Computational Statistics Solutions for Molecular Biomedical Research: A Challenge and Chance for Both

Lutz Edler, Christina Wunder, Wiebke Werft, and Axel Benner

Department of Biostatistics-C060, German Cancer Research Center
Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany,
edler@dkfz.de, c.wunder@dkfz.de, w.werft@dkfz.de, benner@dkfz.de

Abstract. Computational statistics, supported by computing power and availability of efficient methodology, techniques and algorithms on the statistical side and by the perception on the need of valid data analysis and data interpretation on the biomedical side, has invaded in a very short time many cutting edge research areas of molecular biomedicine. Two salient cutting edge biomedical research questions demonstrate the increasing role and decisive impact of computational statistics. The role of well designed and well communicated simulation studies is emphasized and computational statistics is put into the framework of the International Association of Statistical Computing (IASC) and special issues on Computational Statistics within Clinical Research launched by the journal Computational Statistics and Data Analysis (CSDA).

Keywords: computational statistics, molecular biomedical research, simulations, International Association of Statistical Computing, computational statistics and data analysis

1 Introduction

Statistical methods have been recognized and appreciated as unalterable tool for the progress of quantitative molecular biology and medicine (molecular biomedicine) as they were in physics, quantitative genetics and in clinical drug research. With the emergence of larger biomedical data sets, both in terms sample size (n) and number of individual characteristics (p), in particular when $p >> n$, novel and more efficient computational methods and data analysis approaches were needed, and valid conclusions and decision making required the company of statistical inference and statistical theory. Whereas from the beginning on, when molecular data appeared massively due to high-throughput techniques, extraordinary efforts and large investments were put into the quality of biomedical data and bioinformatics, much less was invested into the computational statistics methods for the information extraction. That neglect left gaps in biomedical research projects when the validity of both methods and results were questioned. Modern emerging biomedical approaches and complex models in biological, epidemiological

Y. Lechevallier, G. Saporta (eds.), *Proceedings of COMPSTAT'2010*,
DOI 10.1007/978-3-7908-2604-3_2,

and clinical studies require high quality computational and statistical support. We will address below how computational statistics, computing power, data analysis and data interpretation invaded in a very short time many cutting edge research areas of modern biomedicine and biomedical research. Therefore we will elaborate from two biomedical areas a salient cutting edge research question for which computational statistics plays now an increasing role with a decisive impact.

A statisticians work for biomedicine - name it biostatistics or biometrics - is defined by the biomedical problems and hypotheses as well as by the tools he/she has at hands to solve the corresponding mathematical and statistical problems. According to Finney (1974) it is the duty of a biostatistician "*to interpret quantitative biomedical information validly and usefully*". He also noted that that the applied statistician should express him/herself in "*terms intelligible beyond the confines of statistics* in varying degrees of collaboration with persons being expert in the field, and stressing the fact that a "*substantial contribution from the statistician is essential*" when citing R.A. Fisher with "*when a biologist believes there is information in an observation, it is up to the statistician to get it out*", one of the first statisticians who heavily calculated in Rothamsted for his collaborations with agriculture and genetics. This work has always been initiated and guided by data and required the use of computational methods for doing the calculations right and efficiently. Computational statistics has been an integral part of statistics from its beginning when statisticians had to do calculations and needed to simplify the computational work load, starting with numerical calculus using later statistical tables, mechanical calculators and electronic calculators, called computers, all overruled now by highly interactive computing systems which integrate statistics software with an almost uncountable number of applets acting during data input, data processing and data output. Victor (1984) discussed in a highly recognized essay the role of computational statistics for statistics and statisticians. Although he denied the attribute of a scientific discipline because of missing own methodology and own subject for investigation, he acknowledged the high relevance for applied statistics and its undisputable role for knowledge generation in all sciences. For a discussion of this concept, see e.g. Lauro (1996), also Nelder (1996), who distinguishes science from technology and locates computational statistics nearer to latter although it is performed by scientists. The most conciliatory definition of computational is found in Chambers (1999) citing John Tukeys defining of computational statistics as the "*peaceful collisions of computing with statistics*".

