by performing 10 serial dilutions (range: 0.002% - 1% v/v) and single RT-qPCR was used for miRNAs quantification. We evaluated the relevance

of the expression changes between sample S12 and the other samples (S11-S01) by computing, for each miRNA, the 95% Simultaneous Confidence Interval (SCI) [3] of the Relative Quantity (RQ) RQ=2-ΔΔCt, where ΔΔCt=ΔCt(Sx) – ΔCt(S12) and ΔCt= Ct(miRNA) - Ct(ref) and Ct(ref) is the average of the miRNAs identified as reference by our procedure [4]. By adopting the two-fold threshold rule, miRNAs were considered statistically influenced by haemolysis if the upper limit of the 95% SCI(RQ) was >= 2 or the lower limit of the 95% SCI(RQ) was =< 0.5.

Results and conclusion: miRNAs known as haemolysis-related in literature were confirmed also in our experiment whereas our reference miRNAs seem to be not influenced by haemolysis. Similar results were obtained for the majority of our candidate miRNAs suggesting those as potential haemolysis-independent biomarker for early detection of CRC.

Acknowledgements: This work was supported by grant from Associazione Italiana per la Ricerca sul Cancro (AIRC) (Grant No. 12162).

References

[1] Kirschner MB, et al. PlosOne. 2011; 6:e24145

[2] Appierto V, et al. Bioanalysis. 2014; 6:1215-26.

[3] Pizzamiglio S, et al. Oncol Rep. 2010;23:853-60

[4] Verderio P, et al. Anal Biochem. 2014. 461C:7-9

Ciniselli, Chiara Maura : ***Assessing reproducibility in the microRNA circulating biomarkers identification workflow***

Author list: *Bottelli, Stefano; Pizzamiglio, Sara; Ciniselli, Chiara Maura; Zanutto, Susanna; Gariboldi, Manuela; Verderio, Paolo*

Quantitative real-time polymerase chain reaction (qPCR) is the most commonly used tool to investigate microRNA (miRNA) expression and qPCR low-density arrays are increasingly being used as an experimental technique for both the identification of potentially relevant miRNAs and their subsequent validation. Due to the reduced number of miRNAs to be validated, this phase is generally performed on ad hoc customized cards for which a technical robustness is assumed similar to that of the high-throughput cards used during the discovery phase. Although a good level of intra/inter-reproducibility was reported for both assays, no information is available about their mutual reproducibility. By taking advantage of an ongoing project aimed at finding plasma circulating miRNAs for the early detection of colorectal cancer (CRC) we faced from a statistical point of view the transition from the discovery to the validation phase by evaluating the reproducibility between the high-throughput and the customized assay in terms of concordance [1]. Our preliminary results [2] showed a reproducibility between the two methods that was not fully satisfactory indicating the need of including in the high-throughput based miRNAs identification workflow an additional intermediate phase after the discovery and before the validation step in an independent series. The introduction of an additional step permits to verify the reproducibility of the assays and to correctly select miRNAs that will have more chances to succeed in the subsequent validation. The statistical approach we presented could be view as diagnostics tool to be used for the evaluation of the reproducibility in the biomarkers identification workflow.

Acknowledgements: This work was supported by grant from Associazione Italiana per la Ricerca sul Cancro (AIRC) (Grant No. 12162).

References

[1] Marubini E, Pizzamiglio S, Verderio P. Agreement between observers: its measure on a quantitative scale. Int J Biol Markers. 2005;20:73-8.

[2] Verderio P, Bottelli S, Ciniselli CM, Pierotti MA, Zanutto S, Gariboldi M, Pizzamiglio S. Moving from discovery to validation in circulating microRNA research. Int J Biol Markers. 2015 [Epub ahead of print].

Lugo, Alessandra : ***Prevalence and trends of different type of alcoholic beverages in Italy***

Author list: *Lugo, Alessandra; Asciutto, Rosario; La Vecchia, Carlo; Decarli, Adriano; Gallus, Silvano*

Alcohol consumption is one of the major avoidable risk factors for chronic diseases and injuries, responsible for 4% of total mortality and 5% of the global burden of disease. The type of alcoholic beverage consumed substantially varies according to the socio-cultural and territorial characteristics. In Italy, official sale data

show a dramatic fall in alcohol consumption over the last 5 decades (La Vecchia et al., 2014; Eur J Cancer Prev 23:319-22), but information on drinking patterns remains limited.

We analysed data from seven nationally representative surveys conducted in Italy in 2006-2010 and 2013-2014. The total sample included 21,416 participants aged >=15 years, with available information on weekly consumption of wine, beer and spirits.

Alcohol drinkers were 61.6% overall (78.4% in men and 46.1% in women) and 52.5% among the young (aged 15-24 years). Predominant wine drinkers -those most frequently consuming wine compared to other types of beverages- were 35.3%, beer drinkers were 11.1%, and spirit drinkers were 6.4%. The corresponding estimates for the young were 9.1%, 22.4% and 10.6%, respectively. According to age, a direct trend was observed for predominant wine drinking ($p<0.001$), while an inverse trend was found for predominant beer ($p<0.001$) and predominant spirit drinking ($p<0.001$). Prevalence of alcohol drinkers was significantly higher in intermediate (multivariate odds ratio, OR=1.25) and high education (OR=1.97) compared to low education, in northern Italy compared to southern Italy (OR=1.59), in current (OR=2.17) and ex-smokers (OR=1.46) compared to never smokers, and in normal weight compared to obese participants (OR=1.49). Between 2006 and 2014, the prevalence of drinkers decreased by 9% (from 65.1% to 59.0%) in the Italian general population. Over the same period, the per-capita consumption of total alcoholic beverages decreased by 21% (from 5.6 to 4.4 drinks/week), and of wine by 31% (from 3.9 to 2.7 drinks/week). The consumption of beer (0.9 drinks/week in both 2006 and 2014) and spirits (0.7 drinks/week in both 2006 and 2014) remained stable over time.

Over the last decade we observed a sharp decrease in consumption of wine, and consequently total alcohol drinking, in Italy, confirming data on sales.

Szymański, Andrzej : ***MFMM – Modified Fuzzy Mortality Model based on C*-Banach algebra***
Author list: *Szymański, Andrzej; Rossa, Agnieszka*

The problem of determining the best mortality models is one of the basic fields in the forecasting strategy of insurance companies. Private health insurance or medical expense insurance covers the unexpected expenses that incurred through the sickness of the insured. Actuaries try to predict future costs using demographic characteristics of the insured,

For a given age group x at year t the mortality rate m(x,t) can be expressed in the form of so-called Lee-Carter stochastic mortality model (LC) (1992).

Koissi and Shapiro (2006) formulated the fuzzy version of the LC model, where coefficients of the model have been assumed to be fuzzy numbers with the symmetric triangular membership function (STMF).

To make more precise and elegant inferences from the improved FLC Szymański and Rossa (2014) have applied the Banach algebra of fuzzy numbers, called OFN-algebra introduced by Kosiński et al.(2004), which allowed them to create a mortality model termed Extended Fuzzy Lee-Carter model (EFLC).

In the frame of the work with EFLC model we have observed that better results gives the substitution of Kosiński – Banach OFN algebra by the C* Banach algebra. The essential difference between OFN and C* algebra is in distinct definitions of algebras multiplication.Using C*- Banach algebra we can apply Mazur-Gelfand theorem giving us the isometric isomorphism between C* algebra and the complex analysis.

We will present on the poster the graphical illustrations of our model constructing and the obtained results for Polish data.

The research was supported for both authors by a grant from the National Science Center under contract DEC-2011/01/B/HS4 for which the authors are indebed.

Zolin, Anna : ***Risk factors for reduced pulmonary function in European cystic fibrosis patients***
Author list: *Zolin, Anna; Bossi, Anna; De Boeck, Kris*

Introduction A cross-sectional European Cystic Fibrosis Society Patient Registry (ECFSPR) study [1] demonstrated, using the 2007 data, an association between low lung function and low body mass index (BMI) as well as Pseudomonas aeruginosa (Pa) infection. The recent availability of longitudinal data allowed investigating the risk factors for reduced pulmonary function in a more reliable way.

Method We fitted multilevel models to the 2008-2010 longitudinal FEV1% of CF patients included in the ECFSPR. For each potential risk factor, we used a random intercept and slope model nested within country, and tested whether these factors exerted any effect on the intercept (basal FEV1%, i.e. FEV1%

value in 2008) or on the slope (decline of FEV1% over time). In the final model we included simultaneously all risk factors that, in the univariable models, had an effect on basal FEV1% or on FEV1% decline.

Results More than 10,000 patients had at least one FEV1 measurement in the years of follow up. In the final model, age, gender, genotype, pancreatic status, age at diagnosis, Pa infection, BMI, allergic bronchopulmonary aspergillosis (ABPA) and CF-related diabetes (CFRD) showed a statistically significant ($p<.0001$) effect on the basal FEV1%. Patients with at least one class 4-5 mutation have basal FEV1% by 6-7 percentage points lower than those with at least one nonsense mutation, and patients with a low BMI have basal FEV1% lower by 7 percentage points than those with normal BMI. Only meconium ileus at birth, Pa infection and BMI showed a significant ($p<0.05$) effect on the FEV1% decline.

Conclusion We confirm previous cross sectional findings and expand beyond these. In this dataset, some intrinsic factors like gender and genotype only impacted on the basal FEV1%. Several preventable factors like Pa infection and BMI impacted on the basal FEV1% and determined a faster decrease of FEV1% over time.

References 1 Kerem E, Viviani L, Zolin A et al., Factors associated with FEV1 decline in cystic fibrosis: analysis of the ECFS patient registry. Eur Respir J. 2014;43(1):125-33.

Rosso, Tiziana : ***Comparison of age-period-cohort models for the analysis of mortality rates***

Author list: *Rosso, Tiziana; Malvezzi, Matteo; Decarli, Adriano*

Age–period–cohort (APC) analyses are a family of statistical techniques to study temporal trends in terms of three related time variables: the subject's age (A), the calendar period (P) and the subject's birth cohort (C). APC analysis studies the effects of age, period and cohort simultaneously to disentangle their contributions to the studied outcome. The age, period and cohort variables have an exact linear dependence: A=P-C. This causes an identifiability problem. To overcome this issue, three APC analysis methods from the literature are examined.

Penalised likelihood APC method

This method identifies the solution that minimizes the Euclidean distance between the three two factor models (age-period, age-cohort, cohort-period) by weighing them by their goodness of fit [1].

Generalised additive models (GAM) APC method

Carstensen proposed to use natural splines to smooth the non linear curves of APC models using Holford's parametrization [2].

Partial Least Squares (PLS) APC method

PLS regression is used with a two-stage procedure [3]. First a factorial method is applied to obtain the PLS components. These are selected to maximize covariance between the outcome and the unobserved factors. Subsequently the unobserved factors are used as regressors. Finally these coefficients are transformed into the familiar parameters of age, period and cohort.

APC analysis is useful to study mortality data, particularly cohort effects, but should be used with caution. Penalised likelihood and GAM methods produce similar results, while the PLS method presents differences. The first two methods use different techniques to distribute the effect of the temporal linear drift between cohort and period factors to solve the identifiability problem, whilst the PLS method solves the problem minimizing the matrix of variances and covariances among the possible estimated parameters in the generalized inverse. From an empirical comparison of the models, we conclude that models based on drift distribution are adequate for epidemiological comparisons, where the problem lies mainly in disentangling the drift effect between cohorts and periods. The PLS model is interesting in projecting future rates.

References

1. Decarli A., La Vecchia C. Rivista Statistica Applicata 1987;20:397-410.
2. Carstensen B. Stat Med 2007;26(15):3018-45.
3. Fukuda K. Stat Comp Simulation 2011;81(12):1871–1878

Barbati, Giulia : ***Prognostic evaluation in palliative care: testing for improvement in prediction model performance***

Author list: *Giangreco, Manuela; Barbati, Giulia; Mazzer, Micol; Ermacora, Paola; Aprile, Giuseppe;*

Gregoraci, Giorgia; Sacco, Cosimo Stanislao; Puglisi, Fabio; Fasola, Gianpiero; Isola, Miriam

Introduction:

Healthcare professionals frequently use multidimensional prognostic scores to forecast the patients'outcome. The Palliative Prognostic (PaP) Score is the score most frequently used to forecast the 30 days-survival of patients.

Objectives

With the present cohort study we evaluated the improvement in prediction performance of a new model obtained by adding baseline covariates to the original model that includes only the multidimensional prognostic score.

Materials and Methods: A total of 245 cancer patients were enrolled from April 2011 to August 2014 and PaP Score was calculated. To evaluate the improvement in prediction performance of the score in prediction of 30 days-survival, univariate and multivariate logistic models were carried out. The baseline model included only PaP score while the complete model is obtained considering also metastasis situs, agitation, and adl. We estimated predictive performance for the baseline and enhanced models and the performance improvement.

Results: The AUC of the baseline model was 0.81 while those in the complete model was 0.85. In particular, setting the risk threshold so that 83% of cases haven't a 30 days-survival, the ratio between false positive and true positive in the baseline model was 0.21 while those in the complete model resulted 0.19. Finally the mean risk difference in the baseline model was 0.25 while those in the complete model was 0.33. The performance improvement was 0.04 for AUC, -0.02 for ratio between false positive and true positive and 0.08 for mean risk difference.

Conclusion: The estimates of prediction performance suggest there is a benefit in the prediction of PAP-score adding a several of clinically relevant covariates.

References:

1. Pepe M.S., Kerr K.F., Longton G., Wang Z. Testing for improvement in prediction model performance. Statistics in Medicine, 2013, 32:1467 – 1482.

2. Pencina M.J., D'agostino Sr R.B., D'agostino Jr R.B, Vasan R.S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics in Medicine, 2008, 27: 157 – 172.

Bravi, Francesca : ***Impact of maternal nutrition on breast milk composition: a systematic review***

Author list: *Bravi, Francesca; Dal Pont, Alessia; Wiens, Frank; Agostoni, Carlo; Decarli, Adriano; Ferraroni, Monica*

Maternal nutrition have been suggested to have an impact on nutrients excreted in breast milk. Several studies have considered different aspects of maternal diet and various types of nutrients. In order to summarize and better define the available knowledge on the issue, we carried out a systematic review of the publications investigating the role of maternal diet on breast milk composition. We performed a Pubmed/Medline search of the papers published up to January 2015. We reviewed the manuscripts to identify the eligible studies according to predefined criteria. Studies were included if: (1) provided quantitative information on the relationship between maternal diet and any nutrient in breast milk; (2) were based on original observational or experimental studies; (3) included healthy term infants and healthy mothers. Exclusion criteria were: (1) investigating the role of supplements or probiotics; (2) considering the transfer of pollutants, toxic metals or contaminants from maternal diet to milk; (3) including marginally nourished populations; (4) including children and/or mothers with health problems.

We identified 31 publications, including 1890 mother-child pairs. Twenty-four papers concerned observational cross-sectional studies investigating mothers' usual dietary habits, while 7 papers described experimental (mainly cross-over) studies in which maternal diet was controlled and pre-determined. In 20 studies mature milk was collected 1 to 12 months postpartum, in 5 studies colostrum was collected a few hours/days postpartum, in 4 studies both colostrum and mature milk was collected, and in 2 studies mature milk was collected up to 24 months postpartum. The studies were quite different both in terms of characteristics of mothers' diet considered and breast milk components examined. The majority of studies (17) focused on breast milk fatty acid composition, while the remaining investigated other milk

components including protein (4 studies), total fats (4 studies), vitamins (5 studies), minerals (3 studies), and total energy intake (2 studies). Among the considered breast milk components, polyunsaturated fatty acids, monounsaturated fatty acids, saturated fatty acids, polyunsaturated/saturated fatty acids ratio, DHA acid, linoleic acid, linolenic acid, eicosapentaenoic acid, docosapentaenoic acid, vitamin C, vitamin E, beta-carotene, iron, and zinc were significantly influenced by maternal nutrition.

