The scientific work of Peter Bauer

Werner Brannath, Franz König, and **Martin Posch***

Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria

This June, Peter Bauer – chief editor of Biometrical Journal from 2000–2003 and member of the editorial board since 1996 – celebrated his 65th birthday. Peter Bauer became known as one of the originators of the "adaptive designs", a field of research that recently gained considerable popularity, not only among academia but likewise among regulatory bodies and the pharmaceutical industry. It is characteristic for Peter Bauer, that while "adaptive designs" make their way into main stream statistics, he is already on the search for new challenges and now works, among others, on statistical problems relevant to bioinformatics. His sense for dynamic research areas can be easily seen from his whole professional career where he covered a wide range of research fields including informatics, binomial and poisson processes, model robustness, optimal designs, multiple testing, sequential analysis, adaptive designs and bioinformatics.

Peter Bauer started his scientific career at the medical faculty of the University of Vienna, was appointed 1985 as head of the Department of Medical Documentation and Statistics in Cologne and returned to the University of Vienna in 1994 to become the head of the Department of Medical Statistics at the medical faculty. Besides his co-operation in many clinical research projects he built up an active research group. He currently decided to accept the lead of the Core Unit for Medical Statistics and Informatics, the superordinate unit of his and several other departments since the faculty became an independent university in 2004. In view of the proceeding structural changes at the new Medical University taking this lead is just one more example for his persistent willingness to explore new challenges.

As a physicist by education, Peter Bauer was part of the pioneering epoch of medical computer science in Vienna. In the late sixties he developed a programming language to model chemical reactions and worked on computer based diagnostic programs for liverdiseases. Confronted with the need for statistical expertise at the medical faculty, he turned to statistics and became a British Council Research Fellow at the Department of Statistics in Edinburgh where David Finney was department head at this time. He returned to Vienna, fascinated by the British school of medical statistics, and got involved in several challenging consulting projects that were the starting point of methodological research papers. Motivated by practical experience with high dimensional compartment models he developed a method to specify the power for the discrimination between different regression models and investigated the related statistical design concept nowadays known as T-optimality. The method was published in two research papers of the "Biometrische Zeitschrift" (Bauer, 1975a, b) one year before the journal was renamed to "Biometrical Journal". Another example for his practically oriented methodological research is the analysis of a cross-sectional study on the caries protective effects of fluorid. To model the hazard for the occurrence of caries Peter Bauer forestalled the development of frailty models (Bauer et al., 1980, 1982). In a fruitful cooperation with Peter Hackl he contributed to the theory of quality control and multiple testing. Peter Bauer published his first paper on multiple comparisons in 1985 in the Biometrical Journal (Bauer and Hackl, 1985). The multiplicity issue in statistical inference was the subject of many subsequent works with applications in dose finding (e.g. Budde and Bauer, 1989; Bauer and Budde, 1994; Bauer et al., 1998; Bauer, Brannath and Posch, 2000), model selection (Bauer, Pötscher and Hackl, 1988), equivalence testing (e.g. Bauer and Bauer, 1994)

* Corresponding author: e-mail: Martin.Posch@meduniwien.ac.at, Phone: +43 1 40400 7486, Fax: +43 1 40400 7477

and confirmatory clinical trials (e.g. Bauer, 1991). Another aspect of multiplicity that attracted Peter Bauer's interest were multiple interim analyses in group sequential trials. It was Franz Xaver Wohlzogen, the founder of the Medical Statistics Department at the University of Vienna, who sparked his interest in sequential designs, among other things with the development of a sequential pregnancy test where the women's urine was injected into male toads. Here the sequential design helped to reduce the required number of animals and to save time. Peter Bauer wrote a book on sequential statistical methods together with Viktor Scheiber and Franz Xaver Wohlzogen (Bauer, Scheiber and Wohlzogen, 1986), introduced a new concept of power spending (Bauer, 1992), and made several other methodological contributions. His most important contribution was to add an entirely new feature to group sequential designs – the possibility to adapt the trial design based on unblinded interim data without compromising the type I error rate: in two seminal papers (Bauer, 1989; Bauer and Köhne, 1994) he developed the "combination tests" that control the significance level even if no specific adaptation rule is pre-specified. These tests are based on a simple but powerful principle: the combination of standardized test statistics whose null distributions are invariant under all adaptations. Peter Bauer's original motivation for the methodology was to find a sound way to handle the frequent "protocol amendments" that were and still are common practice in clinical trials. However, the availability of a statistical methodology that allowed to learn from interim data for the design of the remaining part of the trial, opened a wide range of possibilities to gain from design adaptations like changing sample sizes or test statistics (Lang, Auterith and Bauer, 2000). Two major advancement were also made by Peter Bauer with co-authors. The first was the coupling of combination tests with the closed testing principle which opens the possibility of adaptations in a multiple testing framework, e.g., the selection of treatment arms or endpoints in an interim analysis (Bauer and Kieser, 1999; Kieser et al., 1999). The second advancement was a method to include or drop interim analyses during an ongoing trial (e.g. Brannath, Posch and Bauer, 2002). Peter Bauer further enriched the methodology with several other research papers (e.g. Bauer and Röhmel, 1995; Posch and Bauer, 1999; König, Bauer and Brannath, 2006) and recently linked it to bioinformatics (Goll and Bauer, 2007; Zehetmayer, Bauer and Posch, 2005). The simplicity and generality of the combination tests attracted an increasing international group of statisticians and a new statistical research field was created. Only recently the pharmaceutical industry and authorities became aware of the potential of these ideas and "Adaptive Designs" became the "Buzzword" in the community of pharmaceutical statisticians. Peter Bauer, however, did not share the partly naive enthusiasm for adaptive designs but pointed at the limitations in the application of clinical trials. Already in the original publication he discussed issues arising through the use of a non-sufficient test statistic and argued for a cautious interpretation of interim results. Recently he pointed at the difficulties with the estimation of conditional power (Bauer and König, 2006) and the poor reporting of adaptive trials in the medical literature (Bauer and Einfalt, 2006).