In its growth period around 1970-1980 it became the irrevocable tool for statistics. Exact statistical inference methods and permutations, bootstrapping and interactive graphical methods were the dominant tools. The interaction of computational statistics with biostatistics from the view point how it developed in Germany was summarized in Edler (2005). The technological aspect has been emphasized in the German Region of the International Bio-

metric Society when personal computers started to take over the main frame computers in 1990 by Bernd Streitberg. His vision at that time was that computational statistics has the chance to become the driving force for the progress of statistics and, in particular for biostatistics, if it will be possible to overcome fixation to program packages. At that time he already foresaw the innovative power of personal computers- now notebooks - for software development. It is fair to say that the R-project for statistical computing has made this vision coming true. Notably, he denounced software validation where he rather pessimistically stated that for many users and heads of institutes and companies that would not be an issue as long as all compute with the same software, notwithstanding whether that software calculated correct or incorrect. A second still relevant issue Streitberg indicated in 1989 was the wish that the computer has to become the standard test for the applicability and use of a statistical method: If a method cannot be programmed it is not relevant; if it is not programmed its is useless. This way of thought had been expressed already in 1981 by Jürgen Läuter who noted in the introduction to his software development that processes of thinking, decision making and production can be advanced by mathematics and computing techniques.

Concerning the high-dimensional molecular data all these thoughts seem to fit well for an intelligible interpretation of the data and the reduction of the data to their information content.

2 Screening Molecular Data for Predictors

A multitude of biomedical techniques provide high-throughput high-resolution data on the molecular basis of diseases. Most of these DNA microarray array data fall into one of the following categories:

- gene or expression
- allelic imbalance
- methylation imbalances.

These investigations aim at a better understanding of the underlying mechanisms of the genesis of the disease, e.g. of a specific cancer like the AML out of the class of leukaemia. Current biomedical knowledge postulates for most diseases, in particular for cancer, as of being heterogeneous and of different subtypes, on the clinical, histo-pathological and the molecular level. Heterogeneity at the molecular level lead to the development of prognostic and also to predictive gene signatures (also called gene expression profiles, biomarker sets etc) from which some have already been commercialized (for breast cancer see e.g., ONKOtypeDx, MammPrint, GGI) and used in attempts to personalize cancer treatment, although there exists still considerable uncertainty on the use of new molecular markers in routine clinical decision making. The need to examine their role in patient selection and for

stratification for future clinical trials is obvious. For a review of the situation in breast cancer therapy see Kaufmann et al (2010). Urgent biomedical questions concern

- usefulness of currently available molecular biomarkers and biomarker based,
- the establishment of designs and design strategies which account for clinical, histo-pathological and molecular subtypes at the same time, and
- coherent collection, combination and processing of both biomedical information, being it collected prospectively or retrospectively.

The challenge for biostatistics arising from these questions is huge and starts conceptually at a clarification of the difference between prognostic and predictive markers (Sargent (2005)).

Biomedical research has always been targeted to develop e.g. prognostic models which may classifiy patients in different risk groups and so called prognostic marker guide therapy of groups of patients in a general way. Another clinical target is the development of predictive models which guide treatment and optimize therapy by guiding treatment decision in dependency of so-called predictive factors. Next to consider is the translation of the medical task into statistical approaches. For the prognostic models the biostatisticians task, almost exclusively performed in collaborative projects with biologists and clinicians is to build prediction models e.g. for classification in different risk groups based on such molecular data. Whereas formerly the statistical inferences were based on either statstical testing or on class discovery methods e.g. cluster analyses, regression techniques are now somehow rediscovered as the more appropriate approach to build those risk prediction models. Regularization methods are now widespread to solve the so called $p >> n$ problem (penalized regression approaches like the Lasso or the Elastic Net, the use of Support Vector Machines, Boosting etc.).

For the development of biomarkers as predictive factors guiding the choice of therapy, regression type analyses are applied on the outcome variable based on those high-dimensional predictors listed above. Efficient and non-overly conservative adjustment for multiple testing becomes crucial when focusing on a gene wise analysis. Multiple adjustments becomes crucial, see e.g. Dudoit and van der Laan (2007). Simulation studies analyze sample-size determination for the identification and validation of such predictive markers. The classical multivariable regression model works well for identifying prognostic factors and with regard to predictive factors one can go back to another classical tool, interactions between covariables.