Barbati, Giulia : ***Evaluating prognostic accuracy of multiple longitudinal predictors in a standard time-dependent Cox model framework***

Author list: *Barbati, Giulia; Merlo, Marco; Di Lenarda, Andrea; Sinagra, Gianfranco*

The use of longitudinal covariates in time-to-event data is a fundamental issue for improving survival prognostic accuracy. Recently, joint models for longitudinal and survival data have been proposed but at present in the R library "JMbayes" only one longitudinal parameter is allowed in the survival model. For this reason in the present work time-dependent (TD) Cox model has been used. The clinical aim of the study was to analyze longitudinal trends and prognostic power of the right ventricular dysfunction (RVD) in a Dilated CardioMyopathy (DCM) population. All patients in the Trieste Heart Muscle Disease Registry from 1993 to 2008 were included (657 pts with a median follow-up of 138 months). Mixed-effects generalized linear models (GLMM) were estimated to evaluate dynamics of RVD. Multivariable Cox regression was used to evaluate predictors of outcome at baseline and with TD covariates. Comparisons between probabilities from the baseline model and TD model were compared using TD ROC AUC and tested with an ad-hoc bootstrap procedure. To investigate the interaction between follow-up and the hazard ratios of the TD parameters, multiple Cox models stratified by the visit occasions were estimated. RVD had an early significant improvement (within 6-months), a period of stability within 48-month and subsequently a progressive worsening. Baseline Cox model showed that left ventricular ejection fraction (LVEF), duration of disease and age were independent predictors; RVD had a borderline significance and no significant effect at baseline was observed for New York Heart Association (NYHA) class. In the TD Cox model RVD was stronger than baseline and significant, taking into account also the significant time-varying effect of LVEF and NYHA class. Stratifying by follow-up, hazard ratios of RVD were more and more larger from 6 months later on. TD ROC showed a better accuracy of the TD Cox model than the baseline model at each follow up. Standing the limits of TD Cox in dealing with endogenous TD covariates future developments of the present work will be to include in the survival model estimates of longitudinal dynamics obtained from the GLMM models to evaluate if significant variations in the predictive accuracy will be observed.

Dunger-Baldauf, Cornelia : ***Safety analysis for individualized treatment***

Author list: *Dunger-Baldauf, Cornelia*

In individualized treatment, a well-known strategy aims to identify patient characteristics which predict response to a drug. Such factors may be difficult to establish in some disease areas. In such cases, to achieve individualized treatment, the treatment schedule can be adapted at each monitoring time point according to the observed response (individualized or as needed treatment). The treatment time points in this regimen have to be taken into account as random variables. In addition, few patients only might share the same treatment pattern and the associated expected response profile. For instance, let us consider an ongoing 12 months clinical trial to evaluate two regimens of a treatment of pathologic myopia. Following initial treatment, at each monthly visit the decision whether to retreat or not, is based on the patient's response. The number of treatments for a patient may vary between 2 and 12. Such individualized treatment patterns and response profiles over time pose challenges for modeling, parameter estimation and for statistical inference.

We present and discuss a model which addresses these issues for the evaluation of adverse events, illustrated by an example from ophthalmology.

Neubauer, Maria : ***A comparison of the logistic regression model and the random coefficients linear mixed model for analysis of disease progression***

Author list: *Neubauer, Maria; Heinze, Georg*

Urine albumin to creatinine ratio (UACR) is a continuous clinical marker for the progression of chronic kidney disease (CKD). Often, the marker is categorized into 'clinically relevant' categories and disease progression is then simply defined as a change in categories. Therefore, there are two possible approaches to model the probability of disease progression for a patient: either by logistic regression, using a binary outcome variable defining the status of a patient (progressing or stable), or by a longitudinal approach,

using the repeated measurements of UACR as outcome variable. The aim of this presentation is to compare these two approaches for the analysis of early kidney disease progression. We compared the power to detect an association of a risk factor (or treatment) with progression, and the accuracy in estimating progression probabilities. It is shown how quantities usually available only with binary outcomes, such as relative risks, odds ratios, concordance statistics and integrated discrimination improvements can be estimated by either approach, in simulated datasets as well as in a real-life example. All calculations were done using the original log UACR (LOG models) or a Box-Cox log transformation thereof (BCLOG models).

The simulation and the real-life data analysis both showed small but clear differences between the misspecified LOG models compared to BCLOG models. Since the assumption of normally distributed residuals in the longitudinal LOG models was violated, the LOG models underestimated the probability of progression. The longitudinal BCLOG model attained better results than the logistic BCLOG model, both in accuracy of the estimated progression probabilities and in power to detect an effect of a risk factor x on progression. However, assessing model improvement only lead to small differences between the logistic and longitudinal models. Summarizing, the longitudinal model on the Box-Cox transformed continuous disease marker should be the first choice for analysis of disease progression because of its superior accuracy and unbiasedness.

Wallisch, Christine : ***Chances and challenges of using routine data collections for health care research***

Author list: *Wallisch, Christine; Heinze, Georg*

Background: Collections of electronic medical records can provide a rich source of information for health care research. However, their use in statistical analyses requires many preparatory steps, including coding of freetext entries and clear definitions of time windows for harvesting prognostic factors and outcomes. We analyze a large collection of electronic medical records to identify prognostic factors of adequate health care in diabetic patients at risk for chronic kidney disease, and discuss benefits and risks of such re-use of routine data.

Methods: In a representative sample of 695,068 patient records collected in 58 Austrian general practicioners' offices, we could identify 31,374 patients with diabetes mellitus. As outcomes, we investigated whether a patient received a serum creatinine measurement, and the time elapsing between two consecutive serum creatinine measurements. Prognostic factors were defined by extracting previous diagnoses, laboratory measurements, drug prescriptions and demographic characteristics from the records. LASSO logistic and Cox regression were used for analysis.

Results: Serum creatinine was measured annually in 44.4% of diabetic patients with previous signs of reduced kidney function and in 20.5% of the patients without such signs. Within one year after the first measurement, a follow-up measurement was made in 79.4% and 68.4% of the patients, respectively Previous diagnoses, laboratory measurements, drug prescriptions and demographic characteristics explained 41% of the observed variance of kidney function monitoring. With 24% explained variance, previous referals to laboratories were identified as the most important prognostic factor group.

Conclusion: The analysis of large routine data collections poses various challenges, among which the need for coding free text into variables and various sources of biases are most demanding. However, routine data collections represent the daily practice of health care and offer many chances for health services and outcomes research.

Evidence Synthesis and the Use of Co-Data (CEN Invited Session)

Invited session
Wednesday, 17 June 2015
09:00 - 10:30
Room: U6-A10
Session chair: Held, Leonhard; Friede, Tim

Schmidli, Heinz : ***Robust meta-analytic-predictive priors in clinical trials with historical control information***
Author list: *Schmidli, Heinz*

The Bayesian framework allows us to use historical control information in the analysis of a clinical trial, and hence to reduce the number of subjects randomized to control. This decreases costs and trial duration, facilitates recruitment, and may be more ethical. Yet, under prior-data conflict, a too optimistic use of historical data may be inappropriate. We address this challenge by deriving a Bayesian meta-analytic-predictive prior from historical data, which is then combined with the new data. This prospective approach is equivalent to a meta-analytic-combined analysis of historical and new data if parameters are exchangeable across trials. The prospective Bayesian version requires a good approximation of the meta-analytic-predictive prior, which is not available analytically. We propose two- or three-component mixtures of conjugate priors, which provide good approximations and analytical posterior calculations. Moreover, since one of the mixture components is usually vague, mixture priors will often be heavy-tailed and therefore robust. Further robustness and a more rapid reaction to prior-data conflicts can be achieved by adding an extra weakly-informative mixture component. We illustrate the methodology for a randomized placebo-controlled phase II proof-of-concept trial in patients with ankylosing spondylitis, where placebo information from eight historical studies was used.

References

Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics 70(4):1023-1032.

Baeten D et al. (2013) Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. Lancet 382(9906):1705-1713.

Copas, John Brian : ***The use of secondary outcomes in meta-analysis***
Author list: *Copas, John Brian*

Meta-analysis is usually univariate, estimating the mean treatment effect for the outcome of primary interest. However, many clinical studies will also measure secondary outcomes. Multivariate meta-analysis allows us to take these secondary outcomes into account, and also allows us to include studies where the primary outcome is missing. How much is gained by taking these secondary outcomes into account? The talk will show how the contribution of each study's secondary outcomes depends on the level of consistency between the research designs of all the studies in the meta-analysis. A simple graphical interpretation will be suggested, and illustrated using a recent meta-analysis of ten anti-hypertension clinical trials.

Röver, Christian : ***Meta-analysis of few small studies in small populations and rare diseases***
Author list: *Röver, Christian; Neuenschwander, Beat; Wandel, Simon; Friede, Tim*

The between-study heterogeneity plays a central role in random-effects meta-analysis. Especially when the analysis is based on few studies, which is a common problem not only for rare diseases, external a-priori information on heterogeneity may be helpful. In case of little information, the use of plausible weakly informative priors is recommended. Computational simplifications (using the bmeta R package) helped to speed up computations for Bayesian standard random-effects meta-analysis to explore the frequentist properties of Bayesian estimators for different priors. We investigated a range of scenarios (heterogeneities, numbers of studies), to compare bias, MSE and coverage of the Bayesian and classical estimators. The different approaches are illustrated using an application in pediatric transplantation.

Held, Ulrike : ***Improvements in meta-analysis of diagnostic test accuracy in the absence of a perfect reference test***

Author list: *Held, Ulrike; Brunner, Florian; Steurer, Johann; Wertli, Maria M.*

Bone scintigraphy (BS) is a nuclear scanning test that is discussed as a diagnostic tool for complex regional pain syndrome (CRPS 1). Earlier meta-analyses of test accuracy of BS were impeded by the use of different and imperfect reference tests across studies. The aim of our study is to summarize sensitivity and specificity of BS for CRPS 1 and to identify factors to explain heterogeneity. We use a hierarchical Bayesian approach to model test accuracy and threshold. We present different models accounting for the imperfect nature of the reference tests, and assuming conditional dependence between BS and the reference test results. Further, we include duration of symptoms as explanatory variable in the model. A common problem in meta-analyses is the unavailability of certain explanatory variables in some of the included studies. We present a novel approach and impute the missing data within the Bayesian framework. The models are compared using summary ROC curves and the deviance information criterion (DIC). Our results show that those models, which account for different imperfect reference tests with conditional dependence and inclusion of the covariate are the ones with the smallest DIC. The sensitivity of BS is 0.87 (95% credible interval 0.73-0.97) and the overall specificity is 0.87 (0.73-0.95) in the model with the smallest DIC. The estimated effect of duration of symptoms on the threshold parameter is 0.17 (-0.25-0.57). We demonstrate that the Bayesian models presented in this talk are useful to address typical problems occurring in meta-analysis of diagnostic studies, including conditional dependence between index test and reference test, as well as missing values in the study-specific covariates.

Recent Advances in ROC Methodology (Invited Session of the Italian-Spanish-ERM Regions)

Invited session
Wednesday, 17 June 2015
09:00 - 10:30
Room: Aula Martini
Session chair: Antolini, Laura; Reiser, Benjamin

Calle, M.Luz : ***Identification of high-order interactions using the likelihood-ratio score optimal ROC curve***
Author list: *Calle, M.Luz; Urrea, Víctor*

Many statistical methods for epistasis analysis have been proposed that are able to scan for second order interactions or, at most, third order interactions. Since they are very computationally demanding, they become infeasible for exploring higher order interactions when the number of variables to explore is large, as it is usually the case.

We present the "Optimal AUC algorithm", a new strategy for exploring the joint predictive effect of a set of genetic factors, including their possible interactions, in the context of a case-control study. The method is based on the likelihood ratio score. Given a set of predictors and a binary dependent variable, the likelihood ratio score provides an optimal prediction of the outcome in the sense that it provides the largest discrimination accuracy among all possible risk scores derived from the set of predictors.

The Optimal AUC algorithm follows a forward selection process to obtain the subset of factors with the highest joint predictive accuracy. The algorithm is computationally feasible for exploring high order interactions in a large number of variables context.

Barrio, Irantzu : ***Categorization of continuous predictors in the development of prediction models by maximization of the AUC***
Author list: *Barrio, Irantzu; Arostegui, Inmaculada; Rodriguez-Álvarez, María Xosé; Quintana, José María*

Prediction models are nowadays relevant for decision making in various fields, so

they are in medicine. In the development of clinical prediction models it is common to use categorical variables as predictors, although from a statistical point of view, it is hardly recommended since categorizing may bring a loss of information and power. Previous work on categorization of continuous predictors has been done although most of these approaches seek for a unique cut point. However, in many situations more than two categories may be needed.

One of the most commonly used prediction models is the logistic regression, hence we thought about categorizing a continuous predictor in the logistic regression framework. Consider we have a dichotomous response variable Y and a continuous covariate X which we want to categorize. The area under the ROC curve (AUC) is the most widely used discriminatory ability measure in logistic regression. For this reason, we considered categorizing X in such a way that the best logistic regression model, maximal AUC, was obtained for Y. To do so, we propose two alternative algorithms to look for the vector of the optimal k cut points, respectively denominated AddFor and Genetic. With the former we look for each cut-point at a time while the latter simultaneously looks for the vector of k cut points using Genetic Algorithms, the most widely known type of evolutionary algorithm. Finally, we propose a novel approach to select the optimal number of cut points. This is based on a bootstrap confidence interval for the bias corrected AUC difference for two categorized versions of the continuous predictor X.

The proposed methodology was applied to the IRYSS-COPD study, a prospective cohort of patients with a Chronic Obstructive Pulmonary Disease (COPD). The aim was to categorize the covariate arterial blood gas PCO2 in a univariate and multivariate settings. We found out that 2 were the optimal number of cut points obtaining almost the same location when the Genetic or the AddFor algorithms were used. The results obtained were face validated by clinicians.

Nakas, Christos T : ***Developments in ROC surface analysis***
Author list: *Nakas, Christos T*

ROC surface analysis can be employed for the assessment of diagnostic markers in three-class classification problems, e.g. for progressive disorders when clinicians and researchers want to diagnose or classify subjects as members of one of three ordered categories based on a continuous diagnostic marker. Current state of the art of ROC surface analysis is presented and the notion of the generalized Youden index is reviewed. The generalized Youden index is an optimality criterion for the selection of the optimal cut-off points that will be used in practice for decision making. The construction of joint confidence intervals for the resulting True Class Fractions is also discussed and an illustration on a pancreatic cancer diagnostic marker is presented. The use of R-packages and implementation of the techniques is illustrated. The presentation ends with a discussion on future directions for research in this field.

Antolini, Laura : ***Measuring prediction improvement of additional markers: strengths and limitations of the commonly used indexes***

Author list: *Antolini, Laura; Davide, Bernasconi; Valsecchi, Maria Grazia*

The availability of novel biomarkers in several branch of medicine opens room for refining prognosis by adding factors on top of those having an established role. It is accepted that additional factors should not be assessed in their impact relying solely on regression coefficients and their significance. This motivated the fruitful literature appeared in the last decades proposing predictive power indexes, such as ROC based quantities and Brier Score measures and related inference.

A common issue the applied biostatistician faces in practice is that novel factors who are promising at the explorative stage, often results in disappointingly low (or no) impact in the predictive power and without reaching significance. This aspect, which could be intuitively due to real difficulties in improving prediction upon established factors, was thought as: i) due to an insensitivity nature of the aforementioned measures, ii) possibly related to the issue of testing for predictive power increments.

Novel measures, namely net reclassification index and integrated discrimination improvement, where proposed as direct measures of gain due to additional factors based on the intuitive concept of reclassification table. Although the seminal paper appeared only a few years ago, it became extremely popular gaining most traction in cardiovascular disease and even in cancer.

However, several recent contributions appeared in the biostatistical literature to enlighten strong limitations of the reclassification indexes, starting from the apparent easy interpretation, to the tendency to indicate advantageous models obtained adding unrelated factors in simulations. In addition, theoretical literature on testing recommends of not testing on predictive power increments.

These findings were somehow unexpected given both the straightforward relationship the reclassification indexes at first glance with the concept being investigated, and the natural tendency to proceed on testing to reveal predictive power increments. The presentation will review the reclassification table methods together with the main criticisms and the issue of testing. Some alternative promising approaches will be also mentioned.

Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. Stat Med. 2014 Aug 30;33(19)

Pepe MS, Janes H, Li CI. Net risk reclassification p values: valid or misleading? J Natl Cancer Inst. 2014 Apr;106(4)

Pencina MJ1, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008 Jan 30;27(2)

Statistical Methods in Infectious Disease

Contributed session
Wednesday, 17 June 2015
11:00 - 12:00
Room: U6-A11
Session chair: Frigessi, Arnoldo

Corbella, Alice : ***Deterministic modelling of infectious diseases: monitoring H1N1 virus in UK***

Author list: *Corbella, Alice; Birrel, Paul Jeffrey; Boddington, Nicki; Pebody, Richard; Presanis, Anne Margaret; Zhang, Xu-Sheng; De Angelis, Daniela*

In August 2010, after the influenza A/H1N1 outbreak in 2009, WHO announced the beginning of a post pandemic phase and encouraged development of monitoring and surveillance tools at national level.

UK Severe Influenza Surveillance System (USISS) [1] is a hospital based surveillance scheme for severe cases of influenza, through which all the severe cases admitted to Intensive Care Units (ICU) and High Dependency Units (HDU) in all NHS trusts are registered. Further, a sentinel scheme [2], run over a sample of trusts, collects individual level data on patients admitted to ICU/HDU and aggregate counts of cases admitted at all levels of care.

This work investigates the use of USISS data to estimate and monitor disease transmission.

Following from previous work on Influenza pandemic [3] a deterministic transmission model is formulated to approximate the disease dynamics in the population. An observational model links the weekly number of new ICU admissions to the transmission model. The aim of the analysis is the estimation of transmission parameters and, consequently, of the basic and net reproduction numbers R0 and Rn. The analysis is conducted within a Bayesian framework which combines, through hierarchical modelling, information from different sources and appropriately deals with missing information on the transmission processes.

Preliminary results are obtained by assuming very informative priors on all the parameters except for the infection rate. From season 2012/13, Rn is estimated to be 1.0430 (CrI: 1.0347 – 1.0513).

Many challenges are still to be tackled, in particular: the problem of immunity and endemicity (since influenza is seasonal) and the issue of more general priors. A stochastic version of this model is also investigated.

References

1. PHE, UK Severe Influenza Surveillance System (USISS) Protocol for all NHS Acute Trusts 2011-12 (2011).

2. PHE, UK Severe Influenza Surveillance System (USISS) Protocol for sentinel Acute NHS Trusts 2011-12 (2011).

3. Birrell, P. J. et al. (2011). Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. In: PNAS, November 8, 2011, vol. 108, no. 45, pages 18238–18243, doi: 10.1073/pnas.1103002108.

Herzog, Sereina : ***Can mathematical models help to improve sample size calculations for infectious disease trials?***

Author list: *Herzog, Sereina A.; Berghold, Andrea*

An important part at the planning stage of a randomized controlled trial (RCT) is the estimation of the required sample size. Planning the sample size for a trial in the field of infectious diseases raises special issues: the intervention (e.g., treating an infected person) affects not only the individual but also the people around him. This means that individuals in an infectious disease trial are not per se independent of each other. The direct effect of an intervention protects the persons receiving an intervention whereas the indirect effect impacts their social environment as a result of changes in the intensity of transmission in the population. Mathematical models can be used to investigate the contributions of both, direct and indirect effect on the hypothesized effect size over time. We investigated what insights can be gained in planning infectious disease trials with mathematical models when two treatments are compared.

We used a Susceptible-Infected-Recovered (SIR) compartmental model to imitate a two-armed RCT structure including compartments for the background population. In the model we assumed a density-dependent

transmission, i.e. the per capita force of infection increases with the density of infected. The reduction in the mean duration of infection was the effect of interest. The required sample size was calculated using the two group t-test of equal means.

Simulation studies are ongoing considering among other factors the influence of delay in recruiting infected individuals at the start of the infection and the ratio between the background population size and the required sample size.

Brizzi, Francesco : ***Estimating age specific HIV incidence using back-calculation***

Author list: *Brizzi, Francesco; Birrell, Paul Jeffrey; Delpech, Valerie; Brown, Alison; Gill, Noel; De Angelis, Daniela*

Monitoring the evolution of the HIV epidemic is of crucial interest for public health purposes. In [1] a Bayesian back-calculation model is proposed, based on the number of HIV and AIDS diagnoses over time, augmented with the information on CD4 cell count around diagnosis. The epidemic is described by three distinct components: infection, disease progression and diagnosis, which are combined into a population level multi-state model. This model allows estimation of both HIV transmission and testing rates amongst infected men who have sex with men (MSM) in England and Wales. Results show steady sustained transmission and increased diagnoses rate over the recent past.

However, a more targeted public health approach to reduce HIV transmission needs a better characterisation of the groups at risk of infection. We, therefore, extend the model to derive age-specific estimates of HIV incidence. This is a novel approach, generalising the age-specific formulation in [2] to a multi-state back-calculation, allowing estimation of time-dependent diagnosis probabilities. Time and age-specific HIV incidence is modelled non-parametrically through a bivariate spline. The choice of the type of spline to adopt and of the location of the underlying knots pose important challenges, as results are sensitive to these choices.

Data are provided by Public Health England on the age and year of diagnosis (and on CD4 count around diagnosis, when available) for approximately 56,000 MSM in England and Wales. Initial results, obtained using penalized maximum likelihood, show a steady increase in the probabilities of diagnosis and sustained, steady, overall HIV incidence. However, there are contrasting trends across age-groups: incidence appears to be increasing in the age-group between 15 and 35 years, while decreasing trends are estimated for older individuals.

References:

[1] Birrell, Paul J., et al. "HIV incidence in men who have sex with men in England and Wales 2001–10: a nationwide population study." The Lancet Infectious Diseases 13.4 (2013): 313-318.

[2]Marschner, Ian C., and Ronald J. Bosch. "Flexible assessment of trends in age-specific HIV incidence using two-dimensional penalized likelihood." Statistics in medicine 17.9 (1998): 1017-1031.

Methods and Applications in Genetics/Omics (2)

Contributed session
Wednesday, 17 June 2015
11:00 - 12:00
Room: Aula Martini
Session chair: van Wieringen, Wessel N.

Finos, Livio : ***A conditional multivariate score test for RNA-seq data***
Author list: *Finos, Livio; Maragoni, Lorenzo; Risso, Davide*

Conditional (i.e. permutation-based) tests are a well-established approach with relevant inferential properties. These methods provide their best performances for data that are sampled under randomization. On the contrary, the biggest obstacle to a wide use of them in practice is the inability to deal with covariates that affect the response but are not under test (i.e. nuisances parameters in the parametric approach). In these cases only ad hoc solutions can be adopted (e.g. the use of strata when the covariates are few and all categorical). Several works address general approaches to test the coefficients in the (multivariate) linear models, while an approach for the multivariate generalized linear models (MGLM) is still lacking. In this contribution, we present a score test to be applied to the class of MGLMs and to more general models. We show its usefulness through an application to a public RNA-seq dataset obtained following object location memory (OLM) in the mouse hippocampus. In this kind of data, the overdispersion of the counts is a known problem and the Poisson models are proved to fail. More complex models (e.g. negative binomial) are used instead. Since the proposed method computes the distribution of the test statistic (under the null hypothesis) through permutations, it does not suffer from this problem and, in many cases, it is robust even when a wrong data-generating model is fitted. This can be a very relevant property, especially when the number of (univariate) models to be checked are thousands or even more. Among the other features, the proposed test deals efficiently with the dependence of univariate tests. Therefore, it becomes trivial to perform multiplicity correction procedures such as the permutation-based Westfall&Young min-p over large datasets.

Assi, Nada : ***A PLS model for the meet-in-the-middle approach using metabolomics***
Author list: *Assi, Nada; Viallon, Vivian; Ferrari, Pietro*

The availability of metabolomics data was made possible after recent developments in nuclear magnetic resonance (NMR) and mass spectrometry, and resulted in massive sets of input to analyse. New methodologies are increasingly sought to explore the insights into pathological processes that metabolomics may provide, to better understand determinants of disease development.

While quantitative statistical methods are available to process information on metabolomic profiles variability, simultaneously relating large datasets is far from being an obvious pursuit. Such methods are needed, particularly in epidemiological studies, where metabolomics can be used to identify biomarkers associated with lifestyle exposures and risk of patho-physiological conditions.

This is the rationale of the "meeting-in-the-middle" (MIM), for which an analytical framework was developed here centred on the application of Partial Least Square (PLS). PLS is tailored for this purpose, as it generalizes features of PCA and multiple linear regression, by iteratively extracting components maximizing the covariance between two sets of variables, X and Y, known as predictors and responses, respectively. The resulting PLS scores are then related to an outcome variable Z in appropriate generalized linear models. Finally, the mediating role of the Y-scores in the association between the X-scores and the outcome risk is assessed in mediation analysis.

This analytical strategy was applied in a nested case-control study on hepatocellular carcinoma (HCC) within the European Prospective Investigation into Cancer and nutrition (EPIC), where serum 1H NMR spectra (800 MHz) were acquired for 114 cases and 222 matched controls. Through PLS analysis, 21 lifestyle variables (X-set, including information on diet, anthropometry and lifestyle) were linked to 285 metabolic variables (Y-set), and three PLS factors were extracted. The resulting scores were related to HCC risk in conditional logistic regression models.

In this application, we faced challenges related to i) the identification of the sources of systematic variability within the data, ii) the incumbent choice of normalization, and iii) the control for confounding in the mediation analysis step.

This study devised a way to bridge lifestyle variables to HCC risk through NMR metabolomics data. This implementation of the MIM finds applications in high-dimensional settings, increasingly frequent in the –omics generation.

Calza, Stefano : ***Integration of somatic mutation, expression and functional data reveals potential driver genes predictive of breast cancer survival***

Author list: *Suo, Chen; Hrydziuszko, Olga; Lee, Donghwan; Pramana, Setia; Saputra, Dhany; Joshi, Himanshu; Calza, Stefano; Pawitan, Yudi*

Motivation: Genome and transcriptome analyses can be used to explore cancers comprehensively, and it is increasingly common to have multiple omics data measured from each individual. Furthermore, there are rich functional data such as predicted impact of mutations on protein coding and gene/protein networks. However, integration of the complex information across the different omics and functional data is still challenging. Clinical validation, particularly based on patient outcomes such as survival, is important for assessing the relevance of the integrated information and for comparing different procedures.

Results: An analysis pipeline is built for integrating genomic and transcriptomic alterations from whole-exome and RNA sequence data, and functional data from protein function prediction and gene interaction networks. The method accumulates evidence for the functional implications of mutated potential driver genes found within and across patients. A driver-gene score (DGscore) is developed to capture the cumulative effect of such genes. To contribute to the score, a gene has to be frequently mutated, with high or moderate mutational impact at protein level, exhibiting an extreme expression and functionally linked to many differentially expressed neighbors in the functional gene network. The pipeline is applied to 60 matched tumor and normal samples of the same patient from The Cancer Genome Atlas breast-cancer project. In clinical validation, patients with high DGscores have worse survival than those with low scores ($P = 0.001$). Furthermore, the DGscore outperforms the established expression-based signatures MammaPrint and PAM50 in predicting patient survival. In conclusion, integration of mutation, expression and functional data allows identification of clinically relevant potential driver genes in cancer.

Analysis of Registries and Administrative Databases

Contributed session
Wednesday, 17 June 2015
11:00 - 12:00
Room: U6-A10
Session chair: Baccini, Michela

Geroldinger, Angelika : ***Combining information from multiple routine data collections to estimate the prevalence of chronic kidney disease among diabetic patients in Austria***
Author list: *Geroldinger, Angelika; Hronsky, Milan; Oberbauer, Rainer; Heinze, Georg*

Using routine data collections in empirical research can be an economical and reliable alternative to prospective surveys or studies. As with all retrospective studies, where the data have not been collected with the study objective in mind, routine data collections often do not cover all aspects of the research question. In this situation, combination of complementary data sources, possibly concentrating on different subjects, might be a remedy, but is methodologically challenging.

This presentation illustrates this idea by explaining how data on serum creatinine level measurements and data on hospital admissions with only few overlapping subjects were combined to deduce estimates of prevalences for various stages of chronic kidney disease (CKD) among the Austrian diabetic population. We had access to a comprehensive data base on drug prescription data, holding data on almost all Austrian diabetics, which could be linked exactly to data from a single large laboratory and probabilistically to hospital discharge data. First, the serum creatinine levels which are needed to determine the degree of kidney damage in a patient and other laboratory variables were modeled in the subsample of patients from the laboratory using high-dimensional drug prescription patterns as covariates. These models were then used to impute serum creatinine levels in the full Austrian diabetic population using multiple imputation by chained equations. Second, based on hospital discharge diagnoses which were available for hospitalized diabetic patients and again using high-dimensional modeling, we assigned a ratio of probabilities for CKD versus acute kidney injury to each Austrian diabetes patient. Determining the stage of kidney disease from the serum creatinine levels and using the ratio of probabilities for CKD versus acute kidney injury to discount patients suffering from acute but not chronic disease, age and sex specific prevalence estimates for CKD in the Austrian diabetic population could be estimated. These estimates compared well with surveys conducted in other countries.

Our experience shows that health care claims databases, as run in public health systems, can provide valuable information for extrapolating associations observed in subpopulations for which additional data, e.g. on measurements of laboratory parameters or hospital diagnoses, are available to total populations.

Banchelli, Federico : ***Model-based recursive partitioning applied to the Emilia-Romagna Region Hospital Discharge data.***
Author list: *Banchelli, Federico; Miglio, Rossella; Verdini, Eleonora*

The statistical analysis of hospital discharges administrative databases may pursue several targets. One common aim is to partition the data into groups that are homogeneous with respect to hospital resource consumption, as well as with respect to a clinical outcome, in order to assist Health Care organization and hospital planning. Actually, many casemix classification systems have been partly developed following this criterion.

The present paper explores the use of some recursive partitioning statistical techniques which are suitable in order to fulfill this goal, by analyzing a subset of Emilia-Romagna Region Hospital Discharge data (SDO).

Length of hospital stay (LOS) was selected as a reasonable proxy of hospital resource consumption and therefore used as dependent variable. For each record, by means of a clinical coding scheme, some key clinical explanatory variables were derived from diagnoses and procedures codes reported in the SDO, which follows the International Classification of Diseases, 9th version with clinical modification (ICD9-CM). Basic demographic and administrative data was also considered. The skewed non-normal distribution of LOS, the multi-centric collection of the data and the presence of outlier observations, suggested the use of specific Data Mining procedures.

Recent advances in recursive partitioning were therefore considered, such as the model-based recursive partitioning (MOB) method of Zeileis et al. [1]. Within MOB, several parametric models could be fitted in order to improve the partitioning performance. In the present work, two major regression approaches are examined, according to the literature on the analysis of LOS: 1) regression modeling of count data and 2) regression modeling of time-to-event data (survival analysis).

We will consider model-based partitioning methods based on survival analysis flexible parametric regression models, as well as on count data regression models which account for overdispersion. In the latter, inclusion of a zero-inflation component will also be studied.

Results from the different model-based procedures will be presented, together with a discussion on advantages and disadvantages of the use of these methods in the analysis of hospital discharge data.

References

[1] Zeileis A., Hothorn T., Hornik K. (2008). Model-Based Recursive Partitioning. Journal of Computational and Graphical Statistics, 17(2): 492-514.

Småstuen, Milada Cvancarova : ***Challenges when analyzing large registry based data - Results from the IBSEN (Inflammatory Bowel South-Eastern Norway) study.***

Author list: *Småstuen, Milada Cvancarova; Hovde, Øistein; Moum, Bjørn*

Background:

Advantages in medical research have led to longer patient survival, however at the cost of side effects, some of them emerging only long time after the date of diagnosis or end of treatment. It is therefore of great importance to compare patients or survivors with the general population to identify possible side effect. In Norway, all cancer patients are reported to the Cancer Registry of Norway by law and all causes of death are reported to the Death registry. Moreover, each individual born in Norway is given a unique identification number at birth which makes it possible to link all the registries established in Norway. The choice of design for comparison of a given cohort of cases with individuals from the general population has long been debated. Therefore our aim was to test several designs and compared the results on real life data. We modelled cancer mortality in a well-defined population-based cohort of IBD pts 20 years after diagnosis and to compare the risk of cancer development in IBD pts compared to the general population.

