International Chinese Statistical Association

http://www.icsa.org

泛華統計協會

Features:

Marvin Zelen-
Statistical Scientist

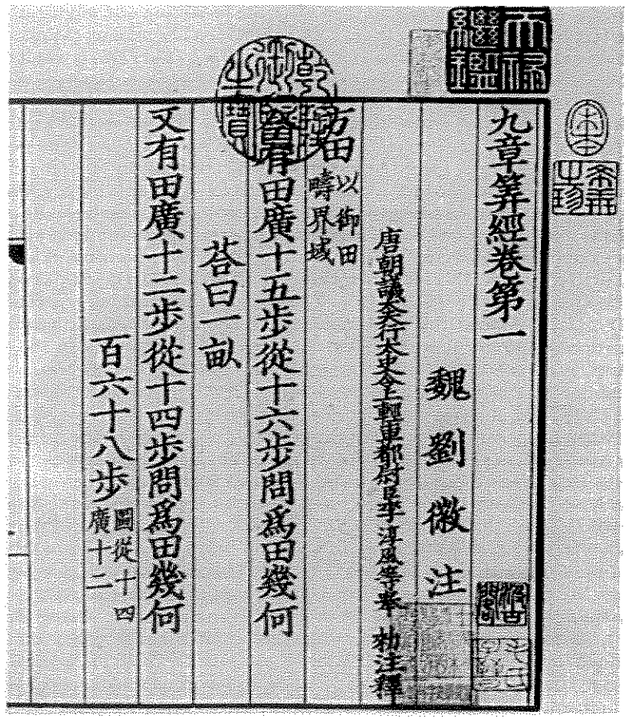
Risk Management

Statistical Issues

Meeting Announcements

Career Opportunities

Bulletin January 2004


九章算經卷第一
魏劉徽注
唐朝議奉行太史令上輕車都尉臣李淳風等奉 勅注釋
方田 以御田疇界域
今有田廣十五步從十六步問爲田幾何
荅曰一畝
又有田廣十二步從十四步問爲田幾何
荅曰一畝百六十八步 圖從十四廣十二

Nine Chapters on the Mathematical Art

九章算術

九章算術 **is a very influential book in the history of Chinese mathematics. It is the earliest specialized mathematical work in China that survives to the present day. It is unclear when this book was produced. However it is estimated that the book was first assembled at the Han dynasty (during the first century). This book contains a total of 246 questions in nine chapters (hence the name Nine Chapters). For each question in the book, there is only answer given. The method of solving the question is omitted. Here is the opening of Chapter 1 of the Nine Chapters on the Mathematical Art.**

From the Editor

Kao-Tai Tsai, Ph.D.

This issue marks the half time of our tenure as the editorial team of the Bulletin. Like any event, it is a good time to review what we have done during the first half of our editorial service to the ICSA and what we hope to accomplish in the second half of our tenure.

We took a bold experiment and received quite a few pleasing responses from our members. Many Committee Chairs and friends from all parts of the world were asked to report their activities. For every issue, the editorial team had to repeatedly ask for contributions. All we can say is that the results are mixed at best. The Committee Chairs and the ICSA officers can certainly be more informative to the ICSA members about their accomplishments and plans in growing this organization. The members can certainly be more active in voicing their opinions too. Regrettably to say, something seems to be lacking, as we can see from the contents of this issue.

As stated in the previous issues, we would like to make this an informative and intellectually beneficial publication to our readers. We all know how difficult that will be without the ingredient from the members. For those friends who help to pull this issue off, we would like to express our sincere appreciation. For the rest of the members, your contributions are urgently needed.

In this issue, in addition to the official reports of the ICSA activities, we present one of the oldest records in ancient Chinese mathematics, and one of the most useful strategies, from the wisdom of a classic military strategist, Suwen-Tse, for our daily dealings. We also show you the topics of financial risk management and recent advances in clinical trials. On top of that, the story about Professor Zelen will surely make an interesting reading.

We hope you will enjoy the fruit of our labor in this issue. We also hope to have more to report next time. And, as usual, your help in any ways will be greatly appreciated.

HAPPY NEW YEAR!