2.1 Using the Analysis of Molecular Data for Identifying Predictive Biomarker

When screening for predictive factors in case of a dichotomous outcome the method of choice is conditional logistic regression. A gene wise interaction

model has then the form

$$\text{Logit}(\mathbf{Y} = \mathbf{1}|\mathbf{X_g}; \mathbf{Z}) = \beta_{\mathbf{0g}} + \beta_{\mathbf{1g}}\mathbf{X_g} + \beta_{\mathbf{2g}}\mathbf{Z} + \beta_{\mathbf{3g}}\mathbf{X_g}\mathbf{Z}$$

where $\mathbf{X_g}$ describes the continuous gene expression of gene g, $g = 1, \ldots, M$, and M is the number of gene expression varaiables, i.e. the number of hypotheses, analyzed in total $\mathbf{Z}$ is a binary treatment variable and $\mathbf{Y}$ is a binary response variable. An interaction effect is then tested with the null hypothesis H0: $\beta_{3g} = 0$ using e.g. likelihood ratio (LR) or the Wald test. When focusing on the multiple testing scenario for the M simultaneous hypotheses one would prefer to control the false discovery rate (FDR) introduced by Benjamini and Hochberg (1995) for such gene expression data as a certain proportion of false discoveries would be accepted. Control of the FDR could be for example obtained by linear step-up procedures such as the Benjamini-Hochberg or Benjamini-Yekutieli method. Lately, resampling-based multiple testing methods for FDR control have become an alternative approach see e.g. Dudoit and van der Laan (2007).

The next challenge arises when determining a sample size for such screening methods and when complexity has barricaded an analytical solution. Only simulations of several scenarios will help to get grip on the sample size estimation which is essential for all trial partners: those investing their time and career, those who invest resources and those who are responsible the sample size, namely the trial statistician. Since one is forced to analyze the system "statistical model" in detail valuable "fall outs" of the simulation approach can be obtained, e.g. a comparison of the performance of different statistical test procedures.

Actually there is now a problem of comparability of the results since the implementation of methods and simulation designs are almost always different. From a user-friendly point of view usage of available methods is impaired by different platforms, different implementations etc. Realized as a Harvest Programme of the PASCAL2 European Network of Excellence a group of researchers has recently come together for a unified, extensible interface covering a large spectrum of multiple hypothesis testing procedures in R: μTOSS (multiple hypotheses testing in an open software system), see Dickhaus et al. (2010). Intended as first step to overcome the problem of comparability of the results μTOSS aims at unifying implementation of methods and simulation platforms as an open source package addressing (i) multiple tests controlling the family wise error rate (single-step and stepwise rejection methods, resampling-based procedures), (ii) multiple adjustment procedures controlling the false discovery rate (classical and adaptive methods, Bayesian approaches as well as resampling-based techniques), (iii) multiplicity-adjusted simultaneous confidence intervals, and (iv) simulation platform to investigate and compare multiple adjustment methods. Features of μTOSS (http://mutoss.r-forge.r-project.org/) are

- Open Source code implementation (using R)

- Well-documented developer interfaces for new procedures to add-on
- Graphical user interface (GUI)
- Online user's guide on which procedure to use according to the user specification
- Inclusion of a large part of the known Multiple Comparison Procedure methods
- Inclusion of tested datasets for verification and exemplary purposes
- Simulation Platform
- Ongoing maintenance

There has been an ongoing discussion in the biomedical community on the best clinical trial design for the identification and validation of predictive biomarkers. At this time, there are three major classes of designs proposed for the evaluation of a biomarker-guided therapy and the assessment of biomarkers in clinical practice, see Sargent et al. (2005), Simon (2008), or Freidlin et al. (2010):

1. Targeted Trial Design (or Enrichment Design)
2. Biomarker Stratified Design
3. Biomarker Strategy Design

Sample size considerations for the biomarker stratified and biomarker strategy designs to assess the clinical utility of predictive biomarkers have been made by Richard Simon, see http://linus.nci.nih.gov/brb/samplesize/index1.html. Current recommendation and practice is to use the biomarker-stratified design since it validates predictivity of a marker best (Freidlin 2010). For the validation of predictive biomarkers one should

i. provide reproducible biomarker information
ii. test in a randomized setting before use in clinical practice
iii. apply a biomarker-stratified design.

However, it may take years until biomedicine will know whether the choice of the design today will have been the best one. Computational statistics should contribute to make this time span shorter.