Methods:

The IBSEN study has prospectively followed all patients diagnosed with IBD from 1990 to 1994, in total 843 IBD patients followed regularly for years. All IBD patients were age-and gender matched with 2, 5 and 25 individuals from the general population (controls). Complete data on death and all cancers in the IBD cohort and in the controls were collected from the Norwegian Cancer Registry. We have fitted Cox regression models and competing risk models using 2, 5 and 25 controls and compared the results.

Results:

In total we have analyzed 843 IBD patients and 20950 controls. There were 117 (14%) cancers in the IBD patients and 2095 (10%) in the controls. Overall, IBD patients had 1.5 times higher risk of cancer development compared to their matched controls (HR = 1.52 (95 % CI (1.26- 1.83) $p<0.001$). There was no difference in all cancers risk between the genders.

Conclusions:

The results were similar regardless the choice of number of controls, with little gain in precision when increasing the number of from 5 to 25 controls.

General Assembly of IBS Italy

Wednesday, 17 June 2015
12:15 - 13:30
Room: U6-A10
Session chair: Valsecchi, Maria Grazia

ROeS General Assembly

Wednesday, 17 June 2015
12:15 - 13:30
Room: U6-A11
Session chair: Held, Leonhard

Robust Methods in Biostatistics

Invited session
Thursday, 18 June 2015
09:30 - 11:00
Room: U6-A10
Session chair: Farcomeni, Alessio

Discussant: Marco Alfò

Greco, Luca : ***A tutorial on the weighted likelihood methodology***
Author list: *Greco, Luca*

A weighted likelihood is characterized by a set of weights, depending on the observed data and the specified model, that are aimed at down weighting those likelihood contributions stemming from outliers in the sample at hand. Outliers are defined as observations that are unlikely to occur under the assumed sampling model. Therefore, robust inferential procedures, that are resistant to the occurrence of outliers and allow to detect them, can be developed based on the weighted likelihood. Under the assumed model, the weighted likelihood shares the main (asymptotic) features of the genuine likelihood function. Then, it is possible to obtain a weighted likelihood ratio test with the standard asymptotic distribution. This feature is particularly attractive in those cases in which robust Wald-type tests do not provide adequate coverage accuracy. Moreover, the weighted likelihood can be used in a Bayesian framework in place of the genuine likelihood in the Bayes' formula. This methodology leads to proper posterior distributions and reliable posterior inferences also in the presence of anomalous data. The use of the weighted likelihood will be illustrated through several examples in different settings.

Perrotta, Domenico : ***Monitoring robust regression***
Author list: *Perrotta, Domenico; Riani, Marco; Cerioli, Andrea; Atkinson, Anthony*

Robust methods are little applied (although much studied by statisticians). We monitor very robust regression by looking at the be- haviour of residuals and test statistics as we smoothly change the robustness of parameter estimation from a breakdown point of 50% to non-robust least squares. The resulting procedure provides insight into the structure of the data including outliers and the presence of more than one population. Moni- toring overcomes the hindrances to the routine adoption of robust methods, being informative about the choice between the various robust procedures. Methods tuned to give nominal high efficiency fail with our most compli- cated example. We find that the most informative analyses come from S estimates combined with Tukey's biweight or with the optimal p functions. For our major example with 1,949 observations and 13 explanatory vari- ables, we combine robust S estimation with regression using the forward search, so obtaining an understanding of the importance of individual obser- vations, which is missing from standard robust procedures. We discover that the data come from two different populations. They also contain six outliers. Our analyses are accompanied by numerous graphs. Algebraic results are contained in two appendices, the second of which provides useful new results on the absolute odd moments of elliptically truncated multivariate normal random variables.

Mayo-Iscar, Agustin : ***A joint application of trimming and constraints for robustifying the estimation of mixtures of factor analyzers***
Author list: *García Escudero, Luis Angel; Gordaliza, Alfonso; Gresselin, Francesca; Ingrassia, Salvatore; Matran, Carlos; Mayo-Iscar, Agustin*

Mixtures of factor analyzers are a flexible way for modelling multivariate data with local dependences. It is well known that trimming and restrictions are useful tools for getting robust estimators when estimating mixture models. It will be introduced a new robust estimator for estimating a mixture of factor analyzers based in the joint application of these mentioned tools. The asymptotic and robustness properties of the proposed methodology will be shown. A feasible AECM algorithm for this estimator will be provided. There will be shown evidences about the effectiveness of this methodology when it is applied to real and to artificial data sets.

Recent Developments to Tackle Time-Dependent Covariates in Clinical Research

Invited session
Thursday, 18 June 2015
09:30 - 11:00
Room: Aula Martini
Session chair: Mittlboeck, Martina

Discussant: Per Kragh Anderson

Galimberti, Stefania : ***Multiple time scale models for the evaluation of time-dependent treatments***
Author list: *Rebora, Paola; Galimberti, Stefania*

Models accounting for multiple timescales were shown to be advantageous for the evaluation of the impact of an intermediate event on time-to-event outcomes. We will show their usefulness also in the presence of an intermediate intervention, i.e. a treatment that changes at different times in the disease course.

We have considered a piece-wise Poisson model that is able to account for the effect of multiple timescales and directly evaluates their impact on the event occurrence by the likelihood ratio test. This model estimates the hazard functions in a flexible way, acknowledging the 'ground' against which the relative hazard effects are estimated, thus improving the description of the process under evaluation. It also provides a valid prediction tool for counselling patients on their prognosis. An alternative strategy to survival estimation, that is valid under certain conditions and has some advantages over the standard landmark approach by Simon and Makuch, has also been proposed.

We will show the advantages of this approach on the comparison of chemotherapy versus transplant in children with high-risk acute lymphoblastic leukemia in first remission.

Pötschger, Ulrike : ***Baseline covariate adjustment for generalised pseudo-values when the effect of a time dependent intervention is assessed***
Author list: *Pötschger, Ulrike; Heinzl, Harald; Valsecchi, Maria Grazia; Mittlböck, Martina*

When cumulative treatment effects, such as long-term survival probabilities or long-term cumulative incidences are of primary interest, the pseudo-value regression technique provides a new and attractive alternative to methods commonly applied in survival analysis. For the investigation of a time-dependent intervention, a generalisation of the pseudo-values regression model technique has been proposed recently. While adjusting for waiting time bias, this generalisation evaluates long-term survival rates without relying on proportional hazards.

For an unambiguous interpretation of the results, a proper adjustment for baseline covariates is essential. The aim is to estimate and compare the cumulative hazards (incidences) or survival probabilities since time 0 with and without the intervention, adjusting for additional baseline covariates. Conditional unbiasedness of resulting generalised pseudo-values and corresponding model results is explored.

A typical example occurs in childhood leukaemia where stem-cell transplantation and conventional chemotherapy are contrasted. The primary endpoint usually is disease free survival, a combined endpoint that consists of two competing events, relapses and non-relapse mortality. The latter is commonly treatment-related and occurs in early time course. It is anticipated that covariates have different impacts on each competing event and treatment by covariate interactions are present.

A simulation study is designed to confirm the theoretically anticipated good statistical performance of the approach in a wide range of complex situations. The simulation study demonstrates unbiasedness and satisfying coverage probabilities of model results and survival probabilities.

Cortese, Giuliana : ***Predicting optimal chemotherapy dosages for breast cancer via multi-state regression models with time-dependent covariates***
Author list: *Cortese, Giuliana*

In breast cancer, the risk for cardiotoxicity due to chemotherapy increases with the cumulative dose of treatment over time (Ryberg et al., 2010). Therefore, it is of interest to estimate an optimal cumulative dosage over time that guarantees a low risk for cardiotoxicity, while controlling the competing risk for mortality by maximizing the antitumor effect.

Cumulative doses, determined according to different time schedules, are treated as internal time-dependent covariates in addition to other risk factors. The aim is to predict an optimal cumulative dosage along a given treatment period that keep the cumulative risk for cardiotoxicity below a certain threshold (e.g. $<5\%$). For this purpose, a landmark approach is used and we consider both a competing risks regression model (with death and cardiotoxicity as competing events) and an illness-death model and discuss their drawback and advantages in providing correct individual predictions.

The conditional cumulative probability to have cardiotoxicity over a certain time window [s, t], Pc(s,t; X(s)), was treated as a function of cumulative dose at time s, X(s). The direct regression models for competing risks allow finding a one-to-one relationship between Pc(s,t; X(s)) and X(s), although it is not anymore possible to control the mortality risk simultaneously. Then, the optimal cumulative doses at a sequence of landmark time points, were found by inverting the dynamic predictions of a cardiotoxicity risk equal to 5%. Confidence interval estimates for the cumulative dosages were also provided.

Other Topics

Contributed session
Thursday, 18 June 2015
11:30 - 12:50
Room: U6-A11
Session chair: Friede, Tim

Rousson, Valentin : ***On the probability for a research finding to be true when the null hypothesis is impossible***
Author list: *Rousson, Valentin*

A current trend in the scientific literature is to suspect that most published research findings, identified as statistically significant results, are false [1]. The probability for a significant result to be true, interpreted as positive predictive value (PPV), is there defined as the proportion of rejected null hypotheses which are false, being positively related to the prior probability P for a null hypothesis to be false. If P is low, as speculated in exploratory research, one ends up with a PPV well below 50%. Note however that the null hypothesis is often a point hypothesis and that a point hypothesis is (almost certainly) false, implying P=1 and making the definition of PPV problematic. For example, a null hypothesis stating that the effects of two treatments are equal is (in a strict sense) false since one treatment is inevitably superior to the other, even if not in a clinical relevant way. This motivated some authors [2-3] to replace the conventional hypothesis testing approach with a "three-decision procedure", where the null hypothesis is considered impossible, and where the three possible outcomes of a statistical test are either "treatment A is superior to treatment B", "treatment B is superior to treatment A" or "we are uncertain about which treatment is superior". In this setting, a false significant result is a significant result in the wrong direction, referred to as type III error. We adopt here this procedure and redefine PPV as proportion of significant results in the right direction, which turns out to be always higher than 50%, often much closer to 95% than to 50%. This may suggest that most research findings, at least among those derived using good data and an appropriate methodology, are indeed true! Of course, this does not imply that most research findings are clinically relevant. To address this latter question, one would need still another definition of PPV.

[1] Ioannidis (2005). Why most published research findings are false? PLOS Medicine.

[2] Jones, Tukey (2000). A sensible formulation of the significance test. Psychological Methods.

[3] Freedman (2008). An analysis of the controversy over classical one-sided tests. Clinical Trials.

Hlavin, Gerald : ***Evidence, eminence and extrapolation - adjusted levels of evidence in small populations***
Author list: *Hlavin, Gerald; Koenig, Franz; Male, Christoph; Bauer, Peter*

A full independent drug development programme to demonstrate efficacy may not be ethical and/or feasible in small populations such as paediatrics populations or orphan indications. Different levels of extrapolation from a larger population to smaller target populations are widely used for supporting decisions in this situation. There are guidance documents in drug regulation, where a weakening of the statistical rigour for trials in the target population is mentioned to be an option for dealing with this problem. To this end we propose clinical trials designs, which make use of prior knowledge on efficacy for inference.

We formulate a framework based on prior beliefs in order to investigate when the significance level for the test of the primary endpoint in confirmatory trials can be relaxed (and thus the sample size can be reduced) in the target population while controlling a certain posterior belief in effectiveness after rejection of the null hypothesis in the corresponding confirmatory statistical test.

We show that point-priors may be used in the argumentation since under certain constraints they have favourable limiting properties among other types of priors.

The crucial quantity to be elicited is the prior belief in the possibility of extrapolation from a larger population to the target population. We try to illustrate an existing decision tree for extrapolation to paediatric populations within our framework.

Brombin, Chiara : ***Parametric and nonparametric approaches for multimodal emotions recognition***

Author list: *Brombin, Chiara; Rancoita, Paola Maria Vittoria; Martoni, Riccardo Maria; Ferrario, Manuela; Di Serio, Clelia*

A correct recognition of emotions is one of the major challenges in modern neuroscience.

In this framework multimodal approaches are becoming winning strategies to provide more objective and quantitative measure of experienced emotions. This setting is based on procedures that allow to integrate information from different sources, such as socio-demographic characteristics, clinical data and physiological responses to affective stimuli, whereas standard assessment tools, e.g. self-report questionnaires, are very limiting.

This contribution presents results obtained from a pilot study carried out on healthy volunteers for which clinical information, subjective responses and physiological measurements (e.g., blood volume pulse, respiratory rate and skin conductance parameters) were available.

These measurements were taken during an experimental session, where 20 stimuli (5 neutral pictures and 15 emotionally charged pictures) were randomly presented to each participant.

Stimulus presentation (lasting 5 seconds) was preceded and followed by a 10-s blank slide, allowing participants' physiological measures to return to baseline. Subjects were also asked to indicate the perceived emotion, to rate the felt emotion scale using a Likert-scale and assess the three affective dimensions of pleasure, arousal, and dominance.

In order to evaluate whether neutral and emotionally charged non-neutral stimuli were able to differently activate subjects, modulating physiological response and affecting subjective ratings, a solution within the NonParametric Combination (NPC) methodology (Pesarin and Salmaso, 2010) is here proposed.

Repeated measures issue are also achieved since the same subject is monitored in 20 different instances (while watching stimuli) and pictures are grouped according to the emotion they are expected to elicit. The key advantage of this methodology is that the underlying dependence structure is nonparametrically and implicitly captured by the combining procedure.

Moreover, the proposed approach allows for quite efficient solution in presence of high-dimensional and small sample size data sets and to easily handle with mixed variables.

Results obtained within a permutation testing framework were compared to those obtained by applying parametric latent class modeling techniques to physiological responses. This allows to evaluate whether emotionally charged stimuli could alter physiological response, with respect to the neutral ones, and to identify different homogeneous groups of subjects with similar observed trajectories.

Spada, Elena : ***Including H, K and J transformations into the Box-Cox formula: an application to the distribution of serum immunoreactive trypsinogen levels in the neonatal screening for cystic fibrosis.***

Author list: *Spada, Elena; Zolin, Anna; Milani, Silvano*

Background. The early diagnosis and treatment were demonstrated to improve quality of life and prognosis of patients suffering with Cystic Fibrosis (CF), a severe congenital disease characterised mainly by chronic pulmonary infections, gastrointestinal and nutritional abnormalities. The neonatal screening for CF may give false-negative results, which imply delayed diagnosis and therapy, or false-positive results, which imply unnecessary anxiety of parents. Thus we carried out a study sponsored by Telethon [1] focussing on the choice of optimal thresholds for the serum level of immunoreactive trypsinogen (IRT), whose determination represents the first step of CF screening algorithm.

Aims. To show how normal approximations of the IRT distribution, useful for tracing parametric ROC curves, can be obtained with the inclusion, into the Box-Cox formula, of terms allowing for kurtosis.

Babies and methods. A database structure was designed and implemented in SAS language to store nearly 700 thousand records regarding the babies born in Lombardy and screened between 2004 and 2011, whose IRT blood level was recorded by the "Laboratorio di Riferimento Regionale per lo Screening Neonatale dell'AO Istituti Clinici di Perfezionamento" of Milano. The analysis was limited to the over 600 thousand neonates without CF, whose blood samples were drawn between the 3rd and the 5th day of life. We first applied the classical Box-Cox transformation, which allows only for skewness. Then we introduced into the Box-Cox formula a single-parameter term based on HJK transformations allowing for kurtosis [2].

Results. The IRT distribution is positively skewed and highly leptokurtic. All the transformations here considered strongly reduced both skewness and kurtosis. Surprisingly, the original Box-Cox transformation

reduced kurtosis better than the other transformations do, but not skewness; whereas the inclusion of the J transformation further reduced skewness, but not kurtosis.

Conclusions. The distribution of IRT is both skewed and leptokurtic, but these components are not independent. Thus, often kurtosis disappears, or is strongly reduced, when a transformation allowing for skewness only is carried out. In these cases the parsimony principle applies.

References.