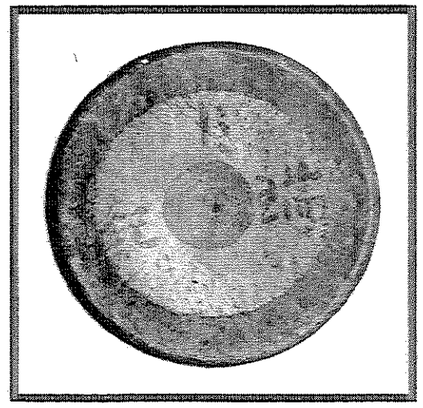
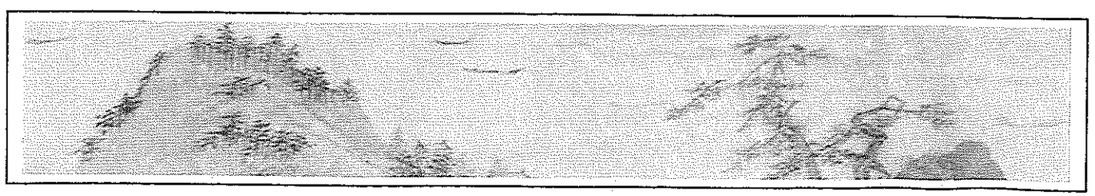


Table of Contents
ICSA Bulletin, January 2004

Editorial Members
ICSA Bulletin, January 2004

Printing
Global Graphics, Inc.
1945 Rt. 27
Edison, New Jersey, USA



Committee Chairs
ICSA, 2004

Board of Directors

Ngai Hang Chan	William W.S. Wei
Chen-Hsin Chen	Zhiliang Ying
Rong Chen	Heping Zhang
Mei-Ling Lee	Jiahua Chen
Guo-Ying Li	Frank Shen
Zhaohai Li	Greg Wei
Don Sun	Ivan Chan
Naitee Ting	Shu-Yen Ho
Sue-Jane Wang	Xumin He

Standing Committee
Program Committee:
Naitee Ting
Finance Committee:
Weiying Yuan
Nominating & Election Committee:
Zhaohai Li
Publication Committee:
Xuming He
Constitution Committee:
Frank Shen

Current Committee
Membership Committee:
Jun Zhao
Fundraising Committee:
Daniel Wang
Public Relations Committee:
Colin Wu
Awards Committee:
Jianping Dong
Communication Committee:
Don Sun
Applied Statistics Symposium Committee:
Nancy Lo
Book & Journal Donation Committee:
Tar Timothy Chen
Annual Meeting Committee:
Jiahua Chen
Archive Committee:
Yi Tsong
Strategic Committee:
Zhiliang Ying

Submission Guidelines
ICSA Bulletin

Articles
The ICSA Bulletin welcomes the submission of articles by the members. Articles submitted should be written in traditional Chinese or English. The contents should be aimed at the general reading.

Articles can be submitted electronically or in papers. If submitted in papers, please leave ¾ inches of blank space on each margin of the letter size paper. The font size of 11 or 12 with Microsoft Word is preferred when typing.

Submission deadline for the January issue is December 15 and the deadline for the July issue is June 15. Articles received after the deadline will be published in the following issue of Bulletin.

Advertisements
The ICSA Bulletin welcomes the submission of advertisements related to the profession of Statistics. The fee for the whole page advertisement is US$300.00. Please contact the Editor for the fee schedule of advertisement with other sizes.

Questions
Please submit your questions to the Editor by email at tsai0123@yahoo.com.

For the complete list of Committee Members and terms, please go to the ICSA website: www.icsa.org.

From the President

Frank Shen, Ph.D.

Leadership Begins with a Small "l"

I am thrilled to have the opportunity to serve ICSA as its 2004 President. While ICSA has experienced another year of success under the leadership of Professor Zhiliang Ying and Dr. Yi Tsong, I can't stop thinking what else I can do in order to improve and reenergize this organization that has great traditions. I have come to the conclusion that the highest agenda item we need to address is to develop our people pipeline. Namely, to increase our membership and attract more passionate current members to join the leadership team of ICSA.