2.2 Combining the Analysis of Molecular Data for Prognosis

The standard approach has been so far the application of a regression model based on a $n \times p$ data matrix $\mathbf{X}$ representing one single source of data, e.g. gene expression, where the sample size n range around 102 and the dimension of the individual observation between 104 and 106. Given the availability of multiple data sources it would be more awarding when searching for prognostic and predictive factors when all available data would be used in an integrative approach to generate one single risk prediction model based on a combination of different sources $\mathbf{X_a}$, $\mathbf{X_b}$, $\mathbf{X_c}$, etc. (e.g. methylation and gene expression data). Since a solution of this problem might be either not

feasible at all or may lead to unstable results with unsatisfying prediction performance compared to single data source based prediction a strategy must be defined on how dealing with more than one data source. In biomedicine an integrative approach is not new at all. Since decades clinicians combined several types of data, e.g. data from physical examinations and hematological laboratory data. The role of the traditional hematological data can be thought of being taken over now by the array data, moving the hematological laboratory nearer to the traditional clinical data.

For future basic medical research it is relevant to know the added value provided by the molecular data. Since usage of p-values is no longer an option a measure characterizing prediction accuracy should inform in particular on the performance of future patients on the treatment selection. Binder and Schumacher (2008) used the bootstrap sampling without replacement for efficient evaluation of prediction performance without having to set aside data for validation. Conventional bootstrap samples, drawn with replacement could be severely biased and such translate to biased prediction error estimates, often underestimating the amount of information that can be extracted from high-dimensional data.

Combining clinical data with one high-dimensional data set (Boulesteix et al. 2008; Benner et al. 2010 or Bovelstad et al 2007) has been quite common since the availability of microarrays. Methods for pre-processing, dimension reduction and multivariable analysis are available as well. It has even become a business in advanced education when e.g. a Cold Spring Harbor Laboratory conference on "Integrative Statistical Analysis of Genome Scale Data", June 8 - 23, 2009 educates in a course for about 3500$ on how to combine different genomic data sources, e.g. to model transcriptional networks through integration of mRNA expression, ChIP, and sequence data.

More appropriate would be a comprehensive integrative approach of risk predictive modeling that would stepwise narrow down the list of candidate predictors. An open question is, however, in which order to proceed with the available data sources. Since the number of sources is small one could try all possible orderings, however the number of predictors could differ by orders of magnitude in this case. Another question would be whether it would be useful to link the data sources sequentially, e.g. by using information from the analysis of data from a first data source for modeling data from a second source, or how to analyze them in parallel. One should also not underestimate technical problems like model misspecification, limited number of replicates, limited computing time or the use of asymptotical test statistics. One has to outweigh the influence of the different factors when planning as well as when interpreting the results, elements of research which are often missing.

3 Outweighing Flexibility and Complexity Using Adaptive Designs

For handling an increasing number of new anticancer compounds, clinical drug testing is pressed by practical, economical and ethical demands for increasing degree of flexibility in the design and the conduct of a clinical study. Adaptive group sequential designs allowing e.g., sample size recalculation, have become critical for overcoming the bottleneck of treatment options and making drugs sooner available to patients. When using adaptive designs, the further course of the trial depends on the data observed so far, the decision about how to continue (effecting e.g. final sample size, selection of treatment arms, choice of data modelling).

Bretz et al. (2009) recommended adaptive designs in confirmatory clinical trials since "*It is a difficult, if not unsolvable problem to completely foresee at the design stage of a clinical trial the decision processes at an interim analysis since other consideration than the observed efficacy results may influence the decision*". However, when evolving scientific expert knowledge and additional unknown background information not available at the planning phase becomes part of an adaptive design, it is essential to understand the operating characteristics before the start of an actual trial. Full scale clinical trial simulations are crucial to describe and analyse the features of such designs. Thus the behaviour of the decision rules can only be described by constructing "real" data for possible interesting scenarios and estimating design features, such as e.g. type I error rate, power, average sample number, from iterated computational simulations of the whole study course. This means there are three major challenges for computer simulations in evaluating the features of a specific adaptive design:

1. The potential decisions during the course of the study have to be specified in advance as detailed as possible to simulate scenarios which depict the closest the reality and hence will allow valid inferences.
2. Computer programs should be built in modules to allow easy implementation of different kind of adaptations. So the flexibility of the adaptive approach will be also maintained in the implementation. Figure 1 displays a study simulations scheme for two-stage adaptive design where design adaptations may be executed in one interim analysis.
3. There is an infinite number of scenarios or parameter settings under which the specific adaptive approach could be simulated. To get relevant results for the considered clinical study situation one has to identify parameter settings which will be probable to occur in reality (e.g. realistic accrual rate, probable true treatment effect, possible loss to follow-up, potential influence of nuisance covariates).