1. Telethon project number GGP12258. Principal investigator: Milani S.

2. Fischer M, Klein I. Allgemeines Statistisches Archiv 2004; 88:35-50

Cavrini, Giulia : ***The effects of chronic diseases on physical and mental well-being: a multilevel quantile regression analysis.***

Author list: *Piombo, Sara; Cavrini, Giulia; Miglio, Rossella; Samoggia, Alessandra*

Aim: The main objective of this analysis is to evaluate the effect of chronic diseases on the perception of physical and mental well-being, taking hierarchical data structure into account. The effects are evaluated on Physical (PCS) and Mental Component Summary (MCS) indexes scores, derived by SF-12 questionnaire.

Methods: Data are drawn from the Italian household survey on "Health Conditions and use of Health Services" in 2005. Our analysis takes into account the hierarchical data structure and considers both individual characteristics and information related to socio-economic conditions and household context. Due to the pronounced asymmetry of the response variables and not normally distributed residuals, more robust estimation methods such as Linear Quantile Mixed Models are used. In order to take into account the hierarchical structure of the sample and at the same time assess the possible influence of the context on individual response, a multilevel strategy has been adopted.

Quantile regression estimates the conditional quantiles of a response variable distribution through a linear model and provides a more complete view of the relationships between variables.

The multilevel linear models, likewise linear regression, estimate the conditional expectation of a response variable taking into account the hierarchical data structure, but they are not able to characterize the entire conditional distribution of a dependent variable. Quantile regression models do allow this but are unable to deal with hierarchical data. Geraci and Bottai (2007, 2013) have introduced a new method for quantile regression with mixed effects, the "Linear Quantile Mixed Model" (LQMM). They propose a conditional quantile regression model for continuous responses where random effects are added to the model taking into account the dependence between units when hierarchical data structure is present. We have adopted the procedure proposed by Geraci and Bottai to perform a multilevel quantile regression model at individual and family level.

Results and discussion: A more detailed analysis of the conditional distribution of the response on other quantiles highlighted a differential effect of some covariates along the distribution, in particular the disability, some acute diseases (arthritis, cancer and infarct for PCS; depression, Alzheimer, and cancer for MCS and physical activities.

Survival Analysis (2)

Contributed session
Thursday, 18 June 2015
11:30 - 12:50
Room: Aula Martini
Session chair: Schemper, Michael

Ambrogi, Federico : ***Competing risks regression with high dimensional covariates***
Author list: *Ambrogi, Federico; Scheike, Thomas*

High dimensional data analysis is an important topic in many research fields. For example, biomedical research generates increasing amount of data to characterise patients bio-profiles (e.g. from genomic high-throughput assay, imaging, physiological measurements, laboratory tests, patient monitoring, etc.).

Variable selection is a long-established problem in statistical research and is every day more and more important. In the last decades many forms of penalised regression have been developed, as a modern form of variable selection, to cope with high and ultra-high dimensional settings.

The increasing complexity in the characterisation of patients bio-profiles, is added to the complexity related to the prolonged follow-up of patients with the registration of the occurrence of possible adverse events, that may offer useful insight in disease dynamic and in identifying subset of patients with worse prognosis and better response to the therapy.

Although in the last years the number of contributions for coping with high and ultra-high dimensional data in standard survival analysis have increased [1], the research regarding competing risks is less developed [2].

The aim of this work is to consider how to do penalized regression when considering the crude cumulative incidence.

The direct binomial regression model of Scheike et al. [3] is reformulated in a penalised framework to possibly fit a sparse regression model. The developed approach is easily implementable using existing high performance software to do either ridge, or lasso or elastic net penalization.

Results from simulation studies are presented together with an application to genomic data when the endpoint is progression free survival.

References

[1] Witten, DM, Tibshirani, R, 2010. Survival analysis with high-dimensional covariates. Stat Methods Med Res.

[2] Binder H, Allignol, A, Schumacher, M, Beyersmann, J, 2009. Boosting for high-dimensional time-to-event data with competing risks. Bioinformatics.

[3] Scheike T, Zhang, MJ, Gerds, T, 2008. Predicting cumulative incidence probability by direct binomial regression. Biometrika.

Chiavenna, Chiara : ***Modelling quantiles of survival in register-based epidemiological studies***
Author list: *Chiavenna, Chiara*

Results from prospective epidemiological studies are mostly reported in terms of event rates. Standard approaches like Cox regression are based on strong assumptions like proportionality of hazards, that may be too restrictive in some case. Modelling quantiles of survival is a possible alternative and presents various advantages. Survival of different exposure groups can be directly compared by computing the difference between percentiles of survival. By estimating a variety of quantiles, a nonparametric estimate of the distribution of the time to event can be obtained. Compared with hazard ratios, conditional quantiles have a much simpler interpretation and can be easily communicated to a non-expert audience.

Implementing an unbiased estimator for quantiles of censored data is not straightforward. The likelihood depends on the entire distribution of the response variable, which is not directly available from the data. Laplace regression is based on the likelihood of the asymmetric Laplace distribution.

In this talk we present an application of Laplace regression, investigating the association between menopausal hormone therapy initiation and time to the occurrence of the first coronary heart disease. The study

population was defined from three large Swedish cohorts of menopausal women, and information about cardiovascular follow-up was obtained from National Registers.

A quantile-based approach was particularly suitable because of its ability to estimate only the lowest percentiles. Avoiding extrapolation beyond the range of observable quantiles is required when considering rare events, e.g., when Register-based studies are conducted since randomized clinical trials are not feasible.

References:

M. Bottai and J. Zhang. Laplace regression with censored data. Biom J, 52(4):487-503, 2010.

N. Orsini, A. Wolk, and M. Bottai. Evaluating percentiles of survival. Epidemiology, 23(5):770-771, 2012.

Marano, Giuseppe : ***Regularized estimation of survival regression models through Bayesian P-splines***

Author list: *Marano, Giuseppe; Boracchi, Patrizia; Biganzoli, Elia Mario*

For the investigation of disease dynamics, the study of the covariate effects on the shape of the hazard function provides useful insights. Recent proposals based on penalized spline functions allow for the simultaneous estimation of the baseline hazard and covariate effects. The advantage relies on an automatic procedure for determining the optimal smoothing. Point and interval estimates of model parameters and survival functions from penalized splines may be obtained in a straightforward way by Bayesian methods from posterior density samples. The use of Bayesian P-spline techniques is based on penalized splines and mixed model theory. However, the estimation procedure involves the calculation of an analytically intractable cumulative hazard. As a practical alternative a piecewise exponential model can be fitted, whose integrated hazard is expressed in closed form, allowing for the use of general MCMC sampling routines without sacrificing flexibility (e.g. Hobbs et al (2014)).

The aim of this work was the evaluation of the robustness of the PE model with respect to: 1) different prior distributions of model parameters; 2) different splines and penalties for estimating the baseline hazard and/or covariate effects. Several models with different specifications of the structural components above have been fitted to two survival datasets from patients affected by breast cancer and by soft tissue sarcoma. The former was chosen because in the case of breast cancer the behavior of the hazard function and covariate effects are well known from several clinical studies. Results obtained by Bayesian models can be compared with such a "benchmark" to evaluate model fitting. The hazard function for sarcomas is less known, its estimated shape depended on the choice of the prior distribution for the smoothing parameter, whereas a negligible influence on the covariates effect was shown. Overall, the proposed method could be a useful tool for the flexible modeling of disease dynamics in complex frameworks like cancer follow-up studies.

Hobbs, B.P., Sargent, D.J., and Carlin, B.P.: Flexible Bayesian survival modeling with nonparametric time-dependent and shape-restricted covariate effects. Submitted to Bayesian Analysis (2014)

Young Statistician Session

Invited session
Thursday, 18 June 2015
11:30 - 12:50
Room: U6-A10
Session chair: Berghold, Andrea

Dvorzak, Michaela : ***Sparse Bayesian modeling of underreported count data***
Author list: *Dvorzak, Michaela; Wagner, Helga*

In epidemiology and public health, count data are collected to assess or monitor risks of disease. Often, however these registry data are subject to underreporting (e.g. due to diagnosing or disease classification and coding errors) as only a fraction of the true but unobserved counts is reported. Consequently, inference from the reported counts will be biased and risks will be underestimated if underreporting is not considered appropriately.

To account for potential underreporting of count data, a joint model for the data generating process of events as well as the fallible reporting process is specified. We combine Poisson regression with a logit model for underreporting, resulting in a version of the Pogit model. Both the intensity of the Poisson process and the reporting probability are related to a set of potential covariates. Identification of the joint model is achieved by additional information on the reporting process provided through validation data and incorporation of variable selection.

We perform a Bayesian analysis of the Pogit model and incorporate Bayesian variable selection in both parts of the joint model. For posterior inference, we propose a convenient MCMC sampling scheme which relies on data augmentation and auxiliary mixture sampling techniques, involving only Gibbs sampling steps.

The proposed MCMC method is evaluated in simulations and applied to real data in order to estimate the risk of cervical cancer death in four different European countries.

Maragoni, Lorenzo : ***Some quantile-based proposals for detecting differential expression in genomic studies***
Author list: *Maragoni, Lorenzo; Chiogna, Monica*

In the last decade the development of statistical methods for microarray data analysis has been object of active research. A crucial problem in this area is the search for genes that show a differential expression between two or more groups like, for example, healthy and sick tissues. A huge number of statistical tools has been proposed for this task, most of which focus on testing equality of means. The aim of the present work is to introduce a simple quantile-based statistic that can test differences at different levels of the distribution, and has the desirable property of invariance with respect to monotone data transformations.

Pavlič, Klemen : ***Properties of the log-rank type test for the comparison of net survival curves***
Author list: *Pavlič, Klemen; Pohar Perme, Maja*

In survival analysis, the log-rank test is the most commonly used test to compare survival distributions between groups. One of its basic assumptions is that the hazard function is homogeneous within each group, if this assumption is violated, stratified log-rank test may be used instead. In this work, we study the properties of the recently proposed log-rank type test in the framework of net survival. In

particular, we are interested in its relation to the tests of regression coefficients and in the difference between the stratified and non-stratified version of the test. We study the properties of both versions with simulations, comment on their interpretation and present guidelines for their usage. We also present an R function which provides an efficient algorithm despite the computational intensity of the test. This function shall be included in the relsurv package for relative survival.

Meyer, Sebastian : ***Combining social contact data with spatio-temporal models for infectious diseases***
Author list: *Meyer, Sebastian; Held, Leonhard*

The availability of geocoded health data and the inherent temporal structure of communicable diseases have led to an increased interest in statistical models for spatio-temporal epidemic data [1]. The spread

of directly transmitted pathogens such as influenza and noroviruses is largely driven by human travel [2] and social contact patterns [3]. To improve predictions of the future spread, we combine social contact matrices with a spatio-temporal endemic-epidemic model for infectious disease counts. We investigate the combined model for public health surveillance data on noroviral gastroenteritis.

References:

[1] Meyer S, Held L, Höhle M (2014): Spatio-temporal analysis of epidemic phenomena using the R package surveillance. arXiv:1411.0416.

[2] Meyer S, Held L (2014): Power-law models for infectious disease spread. Annals of Applied Statistics, 8 (3), pp. 1612-1639.

[3] Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al (2008): Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Medicine, 5, e74.

Mályusz, Miklós Tivadar : ***Applying a modified Potts model on Hungarian healthcare data***

Author list: *Mályusz, Miklós Tivadar; Arató, Miklós*

Quantifying the spatial dependancy of disease related data has been a highly relevant subject of applied statistics in the last decade. In our research, we are using a hidden random Markov field as the explanator for the spatial correlation, which is similar in many ways to the modified Potts model used by Charras-Garrido et al (2012). A notable difference is that in our case, the values in matrix B are not constant throughout the simulation. Our goal was to achieve a mapping of the area the way we can see in Green and Richardson (2002).

The basic Potts model, in which the conditional probability of a region getting the color k given the colors of all the other regions, was modified in a way to account for the values associated with the colors of the surrounding regions, not just the colors themselves.

The model was used on different datasets, most notably on Hungarian subcounty based healthcare numbers.

Statistical Methods in Environmental Sciences, Agriculture and Forestry

Invited session
Thursday, 18 June 2015
14:00 - 15:30
Room: U6-A10
Session chair: Moder, Karl; Castrignanò, Annamaria

Grausgruber, Heinrich : ***Statistics in plant breeding - From field trials to genomic selection***
Author list: *Grausgruber, Heinrich; Ametz, Christian; Vollmann, Johann*

The development of robust, productive crops with broad adaptation to variable environmental conditions is essential to provide food and feed security in view of a growing world population. Advances in molecular marker analysis and the implementation of high-throughput genotyping platforms now enable genome-wide marker analysis of one breeding line at the cost of one field plot. Genomic selection uses dense marker information together with phenotypic information (from a training population) to accurately predict the breeding value of new breeding lines. Selection based on those breeding values has risen hopes to substantially increase the rate of genetic gain in breeding amongst others by increasing the number of lines that can be evaluated and by selection at earlier generations [1,2]. The big challenge now is that phenotyping keeps pace with genotyping as phenotypic evaluation is often time consuming and expensive. Moreover, environmental influences need to be tested in multi-environment trials and incorporated into the selection process. Hence, even in the era of genomics multi-environment field trials for efficient plot selection and unravelling the genotype by environment interaction will be inevitable. Statistical designs and methods that account for natural and extraneous variation in plot errors [3] and enable unreplicated trials at one test site [4] can help to improve early generation plot selection. Identification of (i) superior germplasm with a stable performance and (ii) key discriminatory environments can improve future crop testing with limited resources [5,6]. Different statistical approaches applied at different steps of the breeding cycle will be introduced in the present contribution.

References

[1] Jonas E, de Koning DJ (2013) Trends Biotechnol 31:497-504

[2] Abera Desta Z, Ortiz R (2014) Trends Plant Sci 19:592-601

[3] Gilmour AR, Cullis BR, Verbyla AP (1997) J Agric Biol Env Stat 2:269-93

[4] Payne RW (2006) Crop Sci 46:2476-81

[5] Sharma RC, Morgounov A, Akin B, Bespalova L, Lang L, Litvinenko M, Mustatea P, Ozturk I, Postolatiy A, Rajaram S, Braun HJ (2014) Crop Sci 54:2469-80

[6] Heslot N, Jannink JL, Sorrells ME (2014) Crop Sci 53:921-33

Nothdurft, Arne : ***Seeing the trees despite the forest***
Author list: *Nothdurft, Arne; Saborowski, Joachim; Nuske, Robert S.; Bäuerle, Heidi; Stoyan, Dietrich*

In forest surveys, a standard approach is k-tree sampling based on measuring the radius from a random location to the k-th closest tree. Tree density is then simply estimated as k divided by the sample plot area taken as a circular plot. While easy to apply, tree counts based on k-tree sampling are seriously biased. This is because the distance to the k-th tree is random and dependent on the tree pattern configuration generated by an unknown stochastic point process.

We propose a new density estimator where the forest is reconstructed using the tree-distance measurements from k-tree sample plots in conditional simulations. The tree count is then derived by simply counting the number of trees in the final reconstruction. The reconstruction based density estimator (RDE) proved to be approximately unbiased in various types of tested tree patterns, such as clustering, regularity and randomness.

Forest surveys are often focused on trees showing specific attributes, such as high quality timber or relevant ecological characteristics. In those surveys distance sampling is used, by which trees are more likely to be missed if their distance from the observer is larger. The observed tree pattern is thus thinned due to an unknown random detection function.

Because RDE was restricted to stationary tree patterns it was further enhanced to handle heterogeneous and thinned point patterns. By analyzing the reconstructed tree pattern in a protected landscape area it was found that habitat trees show pseudo-clusters, which are induced by their heterogeneous density. This type of point patterns were then modeled by log-Gaussian Cox processes where the tree locations follow inhomogeneous Poisson processes with log-Gaussian intensity.

References

Bäuerle, H. & Nothdurft, A. (2011): Spatial modeling of habitat trees based on line transect sampling and point pattern reconstruction. Canadian Journal of Forest Research, 41(4), 715–727.

Nothdurft, A., Saborowski, J., Nuske, R.S., Stoyan, D. (2010): Density estimation based on k-tree sampling and point pattern reconstruction. Canadian Journal of Forest Research, 40(5), 953–967.