Let's first examine what we have accomplished so far. Since ICSA was founded at the 1987 Joint Statistical Meetings in San Francisco, we now hold one highly successful Applied Statistical Symposium annually and one international conference every three years. The program committee chaired by Dr. Naitee Ting has the responsibility to coordinate all these events. We have a distinguished journal STATISTICA SINICA co-sponsored by the Institute of Statistical Science, Academia Sinica, in Taiwan. The journal was first launched in January 1991 and has been published quarterly with original work in all areas of statistics including theory, methods, and applications. We also publish an outstanding ICSA Bulletin twice a year. Drs. Sue-Jane Wang and Kao-Tai Tsai are the former and current editor-in-chiefs leading this spectacular effort. The publication committee led by Professor Jun Shao oversees the publication policy for the ICSA. We now have a comprehensive web site. Thanks to Dr. Don Sun's untiring effort and his leadership to the communication committee. In addition to those committees that have been mentioned, ICSA has a Board of Directors, a constitution and by-law, and three other standing committees (Finance, Nominating and Election, Constitution), nine other current committees (Membership, Fundraising, Public Relations, Awards, Applied Statistical Symposium, Book and Journal Donation, Annual meeting, Archive, Strategic), and a Biometrics section. Each plays an important role in strengthening the Association by many leaders who are involved.

Growing in membership is a key to the vitality of our association. Our membership provides the leadership and performs the majority of the work for our various committees. Our publications thrive because of the contributions of the editors who believe in "Today's reader will be tomorrow's leader." The symposium and conference is highly successful because of the conscientious work of the program committee and related current committees. The finances of the Association are carefully scrutinized and monitored by the Finance committee. In other words, ICSA is an "organ" that requires new blood and new nerves all the time to keep all these activities running vigorously. We need abundant membership to participate and many new leaders to keep the blood circulating. Membership recruitment and retention is clearly not a simple task. In fact, most professional societies including ASA have faced plateaued or declined membership in the last few years. ICSA is no exception. It is an ongoing project that requires help from all ICSA members since we are the ones that can still recall the motivating factors that caused us to join the ICSA. I will work very closely with the membership committee this year and I would love to hear your ideas and what those factors might be. While there are many challenges in establishing new channels to attract new members effectively, there must be an ICSA way that we can identify.

Finally, I would like to come back to the title which I chose to highlight my first communication to you. It is about the leadership and how we can keep developing those leaders we need for the future of ICSA. We not only need more members to join the Association, but we also need more passionate leaders to run the many things that I mentioned. Some people assume that leaders are mostly born, not made. Having worked in the industry sector and primarily in an R&D business environment for more than a decade, I disagree. On the contrary, I have observed many senior leaders who were "made" through small

leadership. I never felt that I could accomplish anything alone in my career. Every project I have worked on has components that I can control and those that I can't. I grew up in my ranks by volunteering and taking the lead on those components that I can control well regardless of how big or small the job would be, followed the lead of others on those parts that I had no control of, but eventually learned, so I could gain control next time. To me, leadership is a quality. Leadership skills can be learned by working through people to create the kind of conditions under which an innovative breakthrough might best occur. Leadership is not a destiny such as a managerial position, but a journey of contributing to a community practice. That community could be a workplace or a profession that you belong to and love. I have been doing many "small things" for ICSA and ASA since 1990. Small things that wouldn't result in any recognition of my own, but I always knew that I had learned something in return and I got to know a lot of people. After a while, I knew who knows what, and I knew I don't need to be the smartest person in the room because I have seen many. I followed senior leaders in many organizations, including ICSA, and understand that their leadership is not just the demonstration of acquired assets but rather, is the demonstration of the ability to acquire the assets needed for future situations. In my presidency, I would like to invite all of you to join me and many other excellent leaders in the ICSA on this leadership drive. If you would like to see ICSA grow and be a part of this growth, all you need to do is to give me a call or send me an e-mail to express your willingness to volunteer for a small leadership. I will help you get started and stay involved, and I guarantee that it would be a rewarding experience for you.

Many important activities have been under preparation for 2004 including the Applied Statistical Symposium at San Diego in June and the 6th ICSA International Conference at Singapore in July. Please refer to the ICSA Web site and preliminary programs listed in this Bulletin. Our former Executive Director, Dr. Yi Tsong, has accomplished his three year term and I thank him for his contribution to ICSA in his tenure. Dr. Ivan Chan has been appointed to this important Executive position and I look forward to working closely with him. Personally I have resigned all my ASA appointments for one year so I can concentrate on the ICSA business. In the end, I would like to reiterate my invitation to all of you to join me and take ICSA to a higher summit, and I wish all of you a productive New Year as well.

Frank Shen, Ph.D., is Executive Director of Exploratory Development, Global Biostatistics and Programming, Bristol-Myers Squibb Co.
Email: frank.shen@bms.com

Reports From Committee Chairs

By: Naitee Ting, Ph.D., Jen-Pei Liu, Ph.D.