We report here experience with a simulation set up to enable an investigational randomized two arm phase II study for the rare subtype of non-clear

cell renal cell carcinoma (ncc-RCC) when two novel molecularly targeted agents (Sunitinib and Temsirolinus) were examined for progression-free survival (PFS) as primary endpoint. At planning of that trial there little was known on the activity of each of the drugs in ncc-RCC patients as well as on the achievable difference of the activity between both. A restriction of the study were limitations in both funds and patient horizon which enforced a sample size not largely exceeding $n = 50$ per arm and $N = 100$ in total. Further, the study time was limited to three years accrual and one year follow-up. After about 30 patients, an interim look was foreseen with the possibility to stop for futility or unfeasibility (if the estimated necessary number of patients could not be recruited in the remaining accrual time) and with recalculation of the sample size for the second study stage if the trial continues. A two-stage group sequential design with type I and type II error spending was established where the second stage was adapted for sample size recalculation using the conditional rejection probability principle of Schäfer and Müller (2001). Since the endpoint of interest was a right-censored failure time, sample size recalculation for the second study stage has to account for patients where recruited in the first stage but will still be under observation the second. The sample size of such a failure time study is determined via the number of events needed to achieve the overall power $1 - \beta$ of the study and, when recalculated after the 1^{st} stage as new number of events Δd needed to achieve the conditional power $1 - \beta_{\text{cond}}$ cond. For details see Wunder et al. (2010). The study course is depicted in Figure 2.

The restricted number of patients which could be recruited, enforces to execute simulations under conventional as well as "investigational" high type I and type II error probabilities $(\alpha, \beta) \in \{(0.5, 2), (0.1, 0.2), (0.05, 0.3), (0.1, 0.3), (0.2, 0.3), (0.2, 0.2), (0.3, 0.3)\}$. The "traditional" choices for the error probabilities lead to unachievable sample sizes and only when allowing for unconventionally high error rates the expected sample sizes are near 100 patients. Interim looks are implemented after 30%, 40% and 50% of the expected total event number to finally choose an interim analysis after approximately 30 patients. The uncertainty about the difference between treatment arms causes the need to simulate under a wide range of true treatment effects, i.e. log hazard ratios in $\{0, log(10/7), log(11/7), log(12/7), log(14/7)\}$ to assess the impact of true hazard ratios which differ from the clinically relevant effect of $log(11/7)$. This means, to cover all interesting and relevant design settings, 105 different simulation scenarios were executed.

A large scale simulation study was constructed for illustration of the design and for determining within a set of scenarios that design which meets the desired properties of the planning agreement between principle investigator, sponsor, funding partner and biostatistician. When interpreting simulation results, one has also to keep in mind that simulation results may deal with different sources of inaccuracy. For example, when analyzing type I error rates in adaptive survival trials there may be potential influence of misspecified