De Benedetto, Daniela : ***Application of geophysical methods in agriculture: a protocol of measurement and data analysis for estimating soil water content***

Author list: *De Benedetto, Daniela*

Knowledge of soil water content (SWC) variation is an important asset for the optimization of irrigation and saving natural resources. Since traditional techniques are inefficient in providing rapid large-scale data collection, there is a growing necessity of searching new methodologies and sensors to assess the soil properties variability at very fine resolutions. Proximal sensing, such as Ground-penetrating radar (GPR) and electromagnetic induction (EMI), can provide high resolution data for predicting SWC. The use of GPR data is still limited by the complexity of processing and by the non-unique relation with SWC. The objectives of this work were to develop a protocol of soil survey and data analysis, based on geostatistics, to fuse multi-geophysical sensors and soil point data to improve SWC prediction.

A 2-ha field (south-eastern Italy) was monitored with an EMI sensor and two monostatic GPR systems (the first with a 250MHz frequency antenna and the other with two antennas of 600 and 1600MHz frequencies). SWC of one hundred samples was measured with gravimetric method. GPR data were analysed with a procedure based on instantaneous amplitude and visualization in maps at different depths. For SWC prediction, an optimal subset of predictors, standardized to mean zero and unit variance, was selected through variogram analysis by keeping the variables whose cross-variogram with SWC looked spatially structured. Multi-collocated-cokriging was used to fuse sparse soil data with much denser geophysical information. Finally, SWC prediction was compared with the one obtained with ordinary kriging by calculating different error statistics in a cross-validation test. Moreover, to assess the degree of association between observation and prediction, the correlation coefficient was calculated and compared for the two methods.

SWC was significantly and positively correlated with the signal amplitude at 250MHz frequency, whereas the correlations with amplitude at 600 and 1600MHz frequencies were not significant, showing that relationship is frequency-dependent. The use of geophysical data allowed the upscaling of SWC to field size. However, the advantages, in terms of unbiasedness and accuracy, compared with kriging estimation were not so evident, quite likely due to extremely high sensitivity of geophysical outcomes to several characteristics of soil/subsoil.

The future challenge will be to prove that the spatial pattern showed by these sensors is actually related to SWC variability.

Castaldi, Fabio : ***Estimation of soil and crop variables using linear mixed models and remote sensing data in the context of precision agriculture***

Author list: *Castaldi, Fabio; Casa, Raffaele; Castrignanò, Annamaria*

Detailed information on soil and crop variables is one the main requirements of precision agriculture. The knowledge of field variability allows a more efficient and sustainable utilization of resources and inputs. Spatial Linear Mixed Models (LMMs) have been shown to be more accurate for prediction, as compared to classical regression techniques which assume that residuals normally, identically and independently distributed. The simultaneous estimation of correlation parameters and fixed effect coefficients can be obtained by residual maximum likelihood (REML) variogram estimator, which has proven to be more accurate than Matheron's method of moments estimator (MoM), especially in the case of low number of samples. Often, in the context of field experiments, the number of samples can be lower than 100, mainly due to the cost of sampling and laboratory analyses. Covariates, having a significant correlation with target variable and available at high density, can be integrated into LMMs to improve the estimation of

spatial prediction models. In this context, remote sensing data can provide auxiliary information in a cost effective way over large areas.

The advantage of using such approaches in the context of precision agriculture will be reviewed and an example will be provided concerning the estimation of grain nitrogen content in wheat, an important variable for yield quality assessment.

For this purpose, SPOT satellite images were acquired over a field in Central Italy during the growing season of wheat. Grain nitrogen content was predicted using Ordinary Least Square (OLS) and REML techniques, starting from 100 field samples and by using remote sensing data as covariates. Different covariates were tested: reflectance of four bands of SPOT image and vegetation indices related to leaf nitrogen and chlorophyll content. The results highlight the advantage of using remotely sensed covariates in LMMs to estimate grain nitrogen content as compared to OLS or univariate geostatistical (assumed as reference) predictions. The best results in terms of Root Mean Square Prediction Error (RMSPE) were provided using LMMs models with vegetation indices, calculated from SPOT data acquired between stem elongation and booting stages of wheat, as predictors.

Adaptive Clinical Trials with Subpopulation Selection

Invited session
Thursday, 18 June 2015
14:00 - 15:30
Room: Aula Martini
Session chair: Heinzmann, Dominik; Rufibach, Kaspar

Wassmer, Gernot : ***Problems and solutions for adaptive enrichment designs***
Author list: *Wassmer, Gernot*

In adaptive enrichment designs the interim data is used to select one or more prespecified subpopulations or/and the full population for further analysis. This can be performed in a two-stage or a multi-stage setting using the approach that combines the p-values from each stage. If the primary endpoint is a long-term or a survival endpoint, often the decision to select the population(s) is based on a short-term endpoint. In such a case, as for survival designs in general, this information might jeopardize the Type I error control. Specific methods to solve this problem are available. In this talk we present examples for adaptive enrichment designs and describe the designing issues for assessing the operating characteristics of such a design.

Krisam, Johannes : ***Optimal subgroup selection rules for a targeted therapy in oncology***
Author list: *Krisam, Johannes; Kieser, Meinhard*

Throughout the recent years, there has been a rapidly increasing interest regarding the evaluation of so-called targeted therapies. These therapies are assumed to show a greater benefit in a pre-specified subgroup of patients, commonly identified by a predictive biomarker, as compared to a total patient population of interest. This complex situation has led to the necessity to develop statistical tools which allow an efficient evaluation of such treatments. Amongst others, adaptive enrichment designs have been proposed as a solution (see, e.g. [1-2]). These designs allow the selection of the most promising subgroup based on an efficacy analysis at interim. As has recently been shown, the performance of the applied interim decision rule in such a design plays a crucial role in ensuring a successful trial [3].

We investigate the situation that the primary outcome of the trial is binary or a time-to-event variable. Statistical methods are developed that allow an evaluation of the performance of decision rules in terms of correct selection probability at interim. Additionally, optimal decision rules are derived which incorporate the uncertainty about several design parameters, such as the treatment effects and the sensitivity and specificity of the employed bioassay. These optimal decision rules are evaluated regarding their performance in an adaptive enrichment design in terms of correct selection probability, type I error rate and power and are compared to so-far proposed fixed ad-hoc decision rules.

References:

[1] Wang, S. J., O'Neill, R. T., Hung, H. M. J. Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset. Pharmaceutical Statistics 2007; 6:227-244.

[2] Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharmaceutical Statistics 2011;10: 347–356.

[3] Krisam J, Kieser M. Decision rules for subgroup selection based on a predictive biomarker. Journal of Biopharmaceutical Statistics 2014;24: 188-202.

Rufibach, Kaspar : ***Comparison of clinical development plans for a confirmatory trial with subpopulation selection***
Author list: *Rufibach, Kaspar; Chen, Meng; Ngyuen, Hoa*

Given ever increasing costs to develop a new drug and intense competition, population enrichment designs should be considered during the planning phase of a pivotal trial with potential subgroup defined by a binary biomarker. Population enrichment designs explicitly factor in the possibility that the new drug might differentially benefit distinct biomarker subgroups. We have compared three clinical development plans for a time-to-event endpoint, such as overall survival, that all lead to a final decision in a pivotal trial either in allcomers only, in allcomers and biomarker positive, in the biomarker positive only, or to declare the drug futile. The decision about which hypothesis to test at the final analysis is made based on

a quick time-to-event endpoint, such as progression-free survival, at an interim analysis. We quantify the time gain when using a seamless Phase II/III adaptive design versus alternative development approaches and we outline what type of biomarker needs to be available prior to Phase II in each scenario.

Diagnostic Studies and Meta-Analysis

Contributed session
Thursday, 18 June 2015
16:00 - 17:00
Room: U6-A11
Session chair: Copas, John Brian

Landoni, Elena : ***A flexible statistical-bioinformatic strategy for the development of a volatile classifier discriminating the exhaled breath of breast cancer patients and healthy volunteers.***
Author list: *Landoni, Elena; Miceli, Rosalba; Martinez-Lozano Sinues, Pablo; Dibari, Vincenza Flora; Dugo, Matteo; Agresti, Roberto; Cristoni, Simone; Orlandi, Rosaria*

"Breathomics" is a new frontier in medical diagnosis and might pave the way to the development of novel cancer biomarkers. However, no consolidated guidelines are established for data mining in this field. We developed a flexible statistical-bioinformatic strategy aimed at finding a volatile classifier, by considering 351 signals from the exhaled breath of 25 women (14 breast cancer patients and 11 healthy volunteers) sampled in replicates using secondary electrospray ionization mass spectrometry (SESI-MS). In particular, we set up a two-step procedure in order to carry out data quality control (data pre-processing) and classifier development (supervised class comparison and class prediction analyses). The pipeline included innovative approaches for: 1) data representation, i.e. Concordance and Altman & Bland plots for concordance assessment between replicates, an egg-shaped plot summarizing the results of feature selection and a 'ROC space' plot (a generalization of the ROC curve) representing classification performance assessment; 2) data pre-processing, i.e. methods derived by gene expression data mining, such as the Combat (Combating Batch Effects When Combining Batches of Gene Expression Microarray Data) adjustment for batch effects (e.g. breath collection in different days); 3) supervised class prediction analysis, i.e. a bootstrap selection procedure using three machine learning algorithms (Prediction Analysis for Microarrays, Random Forest and Elastic Smoothly Clipped Absolute Deviation Support Vector Machines) to guarantee robust feature ranking and dissect the interconnections among the most discriminative features, followed by the implementation of different cross validated linear Support Vector Machines (SVM) models. The choice of the final classifier was based on both best classification performance (measured by the highest Youden index) and smallest number of features included in the model. We checked the quality of breath data, obtaining a strong agreement between technical replicates; moreover, we highlighted signals from molecules possibly chemically and/or functionally related in mass regions 147-148, 126-128 and 315 and finally identified four signals able to well discriminate between exhaled breath from breast cancer patients and healthy volunteers (sensitivity=93% and specificity=100%). Further investigations are in progress to evaluate the classifier performance in a larger cohort and to adapt the pipeline to different "omics" contexts.

Rossi, Giuseppe : ***ROC analysis developments in a multistate diagnostic setting: a new threshold estimation method***
Author list: *Pepe, Pasquale; Rossi, Giuseppe; Marchi, Marco*

In medical diagnosis, subjects are usually classified as non diseased or diseased based on the results of a discriminatory variable or marker. The two-state setting is the most consolidated classification method, but it does not cover all kinds of diagnostic problems. Sometimes, diagnostic problems include more than two classification states. In the ROC field, a generalized Youden index has been recently proposed for the assessment of the cut-off points selection in the k-class setting when a monotone ordering exists between the classes under study. The Youden index is the sum of the true classification probabilities and the estimation method demands it to be maximized. An empirical non-parametric estimation of the generalized Youden index has been considered by Nakas et al. (Stat. Med. 2010, 2013). Here we present a new index to estimating the optimum cut-off points for the general case of k disease states. The index is the sum of the Poisson squared deviance residuals, obtained as difference between the number of objects correctly classifieds in the ith category Oi: i = 1...k and the number of objects belonging to the ith category (Ei: i = 1...k), where k is the total number of categories. The index is estimated following the non parametric ROC methodology and the estimation method demands it to be minimized. A simulation study was carried out to compare the performance of the proposed index with that of the generalized Youden index and the Kappa index. The simulated data considered several combinations of the size of the adjacent classes and

percentage of correct classification in adjacent classes for specified cut-offs. Our simulation study showed that the performance of the proposed index in terms of selected cut-off values and correct classification was better than that of the Youden index and the Kappa index. In particular the Youden index and the Kappa index did not always show an unique selection of the cat-off values. In summary, in the common k-class classification problem with monotone ordering, the proposed index is a suitable approach to estimate the optimum threshold when two or more states are involved in a diagnostic issue.

Vock, Michael : ***Small-sample bias of limits of agreement in method comparison meta-analyses***
Author list: *Vock, Michael; Mittlböck, Martina*

In the simplest type of method comparison studies, the within-subject differences of two measurement methods are analysed. Williamson et al. [1] proposed three procedures for pooling the results of several such method comparison studies in order to calculate limits of agreement in meta-analyses. Two of these procedures are regularly used. These two procedures correspond to the fixed-effects meta-analysis using inverse variance weights and the random-effects meta-analysis according to DerSimonian and Laird, respectively. Both approaches are applied to the mean differences as well as the standard deviations of differences from the individual studies. However, if sample sizes of the individual studies are small, the limits of agreement resulting from these meta-analyses are biased towards the estimated mean difference, i.e., they are too narrow on average and therefore too optimistic. This small-sample bias is illustrated in detail and further explored. Additionally, approaches to reduce this bias are introduced and discussed.

References

[1] P. R. Williamson, G. A. Lancaster, J. V. Craig, R. L. Smyth (2002). Meta-analysis of method comparison studies. Statistics in Medicine, 21, 2013-2025.

Missing Data Analysis

Contributed session
Thursday, 18 June 2015
16:00 - 17:00
Room: Aula Martini
Session chair: Rousson, Valentin

Eusebi, Paolo : ***Mixture models for multiple imputation of missing data in longitudinal settings***
Author list: *Eusebi, Paolo; Ranalli, Maria Giovanna; Vidotto, Davide; Vermunt, Jeroen Kornelis*

Longitudinal data arise in many settings. For example, in several randomized clinical trials, the patient's quality of life is regularly assessed. Moreover, in certain prospective studies, a cohort of subjects is followed over time to infer a prognostic model. In such longitudinal studies, however, one or more outcomes may be missing for some subjects or some others may drop out prematurely. Furthermore, incompleteness of clinical records may frequently occur and some covariate values cannot be recovered. Under the assumption of data missing at random, multiple imputation (MI) is a convenient tool for handling missing data (see e.g. Carpenter and Kenward, 2013, for a review).

The present work proposes a novel MI method for dealing with missing data in longitudinal settings, by using mixture models as a tool for estimating the nonparametric joint distribution of the variables.

The approach is a generalization of the work of Vermunt et al. (2008) that used latent class models for MI of categorical data. The proposed methodology uses a mixture model that can deal with continuous and categorical variables. The model handles longitudinal data by relaxing local-dependence assumption between two consecutive time points in the same variables using a first-order time dependence.

The method is applied to a cohort study that investigates the prognostic role of biomarkers for the worsening of cognitive symptoms in Parkinson's disease and a clinical trial that investigates the efficacy of a treatment for back pain.

Results are compared with those of complete-case analyses and MI by chained equations.

The proposed methodology leads to valid estimates under different missingness assumptions and provides enhanced flexibility in handling both continuous and categorical variables.

References

Carpenter J.R. and Kenward M.G. (2013). Multiple Imputation and its Applications. Wiley, New York.

Vermunt, J.K., van Ginkel, J.R., van der Ark, L.A. and Sijtsma, K. (2008). Multiple imputation of incomplete categorical data using latent class analysis. Sociological Methodology 38, 369–397.

Rancoita, Paola MV : ***Improving data imputation in survival tree analysis using Bayesian networks***
Author list: *Rancoita, Paola M.V.; Zaffalon, Marco; Zucca, Emanuele; Bertoni, Francesco; de Campos, Cassio P.*

Prognostic patient stratification is one of the main goals in many retrospective clinical studies. It allows to define groups of patients of similar survival outcome based on a set of clinical parameters, for example, with the aim of enhancing the clinical management of new patients. Unfortunately, retrospective data may exhibit many missing covariate data because of several technical or clinical reasons, that could be also related to the length of the considered period of time and the eventual multicentric nature of the study. The presence of missing data could highly affect the analysis, especially in case of an high percentage of censored data and when the sample size is small.

The survival tree is a state-of-art method for prognostic patient stratification. Although many algorithms exist, they usually address missing data by using surrogate splits, i.e. by defining alternative splitting rules based on other variables yielding similar results to the original ones. Instead, we propose to handle missing data in this context by using a methodology for data imputation based on Bayesian networks and by applying the survival tree on the completed data later on. The Bayesian network allows to model the dependencies among covariates and, once the conditional distributions are estimated, they are used for imputing the missing values. The structure of the network is assumed unknown, thus a structural expectation–maximization procedure is employed for learning it directly from the incomplete data. In

contrast with other methods, this procedure does not assume that the missing values are completely at random, but simply at random (i.e. their missingness, conditional on the observed data, is independent of the unobserved values).