Program Committee

By: Naitee Ting, Ph.D.

For those of you who went to the 2003 JSM at San Francisco, you might have come to the ICSA membership meeting, or even have joined us for the nice Chinese dinner. During this JSM, Dr. Ying Lu helped us organizing the ICSA booth, recruiting new members and arranging for the dinner. On behalf of the Program Committee, I would like to thank him and the local committee for their hard work and for the successful results.

Dr. Ying Lu also helps us drafting this report for the August activity –

- The 2003 annual ICSA member meeting and banquet were held in San Francisco on August 6, 2003. The member business meeting discussed outreach for membership drives and other business issues. After the business meeting, more than 150 ICSA members and their families attended ICSA banquet at the Far East Café in San Francisco Chinatown. Besides delicious authentic Chinese food served in the banquet, our members also enjoyed Karaoke and dancing. The banquet lasted for about four hours.
- The local organization committee hopes our members had a wonderful memorable evening in San Francisco and wishes you all well in your career. We want to acknowledge following students from the Departments of Statistics at Berkeley and at Stanford: Gang Liang, Hui Tang, Jingrong Yang, Peng Zhao, Tao Shi, Jiarui Han, and Xiudong Lei, for their helps in ICSA booth and in the banquet. We also want to thank Drs. Hua Jin and Ru-Fang Yeh for their assistances for banquet. We appreciate also several senior ICSA members who attended ICSA booth during the JSM.

While every party ends, we can always look for the next. We are looking forward to seeing you again next year in Toronto.

As indicated in the previous Program Committee Report, there will be many ICSA activities in the next few years. We welcome all members to participate in, and help with these activities. Here are the planned activities:

- The 2004 Applied Statistics Symposium will take place in June 6-9, 2004, at La Jolla, CA. For information, please visit http://www.icsa.org or contact the Symposium Committee Co-chairs: Nancy Lo (Nancy.Lo@NOAA.Gov) or Gang Li (gangli@sunlab.ph.ucla.edu).
- ICSA activities for the 2004 JSM (Toronto, Canada) will be chaired by Professor Jiahua Chen.
- The Sixth ICSA International Conference will be held in the National University of Singapore, July 21-23, 2004. The organizing committees are currently working on the scientific program as well as local logistics. Updated information can be found at http://www.statistics.nus.edu.sg/ or http://www.icsa.org/.
- The 2005 Applied Statistics Symposium will take place in Washington D.C. Details regarding this symposium will be announced in the future.
- The International Conference on Multiple Comparisons (MCP) plans their next meeting at Shanghai, China, and they are interested in co-sponsorship with ICSA. Some ICSA members have been involved in the planning stage and that the MCP committee has made important progress already. The MCP will be in 2005 and the committee is chaired by Professor Jason Hsu.

Naitee Ting, Ph.D. is an Associate Director in Biostatistics, Pfizer Global Research & Development, New London, CT 06320, USA
Email: naitee_ting@groton.pfizer.com.

Biometrics Section

By: Jen-Pei Liu, Ph.D.

As drug development has become a global business, bridging studies will be more and more prevalent in the future. However, bridging studies is not limited to extrapolatability between regions or countries only. They are equally important to bridge the results between age groups, between genders, and between races. The importance of statistical methodology for evaluation of bridging studies can be seen from an invited session on bridging studies held at the 2003 Joint Statistical Meetings at San Francisco, U.S.A. and another incoming invitation session at the 2004 International Biometrics Conference in Australia in July 2004. Recently, the 2003 Symposium on Statistical Methodology for Evaluation of Bridging Evidence, organized by the Division of Biostatistics and Bioinformatics, National Health Research Institutes, Taipei, Taiwan, was held on November 15, 2003, just before the 2003 Symposium on APEC Network of Pharmaceutical Regulatory Sciences. The following is the program for the 2003 symposium on bridging studies:

- Bridging studies in Taiwan by O.Y.P. Hu, National Defense Medical Center, Taiwan.
- The use of prior information in medical studies by K.K.G. Lan, Aventis, USA.
- Three scenarios for statistical evaluation of clinical trials for global drug development by Weichung Shih, University of Medicine and Dentistry of New Jersey, USA.
- A group sequential approach to evaluation of bridging studies by Chin-Fu Hsiao, National Health Research Institutes, Taiwan.
- Clinical relevance of ethnic factors by Mey Wang, Center for Drug Evaluation, Taiwan.
- A comparison of statistical methods for evaluation of bridging studies by Jen-pei Liu, National Cheng-kung University and National Health Research Institutes, Taiwan.
- Cross-trial statistical inference: bridging vs. non-inferiority scenarios by Sue-Jane Wang, the US FDA, USA.
- Statistical issues with design and analysis of bridging clinical trials by H.M. James Hung, the US FDA, USA.
- Variance components testing in mixed effects models in bridging studies by M. Takeuchi, Kitasato University, Japan.