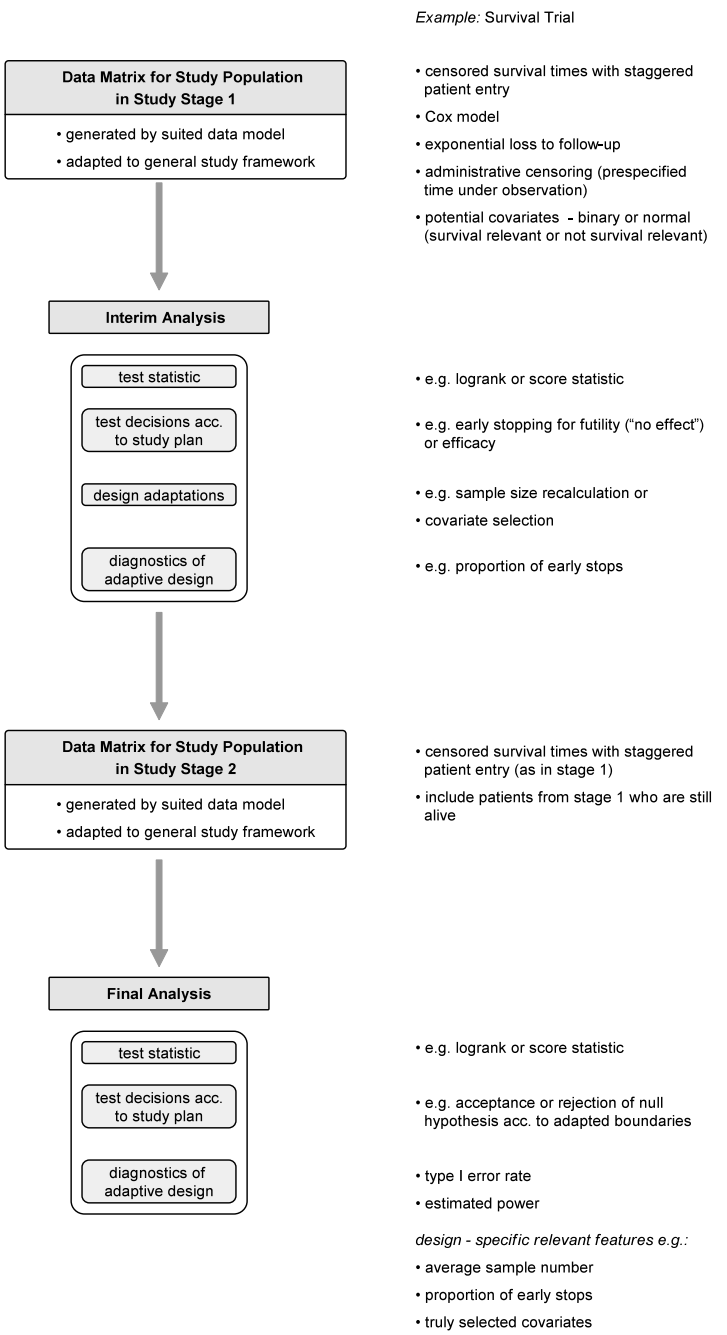

Fig. 1. Diagram of study simulations for two-stage adaptive designs where design adaptations may be executed in the interim analysis.

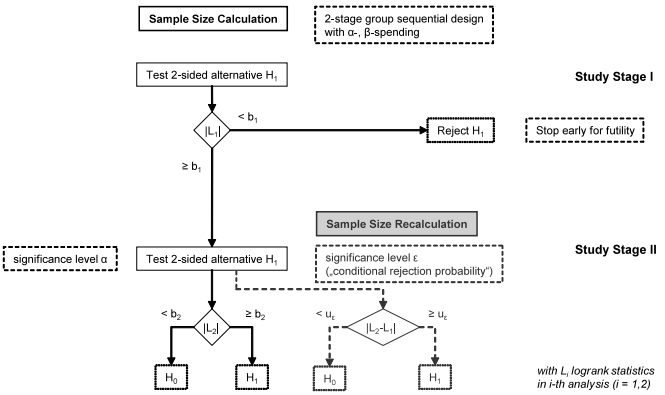

Fig. 2. Diagram of two-stage phase II study in ncc-RCC patients with possibility to stop for futility after the first stage and sample size recalculation for the second stage.

models (see e.g. Lagakos and Schoenfeld (1984)), simulation error (number of replicates may be limited by time) or asymptotical effects in the test statistics (e.g. when considering logrank or score tests, see e.g. Tsiatis et al. (1985). Thus, one has to outweigh the influence of different factors when interpreting simulation results.

4 Discussion

With the appearance of molecular sequence data and microarray data and with the intrinsic problems of screening and archiving these new and massive data sets grew in biomedicine the impression that bioinformatics tools would be the most appropriate methods to analyze these data. Data mining and clustering methods were overestimated in their potency and computer programs were just applied without a thorough statistical analysis of the research problems and the properties of ad hoc generated optimization algorithms. The two examples used above to illustrate the fruitful interaction between computational statistics and molecular biomedicine is the tip of the iceberg. There are many more examples and there are computational methods which are much more involved in those as well as in the examples above where one may look forward for further development.