On simulated data, we show that the proposed methodology generally outperforms other existing methods for data imputation, that can be applied for completing the data in this setting. The use of Bayesian networks especially achieves a more accurate imputation when there are dependencies among covariates (which is a realistic assumption in many clinical settings). It also allows a better estimation of the prognostic patient stratification than using surrogate splits on both the simulated and real data.

Studies in Veterinary, Agricultural and Environmental Research

Contributed session
Thursday, 18 June 2015
16:00 - 17:00
Room: U6-A10
Session chair: Nothdurft, Arne

Farcomeni, Alessio : ***On the design of closed recapture experiments***
Author list: *Alunni Fegatelli, Danilo; Farcomeni, Alessio*

We consider closed recapture experiments where subjects in a static population are sampled S>1 times with replacement. The capture history of each subject observed at least once can be used to infer on the population size. A limitation of current experiments is that the number of capture occasions is often set in advance without formal planning, resulting in experiments that could be either too expensive or too imprecise. There are very few planning procedures available, which are either specific to certain models, based on informal computer experiments, or yielding prescription on uncontrollable pre-experimental features (e.g., the sampling fraction). In this work we outline a general method for planning the only controllable feature, namely S, based on the expected length of the profile likelihood confidence interval. After discussing the general principle, which can be applied numerically for any model specification, we detail our approach for models M0 , Mt , Mh , Mb , Mbh. These models take into account the most common sources of heterogeneity (time-heterogeneity due to different experimental conditions at different occasion, behavioral responses to capture, unobserved heterogeneity) and in many cases closed form solutions of relevant quantities can be obtained. We formally show the validity of the approach by proving distributional convergence of the planned and expected (before the experiment) profile confidence interval. We illustrate with simulations and benchmark examples. An interesting conclusion of practical relevance is that in many of the benchmark examples considered adding as few as two sampling occasions might have substantially reduced the length of confidence intervals around population size estimates.

Lucà, Federica : ***Influence of the calibration set size on the prediction of soil organic carbon through Vis–NIR spectroscopy***
Author list: *Lucà, Federica; Conforti, Massimo; Castrignanò, Annamaria; Matteucci, Giorgio; Buttafuoco, Gabriele*

Visible and near-infrared (Vis-NIR) spectroscopy has a great potential for estimating soil properties. Different chemometrics and statistical techniques have been implemented but their effectiveness and accuracy is strongly dependent on the calibration set selection. In fact, finding a calibration set representative of both the variation of soil property and spectra is crucial.

The aim of the work was to evaluate the effect of the calibration set size on the predictive performance of Vis–NIR models for soil organic carbon (SOC).

The dataset is made up of 217 topsoils (0-20 cm) collected in the "Marchesale" Biogenetic Nature Reserve (Serre Massif, Calabria) within the project LIFE09 ENV/IT/078 "Managing forests for multiple purposes: carbon, biodiversity and socio-economic wellbeing" (ManFor C.BD.). Soils, developed in a beech forest onto Paleozoic granitoid rocks or colluvial deposits, are young (Entisol and Inceptisol) and shallow to moderately deep. The samples, oven dried at 40° for 48 hours and sieved at 2 mm, were used for spectroscopic measurements and analysed for SOC content. The Vis-NIR reflectance was measured in laboratory, under artificial light, using an ASD FieldSpec IV 350-2500 nm spectroradiometer (Analytical Spectral Devices Inc., Boulder, Colorado, USA), whereas SOC was determined using a TOC-analyzer (Shimadzu Corporation, Kyoto, Japan). Reflectance was converted to absorbance and a standard normal variate (SNV) pre-processing was applied to each spectrum.

Principal component regression (PCR) and support vector machine regression (SVMR) were used to analyze the relationships between spectra and SOC whereas root mean square error (RMSE) and coefficients of determination (R2) were used to evaluate the sensitivity of the models. Calibration data, from 45 to 80% of total samples, were selected within the deciles of the sample distribution, and used to construct the regression models, validated by means of independent data.

Results revealed that when the number of calibration set samples was relatively small, PCR generated models with poor generalization ability. An increasing of R2 and low RMSEs were observed until the

calibration set samples were about 70% of the total. On the other hand, when the SVMR was used, the sampling size was not a critical issue.

Pittavino, Marta : ***Comparison between generalized linear models and additive bayesian networks. analysing risk factors for leptospira incidence in meat workers in New Zealand.***
Author list: *Pittavino, Marta; Dreyfus, Anou; Lewis, Fraser; Heuer, Cord; Torgerson, Paul; Furrer, Reinhard*

Introduction:

Leptospirosis burdens New Zealand's (NZ) rural communities with most cases occurring in farmers and meat workers. In a recent cohort study in four sheep slaughtering abattoirs in NZ, sera were collected bi-annually from 384 meat workers and tested by the Microscopic Agglutination Test for Leptospira interrogans sv Pomona and/or Leptospira borgpetersenii sv Hardjobovis. The annual infection risk (incidence) was 12%. Significant risk factors for new infection in the generalized linear model (GLM) were worker position, abattoir and time worked in the meat industry.

The aim of the study was to compare the results from the GLM with those from additive Bayesian network (ABN).

Methods:

ABNs are a form of graphical modeling which generalize the usual GLM to multiple dependent variables. In a first step we identified the best fitting model using the marginal likelihood metric and in a second step we adjusted this model for any over-fitting applying a parametric bootstrapping approach. The marginal posterior 95% confidence intervals of the log odds ratio for each parameter were estimated and compared with the GLM results.

Our R package abn, for modelling data using additive Bayesian networks, will be presented and used to conduct the analysis with ABN.

Results:

Results indicate that significantly associated with Pomona were "Worker position" and "Abattoir" (GLM) and "Worker position" (ABN). The odds of Pomona infection (OR, [95

Conclusions:

ABN confirmed the main outcome from the GLM and showed also a statistical added value through their graphical structure, extending GLM results and giving deeper insights. The advantage of ABN compared to GLM is that all relationships between all variables are modelled, revealing more about key features of complex disease systems.

Satellite meeting

Talks

Friday, 19 June 2015
09:30 - 12:30
Room: Napoleonic Hall , University of Milan, Via Sant'Antonio 10
Session chair: Decarli, Adriano

The meeting focus will be on recent developments in nutritional epidemiology, including issues in dietary assessment, the role of microbiota and the problems related to prevention policy related to epidemiological evidences. The workshop includes a round table devoted to informal discussion of issues or problems that workshop participants have experienced in conducting their own research or analyses.

Peto, Julian : ***The VIDAL Trial: a cluster randomised comparison of open v. placebo-controlled allocation of high-dose vitamin D in healthy adults***
Author list: *Peto, Julian*

In the VIDAL (Vitamin D and Longevity) feasibility trial we have randomised 20 GP practices (each randomising 80 patients aged 65-84) in 10 matched pairs between open randomisation of 2 years of vitamin D v. no treatment (800 patients), or double-blind randomisation of 2 years of vitamin D v. placebo (800 patients). This is a feasibility study for a trial of 20,000 patients treated for 5 years. Open randomisation gives reliable evidence on long-term effects of prolonged prophylaxis on cause-specific mortality, cancer incidence and subgroups of serious cardiovascular disease. It is cheaper and simpler than placebo controlled randomisation, and is likely to achieve longer compliance over 5 years and hence substantially larger long-term effects on disease and survival rates (if vitamin D has such effects), giving greater power. Moreover, many patients who have taken open-label high-dose vitamin D for 5 years would probably accept a further 5 years of treatment, further widening the difference between the vitamin D arm and untreated controls. A doubling of the eventual treatment effect would give a trial of 10,000 with open allocation and extended treatment equivalent power to a trial of 20,000 or more with placebo control. Opponents of open randomisation maintain that a placebo effect on lifestyle behaviour can never be ruled out, so placebo control should always be used despite these disadvantages even when the main endpoint is mortality.

Ferrari, Pietro : ***Are complex models in nutritional epidemiology always worth the trouble?***
Author list: *Ferrari, Pietro*

It has been repeatedly emphasized that diet could account for up to 40% among preventable causes of cancer, although the consensus around this estimate is not unanimous. Despite several decades of research, comparatively few nutrition-related factors have been established as playing a causal role in human cancer. The evaluation of role of diet on the occurrence of cancer has entailed a number of methodological challenges. First, extensive focus was given to procedures designed to perform correction of risk parameters for random and systematic measurement errors in individuals' dietary exposure estimates. Second, the evaluation of exposure/disease relationships in international multi-center study consortia motivated the need to exploit any level of etiological evidence, notably at the individual level (within-center) and at the aggregate level (between-center). Third, standard approaches have long focused on the relation between one or a restricted group of foods or nutrients and the risk of cancer, which requires a relevant use of statistical assumptions when controlling for potential confounding by other dietary and lifestyle factors. Recognizing the multi-factorial nature of cancer and other chronic diseases, complementary holistic methodologies have been employed to address the notion of dietary patterns, a concept conceived to address the inherent inter-correlations between dietary variables. Strategies relying on a priori (evidence driven) or a posteriori (unsupervised or data driven) approaches have been proposed, thus contrasting analytical simplicity with computational sophistication. The merits and the pitfalls of each of the above points will be illustrated and discussed. In an effort to provide workable tools to understand the etiology and possibly prevent chronic diseases, the day-to-day experiences of applied statisticians should be characterized by continuous concerns on the efficacy of cutting-edge statistical models to tackle biological complexity.

Muller, David Clemens : ***More than just hazard ratios: using flexible parametric survival models to comprehensively investigate associations in prospective cohort studies.***
Author list: *Muller, David Clemens*

Typically the analysis of a prospective epidemiological study begins and ends with the calculation of a hazard ratio for each covariate of interest. Whilst hazard ratios are useful summaries of associations, it is often desirable to investigate and evaluate the contribution of risk factors in a more comprehensive manner. For instance conditional survival functions can provide the absolute risk of disease for a given pattern of covariates. Another interesting quantity is the attributable fraction, or the proportion of events observed in a population that would not occur given the removal of some causal risk factor or set of risk factors. Flexible parametric survival models, also known as Royston-Parmar models, are an alternative to the Cox proportional hazards model for the analysis of time-to-event data. In addition to the estimation of hazard ratios, these models use restricted cubic splines to directly and flexibly model the baseline hazard thus allowing immediate estimation of absolute risks, unlike the Cox model which treats the baseline hazard as a nuisance parameter. Further, they provide a simple framework for fitting models with non-proportional hazards. I will give an introduction to flexible parametric survival models, and will demonstrate how it is possible to estimate conditional survival functions and attributable fractions for an arbitrary set of covariates by making predictions from the fitted model. I will also propose an attributable fraction that is defined as a function of time, which can thus incorporate non-proportional hazards. To illustrate these points, I will use examples from a study quantifying the relative importance of various factors, including a composite diet score, toward premature mortality in the European Prospective Investigation into Cancer and nutrition study.

Wiens, Frank : ***The effect of diet on human milk composition – is it relevant for public health?***
Author list: *Wiens, Frank*

Maternal and child health are tightly linked. This link is mainly nutrition during the first thousand days from conception to a child's second birthday. Maternal milk is recommended to form a significant part of the infant nutrition during this period from right after birth onwards. Maternal milk is highly variable between women; part of this variability is caused by dietary differences. So far, efforts to understand the public health relevance of maternal diet via changed breastmilk composition on infant's long-term health have focused mainly on the role of micronutrient deficiencies, but the relevant scope might be much broader. One neglected aspect, for example, is the oxidative load and antioxidative capacity of human milk that can be hypothesized both to be influenced by diet and to impact infant health. Some traditional, once widespread dietary patterns feature high levels of anti-oxidative compounds possibly translating into a typical redox status of the milk of breastfeeding women. For estimating the effect of modern dietary trends towards reduced intake of anti-oxidative plant compounds and more reliance of highly processed foods on the health of breastfed infants even basic biomedical data are still missing. A step-wise approach towards collecting these data and using them to test public health hypotheses on the impact of modern changes in human milk composition is outlined.

Pigeot, Iris : ***Validity and modelling of dietary exposures in young children***
Author list: *Pigeot, Iris; Börnhorst, Claudia*

The assessment of dietary intake is a challenging task especially in population-based surveys. Various instruments such as food frequency questionnaires or 24-hour dietary recalls are at hand where advantages and drawbacks have to be balanced against each other to find the most appropriate assessment method depending on the purpose of the study. Problems emerge from the daily variation in diet and in particular from differential measurement errors resulting from (un)intentional misreporting which need to be accounted for in the statistical analyses to avoid biased estimates of exposures and diet-disease associations. In young populations, these problems may be even more pronounced as dietary data mainly rely on proxy respondents where little is known about the validity and determinants of misreporting. However, valid estimates of intakes in children populations are essential for monitoring trends as well as for nutritional interventions, for instance, to combat childhood obesity.

In this talk, we will briefly introduce the various instruments used in the IDEFICS study to assess dietary intake in 16,228 2-9.9 year old children in eight European countries. We will then present (1) results of various validation studies and estimates of the prevalence and determinants of misreporting in proxy-reported data, (2) statistical methods to model dietary exposures like usual intakes and dietary patterns, e.g. based on model-based clustering techniques, and (3) methods to account for misreporting when estimating diet-obesity association e.g. by using propensity scores. All methods will be applied to the IDEFICS dataset.

References

Börnhorst C, Huybrechts I, Hebestreit A, Vanaelst B, Molnár D, Bel-Serrat S, Mouratidou T, Moreno LA, Pala V, Eha M, Kourides YA, Siani A, Eiben G, Pigeot I, on behalf of the IDEFICS consortium. Diet-obesity associations in children: approaches to counteract attenuation caused by misreporting. Public Health Nutr 2013; 16: 256-66.

Börnhorst C, Huybrechts I, Ahrens W, Eiben G, Michels N, Pala V, Molnár D, Russo P, Barba G, Bel-Serrat S, Moreno LA, Papoutsou S, Veidebaum T, Loit HM, Lissner L, Pigeot I, on behalf of the IDEFICS consortium. Prevalence and determinants of misreporting among European children in proxy-reported 24 h dietary recalls. Br J Nutr 2013; 109: 1257-65.

Rossi, Marta : ***Nutritional epidemiology of cancer in case-control studies: from dietary assessment to a metagenomics approach.***

Author list: *Rossi, Marta*

In a network of Italian case-control studies on common cancers, various dietary aspects have been investigated during the last two decades. A valid and reproducible food frequency questionnaire, including 78 food items and 5 questions on alcoholic beverages, was developed to collect subject's dietary habits. Among major results, a high consumption of fruit and vegetables has been found to protect against common epithelial cancers, including those of the digestive tract. Limiting red meat, refined grain carbohydrates, and saturated fats has been associated to a reduced risk of the upper aerodigestive tract cancers. Besides food groups, single foods (e.g. apples, tomatoes) and several nutrients (e.g. vitamins, polyphenols) have been successively investigated according to the current research trends. The investigation then focused on the development of a priori and a posteriori dietary patterns, in particular on the Mediterranean diet score, which has been inversely related with the risk of various cancers of the digestive and respiratory tracts. More recently, other multivariate methods have been used to identify a pattern that maximizes the association between dietary exposures and cancer risk. Another approach studied the total antioxidant capacity (TAC) from the entire diet – instead of the intakes of single food components. These two innovative techniques allow taking into account the activity of all dietary factors and their potential synergistic effects. Possible future developments to overcome the major limits of these nutritional studies will be discussed. In particular, a case-control study applying an innovative metagenomic approach to explore the existence of a diet-microbiota axis as a prominent factor for colorectal cancer will be presented.

Cavalieri, Duccio : ***The hygene hypothesis revisited: Role of microbiota and mycobiota in immune regulation***

Author list: *Cavalieri, Duccio*

The human microbiome is a complex consortium of trillions of microbes, whose collective genomes contain at least 100 times as many genes as our own eukaryote genome.