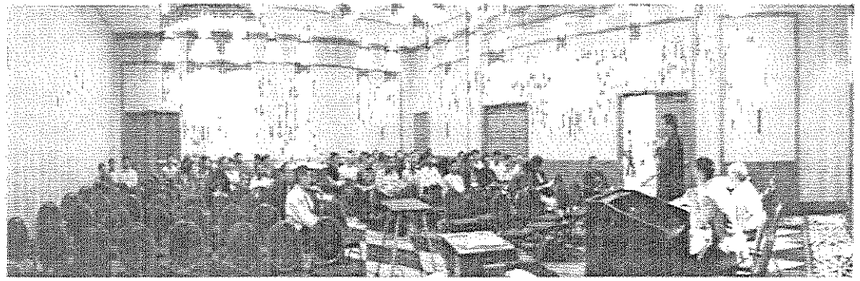
The symposium was well attended and discussion was intensive but constructive. Anyone who wants further information about the presentations at the symposium please contact me at jpliu@nhri.org.tw or jpliu@email.stat.ncku.edu.tw.

Jen-Pei Liu, Ph.D. is Professor of the Statistics Department, National Cheng-Kung University, Tainan, Taiwan.
Email: jpliu@email.stat.ncku.edu.tw

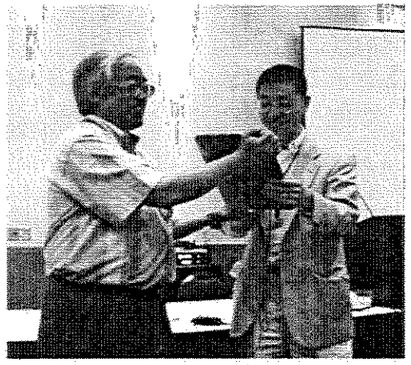
Ancient Chinese Oracle Record of Committee Reports (ca. Shang Dynasty)

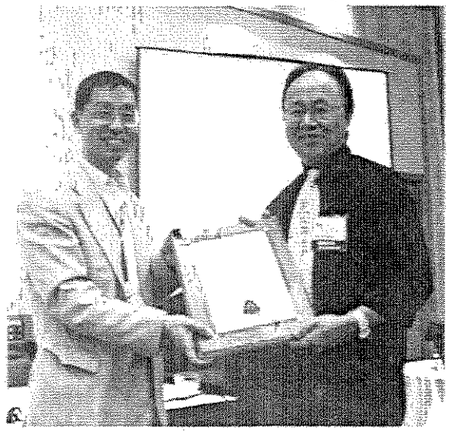
Photos of Membership Meeting & Dinner Party

By: Ying Liu, Ph.D.