The role of biostatistics and computational statistics has been recognized also in the bioinformatics community, see e.g. the announcement of courses in the internet with a list of contents like: *Descriptive statistics, Distributions, Study design, Hypothesis testing/interval estimation, Non-parametric methods, Analysis of variance, Linear regression, Multiple testing, The statis-*

tical program R. As biostatistician one can be proud about the fact that the classical disciplines of statistics are taught in bioinformatics departments but one may also wonder why researchers untrained in statistics actually could start to work in Bioinformatics. Of more serious concern is however, when instructors propose the usage of normal distribution theory for analysis of complex molecular data and define as goal: to "*Handle the symbolic language of statistics and the corresponding formalism for models based on the normal distribution*". Good is when such a course is committed to a statistical programming language. In that "computational" respect, bioinformaticians were from the beginning more determined and more prudent than biostatisticians. As long as leading biomedical researchers confuse biostatistics with bioinformatics, if they realize statistics at all as necessary for the analysis of the molecular data, biostatistics and computational statistics has to articulate is contribution to biomedical science for better designs and for better analyses of high-dimensional genomic data. Computational statistics methods are strongly required for exploratory data analysis and novel means of visualizing high-dimensional genomic data as well as for quality assessment, data pre-processing, and data visualisation methods. The R packages and the Bioconductor project have taken a promising lead to improve the situation. Yet, one should recognize that computational statistics being it science, technology or something special in between is still young, below age 50 when one remembers the start in UK in December 1966 and in the USA on February 1967, and we should give it time.

Recently the journal "Computational Statistics & Data Analysis" (CSDA) launched a second special issues on "Computational Statistics within Clinical Research" where the call explicitly asks for submission of work for "*understanding the pathogenesis of diseases, their treatment, the determination of prognostic and predictive factors, and the impact of genetic information on the design and evaluation of clinical outcomes*". Such activities at the interface between biomedicine and computational statistics may add further to bridge the gaps.

5 First Author's epilogue

When I started my career in biostatistics at the German Cancer Research Center three decades ago it took only a few months to realize the importance and relevance of computational statistics for both the research work in applied statistics and the biostatistical consulting of clients and partners coming from all relevant biomedical fields of experimental and clinical cancer research. At that time - just after the appearance of John Tukeys book on Explorative Data Analysis (Tukey (1977) resources on computational statistics methods and literature were rare.

The journal Computational Statistics & Data Analysis (CSDA) which later became the flagship publication of the International Association of Sta-

tistical Computing (IASC) as well as the journal "Computational Statistics Quarterly", now "Computational Statistics" were then just starting. However there were two easy accessible resources which imprinted my relationship with computational statistics for ever: the receipt of the "Statistical Software Newsletter" (SSN), founded already in 1975, with methods and algorithms at that time hardly needed for the rising computational needs for the analysis of clinical survival data (Edler et al. 1980), and the attendance of COMPSTAT conferences for the exchange with the colleagues interested at the interface between statistics and computing.

It was the 5^{th} COMPSTAT in Toulouse in 1982 where I started to report methods developed for biostatistical applications and I enjoy now how COMPSTAT has grown and developed by 2010 and its 19th COMPSTAT in Paris, again in France. The passage of the years has not diminished my respect and my inclination to that forum of scientific exchange nor my pride of having had the honor to serve IASC as officer for some time, our society IASC which shields as member of the ISI family the COMPSTAT conference.

References

BENJAMINI, Y. and HOCHBERG, Y. (1995): Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B 57, 289-300.*

BENNER, A., ZUCKNICK, M., HIELSCHER, T., ITTRICH, C. and MANSMANN, U.(2010): High-dimensional Cox models: the choice of penalty as part of the model building process. *Biometrical Journal 52, 50-69.*

BINDER, H. and SCHUMACHER, M. (2008): Adapting Prediction Error Estimates for Biased Complexity Selection in High-Dimensional Bootstrap Samples. *Statistical Applications in Genetics and Molecular Biology 7 (1) Article 12.* DOI: 10.2202/1544-6115.1346

BOULESTEIX, A.L., STROBL, C., AUGUSTIN, T., DAUMER, M. (2008): Evaluating microarray-based classifiers: An overview. *Cancer Informatics 6, 77-97.*

BOVEHSAD, H.M., NYGARD, M., STORVOLD, S., ALDRIN, H.L.L., BORGAN, O., FRIGESSI, A. and LINGGIAERDE, O.C.C. (2007): Predicting survival from microarray data - a comparative study. *Bioinformatics, 23(16), 2080-2087*

BRETZ, F., KOENIG, F., BRANNATH, W, GLIMM,E. and POSCH, M. (2009: Tutorial in biostatistics: Adaptive designs for confirmatory clinical trials, *Statistics in Medicine 28, 1181-1217.*

CHAMBERS, J. (1999): Computing with data: concepts and challenges. *The American Statistician 53, 73-84.*

DICKHAUS, T., BLANCHARD, G., HACK, N., KONIETSCHKE, F., ROHMEYER, K., ROSENBLATT, J., SCHEER, M. and WERFT, W. (2010): μTOSS - Multiple hypotheses testing in an open software system. 2nd Joint Statistical Meeting DAGSTAT Statistics under one Umbrella March 23-26, 2010, Technical University, Dortmund, p 77 (abstract).