Peter Bauer's sense of responsibility goes far beyond his methodological and applied statistical research. He regularly serves as expert in data monitoring boards, ethics committees, and consultant for regulatory agencies. Due to his initiative and continuous contribution statistics finally became an obligatory part of the medical curriculum in Vienna.

We acknowledge that our review of Peter Bauer's work reflects just our subjective selection and is far from complete. We further like to express our gratitude to Peter Bauer for his energetic and cooperative collaborations. We finally wish him that he will continue with his enjoyment and success in his scientific research and professional career, in his "second profession" as jazz musician, and in his family.

References

Bauer, P. (1975). Zur Adäquatheitsprüfung von Modellen (I). *Biometrische Zeitschrift* **17**, 275–285.

Bauer, P. (1975). Zur Adäquatheitsprüfung von Modellen (II). *Biometrische Zeitschrift* **17**, 345–362.

Bauer, P. (1989). Multistage testing with adaptive designs. *Biometrie und Informatik in Medizin und Biologie* **20**, 130–148.

Bauer, P. (1991). Multiple testing in clinical trials. *Statistics in Medicine* **10**, 871–890.
Bauer, P. (1992). The choice of sequential boundaries based on the concept of power spending. *Biometrie und Informatik in Medizin und Biologie* **23**, 3–15.
Bauer, P. and Bauer, M. M. (1994). Testing equivalence simultaneously for location and dispersion of two normally distributed populations. *Biometrical Journal* **36**, 643–660.
Bauer, P., Binder, K., Husinsky, I., Kleinert, W., Künzel, W., Scheiber, V., Überhuber, C. W., Westphal, G., and Wohlzogen, F. X. (1980). Beurteilung der kariesprotektiven Wirkung von Fluoriden aufgrund von Modellen für den zeitlichen Verlauf des Kariesbefalls der bleibenden Zähne. *Zahn-, Mund- und Kieferheilkunde mit Zentralblatt* **68**, 214–230.
Bauer, P., Binder, K., Husinsky, I., Kleinert, W., Künzel, W., Scheiber, V., Überhuber, C. W., Westphal, G., and Wohlzogen, F. X. (1982). Models for irreversible processes, based on cross-sectional data. *Biometrical Journal* **24**, 399–411.
Bauer, P., Brannath, W., and Posch, M. (2001). Multiple testing for identifying effective and safe treatments. *Biometrical Journal* **43**, 605–616.
Bauer, P. and Budde, M. (1994). Multiple testing for detecting efficient dose steps. *Biometrical Journal* **36**, 3–15.
Bauer, P. and Einfalt, J. (2006). Application of adaptive designs – a review. *Biometrical Journal* **48**, 493–506.
Bauer, P. and Hackl, P. (1985). The application of Hunter's inequality in simultaneous testing. *Biometrical Journal* **27**, 25–38.
Bauer, P. and Kieser, M. (1999). Combining different phases in the development of medical treatments within a single trial. *Statistics in Medicine* **18**, 1833–1848.
Bauer, P. and Köhne, K. (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics* **50**, 1029–1041.
Bauer, P. and König, F. (2006). The reassessment of trial perspectives from interim data – a critical view. *Statistics in Medicine* **25**, 23–36.
Bauer, P., Pötscher, B. M., and Hackl, P. (1988). Model selection by multiple test procedures. *Statistics* **19**, 39–44.
Bauer, P. and Röhmel J. (1995). An adaptive method for establishing a dose-response relationship. *Statistics in Medicine* **14**, 1595–1607.
Bauer, P., Röhmel, J., Maurer, M., and Hothorn, L. (1998). Testing strategies in multi-dose experiments including active control. *Statistics in Medicine* **17**, 2133–2146.
Bauer, P., Scheiber, V., and Wohlzogen, F. X. (1986). *Sequentielle statistische Verfahren*. Gustav Fisher Verlag, Stuttgart.
Brannath, W., Posch, M., and Bauer P. (2002). Recursive combination tests. *Journal of the American Statiscal Association* **97**, 236–244.
Budde, M. and Bauer, P. (1989). Multiple test procedures in clinical dose finding studies. *Journal of the American Statiscal Association* **84**, 792–796.
Goll, A. and Bauer, P. (2007). Two-stage designs applying methods differing in costs. *Bioinformatics* **23**, 1519–1526.
Kieser, M., Bauer, P., and Lehmacher, W. (1999). Inference on multiple endpoints in clinical trials with adaptive interim analyses. *Biometrical Journal* **41**, 261–277.
König, F., Bauer, P., and Brannath, W. (2006). An adaptive hierarchical test procedure for selecting safe and efficient treatments. *Biometrical Journal* **48**, 663–678.
Lang, T., Auterith, A., and Bauer, P. (2000). Trendtests with adaptive scoring. *Biometrical Journal* **42**, 1007–1020.
Posch, M. and Bauer, P. (1999). Adaptive two stage designs and the conditional error function. *Biometrical Journal* **41**, 689–696.
Zehetmayer, S., Bauer, P., and Posch, M. (2005). Two-stage designs for experiments with a large number of hypotheses. *Bioinformatics* **21**, 3771–3777.