We inherit the seed of this community from our mothers at birth, and recent reports suggest that early microbial colonization has an important role for in promoting health. Advances in understanding host-microbe interactions imply that maternal microbiota plays a crucial role on health programming. Having the right seed microbiomein the child may contribute to reduce the risk of chronic diseases such as obesity, allergies and inflammatory conditions in the elderly. This process begins in utero and it is modulated by mode of delivery and diet. Maternal microbiota is transferred to infants during birth and supported by lactation practices.

Historically, the microbial ecosystem of the digestive tract was specific for a geographic area, as much as the flora and fauna of an ecosystem are geographically distinct.

Globalization of the microbial population due to industrialization and standardization of food chain products is thought to have a causative effect for the recent raise in immune related disorders.

In our talk we will discuss how globalization of diet is standardizing the microbiome, how robust and inheritable is the effect of this standardization, and how this change is affecting immune and metabolic health.

We will discuss how solely an integrated Systems Biology approach will allow to turn data deluge into knowledge permitting comparability of datasets on diet, microbiota and immune system. Finally we will discuss how the so far underscored fungal component of the microbiota plays a fundamental role in immune training.

Understanding the effect of globalization of microbiota on health will impact importantly on nutritional, microbiological and food production policies, helping in the end reducing the risk of chronic diseases such as obesity, allergies and inflammatory conditions.

Satellite Meeting - Round Table

Friday, 19 June 2015
14:00 - 16:00
Room: Napoleonic Hall , University of Milan, Via Sant'Antonio 10
Session chair: Ulmer, Hanno

In an integrated picture of different points of view, the speakers will bring to the attention of the participants their ideas on the available evidence on healthy dietary habits and on the open issues that researchers need to address for a deeper understanding of the relationship between foods, nutrients and health.

Speakers:

- Wolfgang Ahrens, Leibniz-Institut für Präventionsforschung und Epidemiologie, Bremen, Germany.
- Pietro B. Ferrari, Nutrition and Metabolism Section, IARC, Lyon, France.
- Vittorio Krogh, Fondazione IRCCS- Istituto Nazionale dei Tumori, Milan, Italy.
- Julian Peto, Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.
- Iris Pigeot-Kübler, Leibniz-Institut für Präventionsforschung und Epidemiologie, Bremen, Germany.
- Michael Themessl-Huber, Section for Medical Statistics, Medical University of Vienna, Austria.

List of Participants

Name			E-Mail
Ms.	Ahrens	Maike	maike.ahrens@uni-dortmund.de
Prof.	Alfo'	Marco	marco.alfo@uniroma1.it
Dr.	Alunni Fegatelli	Danilo	danilo.alunnifegatelli@uniroma1.it
Prof.	Ambrogi	Federico	federico.ambrogi@unimi.it
Prof.	Andersen	Per Kragh	pka@biostat.ku.dk
Dr.	Antolini	Laura	laura.antolini@unimib.it
Ms.	Assi	Nada	assin@students.iarc.fr
Ms.	Baccini	Michela	baccini@disia.unifi.it
Dr.	Banchelli	Federico	federico.banchelli2@unibo.it
Dr.	Barbati	Giulia	giulia_barbati@yahoo.it
Ms.	Barrio	Irantzu	irantzu.barrio@ehu.es
Prof.	Bellocco	Rino	rino.bellocco@unimib.it
Prof.	Berghold	Andrea	andrea.berghold@medunigraz.at
Dr.	Bernasconi	Davide Paolo	davide.bernasconi@unimib.it
Prof.	Biggeri	Annibale	abiggeri@disia.unifi.it
Dr.	Bottelli	Stefano	stefano.bottelli@istitutotumori.mi.it
Dr.	Bravi	Francesca	francesca.bravi@unimi.it
Mr.	Bravi	Fabio	fabio.bravi@ibisinformatica.com
Mr.	Brizzi	Francesco	francesco.brizzi@mrc-bsu.cam.ac.uk
Dr.	Brombin	Chiara	brombin.chiara@hsr.it
Dr.	Burger	Hans Ulrich	hans_ulrich.burger@roche.com
Dr.	Calle	M.Luz	malu.calle@uvic.cat
Prof.	Calza	Stefano	stefano.calza@unibs.it
Dr.	Castaldi	Fabio	castaldi@unitus.it
Dr.	Castrignanò	Annamaria	annamaria.castrignano@entecra.it
Prof.	Cavalieri	Duccio	duccio.cavalieri@fmach.it
Prof.	Cavrini	Giulia	gcavrini@unibz.it
Mr.	Chaouch	Aziz	aziz.chaouch@chuv.ch
Ms.	Chiavenna	Chiara	chiara.chiavenna@gmail.com
Dr.	Ciniselli	Chiara Maura	chiara.ciniselli@istitutotumori.mi.it
Prof.	Copas	John Brian	jbc@stats.warwick.ac.uk
Ms.	Corbella	Alice	alice.corbella@mrc-bsu.cam.ac.uk
Dr.	Cortese	Giuliana	gcortese@stat.unipd.it
Dr.	Crivori	Patrizia	patrizia.crivori@clioss.com
Dr.	De Benedetto	Daniela	daniela.debenedetto@entecra.it
Dr.	De Bin	Riccardo	debin@ibe.med.uni-muenchen.de
Dr.	De Capitani	Lucio	lucio.decapitani1@unimib.it
Prof.	Decarli	Adriano	adriano.decarli@unimi.it
Dr.	De Lorenzo	Paola	paola.delorenzo@unimib.it
Mr.	Di Maso	Matteo	matteo.dimaso@unimi.it

Name			E-Mail
Prof.	Di Serio	Clelia	diserio.clelia@unisr.it
Dr.	Dunger-Baldauf	Cornelia	cornelia.dunger-baldauf@novartis.com
Dr.	Dunkler	Daniela	daniela.dunkler@meduniwien.ac.at
Ms.	Dvorzak	Michaela	michaela.dvorzak@joanneum.at
Dr.	Edefonti	Valeria	valeria.edefonti@unimi.it
Dr.	Eusebi	Paolo	paoloeusebi@gmail.com
Dr.	Farcomeni	Alessio	alessio.farcomeni@uniroma1.it
Dr.	Ferrari	Pietro	ferrarip@iarc.fr
Ms.	Ferré	Chiara	chiara.ferre@unimib.it
Prof.	Finos	Livio	livio.finos@unipd.it
Prof.	Friede	Tim	tim.friede@med.uni-goettingen.de
Prof.	Frigessi	Arnoldo	frigessi@medisin.uio.no
Mr.	Fritz	Josef	josef.fritz@i-med.ac.at
Prof.	Galimberti	Stefania	stefania.galimberti@unimib.it
Mr.	Gasparini	Alessandro	a.gasparini9@campus.unimib.it
Prof.	Gasparini	Mauro	gasparini@calvino.polito.it
Dr.	Geroldinger	Angelika	angelika.geroldinger@meduniwien.ac.at
Dr.	Gosoniu	Maria Laura	MariaLaura.Gosoniu@rdls.nestle.com
Dr.	Grant	Timothy Sean	tsgrant.dublin@gmail.com
Prof.	Grausgruber	Heinrich	heinrich.grausgruber@boku.ac.at
Ms.	Greco	Teresa	teresa.greco@unimi.it
Dr.	Greco	Luca	luca.greco@unisannio.it
Prof.	Gu	Fei	fei.gu@mcgill.ca
Dr.	Hayoz	Stefanie	stefanie.hayoz@sakk.ch
Mr.	Heinze	Georg	georg.heinze@meduniwien.ac.at
Dr.	Heinzmann	Dominik	dominik.heinzmann@roche.com
Prof.	Held	Leonhard	leonhard.held@uzh.ch
Dr.	Held	Ulrike	ulrike.held@usz.ch
Dr.	Herzog	Sereina	herzog.sereina@gmail.com
Mr.	Hlavin	Gerald	gerald.hlavin@meduniwien.ac.at
Dr.	Hsu Schmitz	Shu-Fang	shu-fang.hsu_schmitz@novartis.com
Dr.	Jurman	Giuseppe	jurman@fbk.eu
Dr.	Klinglmueller	Florian	florian.klinglmueller@meduniwien.ac.at
Mr.	Krisam	Johannes	krisam@imbi.uni-heidelberg.de
Dr.	Landoni	Elena	elena.landoni@unimi.it
Dr.	Lucà	Federica	federica.luca@isafom.cnr.it
Dr.	Lugo	Alessandra	alessandra.lugo@unimi.it
Dr.	Magri	Maria Chiara	m.magri12@campus.unimib.it
Mr.	Mályusz	Miklós Tivadar	mmtivadar@gmail.com
Dr.	Maragoni	Lorenzo	lorenzo.maragoni@unipd.it

Name			E-Mail
Dr.	Marano	Giuseppe	giuseppe.marano@istitutotumori.mi.it
Dr.	Mastandrea	Paolo	paolomastandrea@gmail.com
Dr.	Mattei	Alessandra	mattei@disia.unifi.it
Dr.	Maurer	Willi	maurers@balcab.ch
Prof.	Mayo-Iscar	Agustin	agustin@med.uva.es
Prof.	Mealli	Fabrizia	mealli@disia.unifi.it
Mr.	Meyer	Sebastian	sebastian.meyer@uzh.ch
Prof.	Mezzetti	Maura	maura.mezzetti@uniroma2.it
Dr.	Migliavacca	Eugenia	eugenia.migliavacca@rd.nestle.com
Prof.	Miglio	Rossella	rossella.miglio@unibo.com
Prof.	Milani	Silvano	silvano.milani@unimi.it
Ms.	Mittlboeck	Martina	martina.mittlboeck@meduniwien.ac.at
Prof.	Moder	Karl	karl.moder@boku.ac.at
Dr.	Moran	John Leith	john.moran@adelaide.edu.au
Dr.	Muller	David Clemens	davidmull@gmai.com
Dr.	Nakas	Christos T	cnakas@uth.gr
Ms.	Neubauer	Maria	maria_neubauer@gmx.at
Prof.	Nothdurft	Arne	arne.nothdurft@boku.ac.at
Ms.	Ofner-Kopeinig	Petra	petra.ofner@medunigraz.at
Ms.	Orenti	Annalisa	annalisa.orenti@unimi.it
Mr.	Pasquale	Pepe	pasquale_pepe1@virglio.it
Mr.	Pavlič	Klemen	klemen.pavlic@mf.uni-lj.si
Dr.	Perrotta	Domenico	domenico.perrotta@ec.europa.eu
Prof.	Peto	Julian	julian.peto@LSHTM.ac.uk
Prof.	Pigeot	Iris	pigeot@bips.uni-bremen.de
Ms.	Pittavino	Marta	marta.pittavino@math.uzh.ch
Prof.	Posch	Martin	martin.posch@meduniwien.ac.at
Ms.	Pötschger	Ulrike	ulrike.poetschger@ccri.at
Mr.	Pramov	Aleksandar	aleksandar.pramov@phd.unibocconi.it
Ms.	Praud	Delphine	delphine.praud@unimi.it
Ms.	Pregartner	Gudrun	gudrun.pregartner@medunigraz.at
Dr.	Rancoita	Paola MV	rancoita.paolamaria@unisr.it
Dr.	Rebora	Paola	paola.rebora@unimib.it
Prof.	Reilly	Marie	marie.reilly@ki.se
Prof.	Reiser	Benjamin	reiser@stat.haifa.ac.il
Dr.	Riedl	Regina	regina.riedl@medunigraz.at
Dr.	Rosenkranz	Gerd	gerd.rosenkranz@novartis.com
Ms.	Rossa	Agnieszka	agrossa@uni.lodz.pl
Dr.	Rossi	Giuseppe	giuseppe.rossi@ifc.cnr.it
Dr.	Rossi	Marta	marta.rossi@unimi.it

Name			E-Mail
Dr.	Rosso	Tiziana	tiziana.rosso@unimi.it
Dr.	Rota	Matteo	matteo.rota@marionegri.it
Prof.	Rousson	Valentin	valentin.rousson@chuv.ch
Dr.	Röver	Christian	christian.roever@med.uni-goettingen.de
Dr.	Rufibach	Kaspar	kaspar.rufibach@roche.com
Ms.	Scarale	Maria Giovanna	maria.scarale@unimi.it
Prof.	Schemper	Michael	michael.schemper@meduniwien.ac.at
Prof.	Schimek	Michael G.	michael.schimek@medunigraz.at
Prof.	Schmid	Volker J.	volker.schmid@lmu.de
Dr.	Schmidli	Heinz	heinz.schmidli@novartis.com
Ms.	Seibold	Heidi	heidi.seibold@ifspm.uzh.ch
Prof.	Senn	Stephen	stephen.senn@lih.lu
Dr.	silvestri	daniela	daniela.silvestri@unimib.it
Dr.	Småstuen	Milada Cvancarova	miladacv@math.uio.no
Dr.	Solari	Aldo	aldo.solari@unimib.it
Ms.	Strohmaier	Susanne	susanne.strohmaier@medisin.uio.no
Dr.	Summers	Jennifer Ann	jennifer.a.summers@kcl.ac.uk
Ms.	Svendova	Vendula	vendula.svendova@medunigraz.at
Dr.	Szymański	Andrzej	anszyman@math.uni.lodz.pl
Mr.	Themessl-Huber	Michael	michael.themessl-huber@meduniwien.ac.at
Prof.	Ulm	Kurt	kurt.ulm@tum.de
Prof.	Ulmer	Hanno	hanno.ulmer@i-med.ac.at
Prof.	Valsecchi	Maria Grazia	grazia.valsecchi@unimib.it
Dr.	van Wieringen	Wessel N.	w.vanwieringen@vumc.nl
Prof.	Victoria-Feser	Maria-Pia	Maria-Pia.VictoriaFeser@unige.ch
Dr.	Vock	Michael	michael.vock@stat.unibe.ch
Ms.	Wallisch	Christine	christine.wallisch@univie.ac.at
Prof.	Wassmer	Gernot	gernot.wassmer@uni-koeln.de
Ms.	Wei	Wei	wei.wei@uzh.ch
Dr.	Wiens	Frank	frank.wiens@danone.com
Dr.	Wolkewitz	Martin	wolke@imbi.uni-freiburg.de
Prof.	Zambon	Antonella	antonella.zambon@unimib.it
Dr.	Zebrowska	Magdalena	magdalena.malina@meduniwien.ac.at
Ms.	Zell	Elizabeth	erzell@mac.com
Ms.	Zolin	Anna	anna.zolin@unimi.it

Sponsors

- Roche
- Novartis
- SAS
- Spreafico

Committees

Scientific Program Committee

Held Leonhard (University of Zurich) Co-chair
Valsecchi Maria Grazia (University of Milano-Bicocca) Co-chair
Ambrogi Federico (University of Milan)
Baccini Michela (University of Florence)
Berghold Andrea (Medical University of Graz)
Bretz Frank (Novartis)
Castrignanò Annamaria (Agricultural Research Council Research, Bari)
Farcomeni Alessio (University of Rome)
Finos Livio (University of Padua)
Hothorn Torsten (University of Zurich)
Posch Martin (Medical University of Vienna)
Schemper Michael (Medical University of Vienna)

Local Organizing Committeee

Galimberti Stefania (University of Milano-Bicocca) Chair
Antolini Laura (University of Milano-Bicocca)
Biganzoli Elia (University of Milan)
Bonetti Marco (Bocconi University)
Di Serio Clelia (VS San Raffaele University)
Rebora Paola (University of Milano-Bicocca)
Solari Aldo (University of Milano-Bicocca)
Miglio Rossella (University of Bologna, Treasurer IBS-ITR)
Hsu Shu-Fang (University of Bern, Treasurer IBS-ASR)

Satellite Meeting Program Committee

Decarli Adriano (University of Milan) Chair
Ulmer Hanno (Medical University of Innsbruck) Co-Chair
Edefonti Valeria (University of Milan)
Ferraroni Monica (University of Milan)
Krogh Vittorio (Fond. IRCCS - Istituto Nazionale dei Tumori, Milano)

Content and design of this book

This book was created by Heidi Seibold (heidi.seibold@uzh.ch) using ConfTool (`https://www.conftool.com`), R (`http://www.r-project.org/`), knitr (`http://yihui.name/knitr/`), and LaTeX (`http://www.latex-project.org/`).