DUDOIT, S, and Van DER LAAN, M. (2007) *Multiple Testing Procedures with Applications to Genomic.* Springer, New York

EDLER, L. (2005): Computational Statistics und Biometrie. Wer treibt wen? *GMDS Medizinische Informatik, Biometrie und Epidemiologie 1 (2)* doi10 - (20050620), 2005.

EDLER, L, WAHRENDORF, J. and BERGER, J. (1980): SURVIVAL. A program package for the statistical analysis of censored survival times. *Statistical Software Newsletter 6, 44-53*

FINNEY, D.J. (1974): Problems, data, and Inference (with discussion). *Journal of the Royal Statistical Society Series A 137, 1-22.*

FISHER R.A. (1925): *Statistical Methods for Research Workers.* Oliver & Boyd, Edinburgh.

FREIDLIN, B,, MCSHANE, L. and KORN, E.L.(2010): Randomized Clinical Trials With Biomarkers: Design Issues. *Journal National Cancer Institute 102, 152-160*

KAUFMANN, M., PUSZTAI, L., RODY, A., CARDOSO, F., DIETEL, M., EDLER, l., HAHN, M., JONAT, W., KARN T., KREIPE. H., LOI, S., VON MINCKWITZ, G., SINN, H.P. and VAN DE VIJVER (2010/11): Use of standard markers and incorporation of molecular markers into breast cancer therapy (to appear).

LAURO, N.C. (1996): Computational statistics or statistical computing, is that the question? *Computational Statistics and Data Analysis 23, 191-193.*

L€UTER, J. (1981): *Programmiersprache DIST. Dateneingabe und Datenstrukturierung.* Akademie Verlag, Berlin.

LAGAKOS, S. W. and SCHOENFELD, D. A. (1984): Properties of proportional-hazards score tests under misspecified regression models. *Biometrics 40,1037-1048.*

NELDER, J. (1996): Statistical computing. In P. Armitage and D. Hand (Eds.): *Advances in Biometry. 50 Years of the International Biometric Society.* Wiley, New York, 201-212.

POTTER, M. (2005): A permutation test for inference in logistic regression with small- and moderate-sized data sets. *Statistics in Medicine 24 (5), 693-708.*

TSIATIS, A., ROSNER, G. L. and TRITCHLER, D. L. (1985): Group sequential tests with censored survival data adjusting for covariates. *Biometrika 72(2), 365-373.*

SARGENT, D.J., CONLEY, B.A., ALLEGRA, C., et al. (2005): Clinical trial designs for predictive marker validation in cancer treatment trials. *Journal of Clinical Oncology 23(9), 2020-2027.*

SCHÄFER, H. and MÜLLER, H. (2001): Modification of the sample size and the schedule of interim analyses in survival trials based on data inspections. *Statistics in Medicine 20, 3741-3751.*

SIMON, R. (2008). Using genomics in clinical trial design. *Clinical Cancer Research 14, 5984-5993.*

VICTOR, N. (1984): Computational statistics - science or tool? (with discussion). *Statistical Software Newsletter 10, 105-125.*

WUNDER, C., KOPP-SCHNEIDER, A, and EDLER, L. (2010): *Adaptive group sequential trial to improve treatment in rare patient entities* (to appear).

TUKEY, J.W. (1977): *Exploratory Data Analysis.* Addison-Wesley, New York.

http://www.springer.com/978-3-7908-2603-6

Proceedings of COMPSTAT'2010
19th International Conference on Computational StatisticsParis France, August 22-27, 2010 Keynote, Invited and Contributed Papers
Lechevallier, Y.; Saporta, G. (Eds.)
2010, XXX, 621 p. 105 illus., Softcover
ISBN: 978-3-7908-2603-6
A product of Physica-Verlag Heidelberg