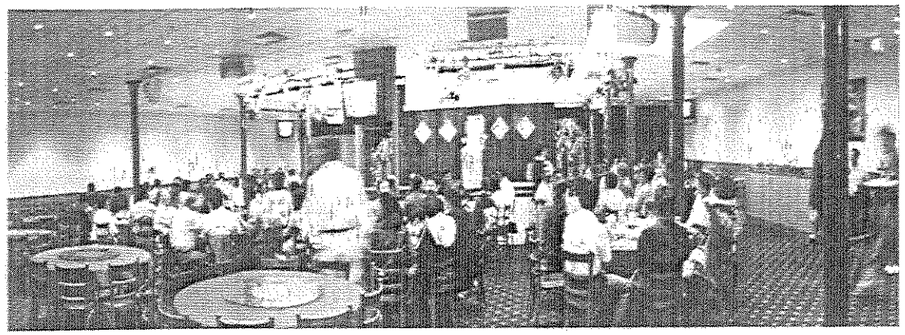


ICSA Membership Meeting





Awards Presentation



ICSA Dinner Banquet at San Francisco, California

Joyful Dancing

Beautiful Singing

STATISTICA SINICA Vol. 13, No. 4, pp. 927-1280 October 2003

STATISTICA SINICA

Volume 13 Number 4 October 2003

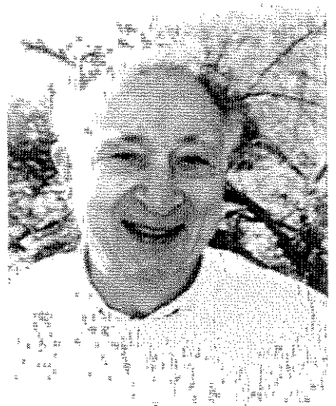
Marvin Zelen - Statistical Scientist*

By: Stephen Lagakos, Ph.D.

I've known Marvin Zelen for more than 30 years. He hired me for my first job after graduate school, at the Statistical Laboratory of the Statistical Science Division at the State University of New York at Buffalo (SUNY/Buffalo). Marvin was also my boss at the Dana Farber Cancer Institute—a teaching hospital of Harvard—and was my department chair at Harvard's Department of Biostatistics. While for the past four years I have been chair of the department of biostatistics, there is no doubt in my mind—or in the minds of others there—that we all in some sense still work for Marvin! He and I are close personal friends and have collaborated on several projects over the years. There is inadequate time to discuss in adequate detail Marvin's many activities and accomplishments, but I will try to highlight some.

**Extracted from introductory remarks made at a conference in Bordeaux, France on September 23, 2003 in honor of Marvin Zelen, during which time Dr. Zelen was awarded an honorary doctoral degree by University Victor Segalen Bordeaux 2.*

First, let me summarize some facts regarding his education and employment. Marvin graduated from City College (in New York) in 1949, but his college career was interrupted when he proudly served in the United States Army immediately following World War II. Following his undergraduate studies, Marvin obtained a Master's degree from the University of North Carolina in 1951. Marvin's first professional position was at the Experimental Towing Tank at Steven's Institute of Technology, from 1951-1952. From 1952-1961, he was a member of the Statistical Engineering Laboratory of the Division of Mathematics at the National Bureau of Standards (NBS), where his real research career really began. Here he worked with and was supervised by two famous statisticians—W.J. Youden and Churchill Eisenhart. Marvin's research career blossomed at the NBS, where he worked on problems in experimental design and reliability.



After his tenure at the National Bureau of Standards, Marvin briefly held positions at the University of Maryland, the Army's Mathematics Research Center in Madison, Wisconsin, the University of California at Berkeley, and at Imperial College, where he was a Fulbright Fellow. From 1963-1967, Marvin was a statistician at the National Cancer Institute, and this is where his career shifted to a focus in biostatistics.

In 1967, after President Nixon declared the "War on Cancer", Marvin joined the faculty at SUNY/Buffalo, where he formed the "Statistical Laboratory", becoming involved in many cancer clinical trials through several cancer cooperative groups such as the Eastern Cooperative Oncology Group, the Veteran's Administration Lung Group, the Radiation Therapy Oncology Group, and the Gastrointestinal Tumor Study Group. Today, the housing of statistical coordinating centers

in academic settings is a common phenomenon. Then it was very novel.

In 1977, he left SUNY/Buffalo to join Harvard's Biostatistics Department and the DFCI. At the DFCI, he formed and headed what is now the Department of Biostatistical Science. Marvin headed these two groups for more than a decade, and remains actively involved to this day.

During his career, Marvin has made many important contributions to the statistical literature and as the statistical collaborator in multidisciplinary groups doing applied research, primarily in clinical trials, where he interacted with clinical scientists. His early methodological work was not in biostatistics, and includes numerous contributions to the fields of experimental design and reliability theory. Since the 1960's, however, the work has been in biostatistics. His areas of interest are diverse, and major areas include models for the natural history of disease, models for screening for early detection of disease, and a variety of issues in the design and analysis of clinical trials. The contributions are numerous—and too much to discuss in any detail now—but including fundamental and innovative contributions in length-biased sampling, lead-time bias, the allocation of patients to clinical trials, and stochastic processes. His work is frequently cited and has had an enormous impact. It is noteworthy that much of his statistical work was implemented in the clinical trials he was involved in, sometimes simultaneously or even before the methodological papers were published. This reflects his close involvement in the 'science' behind the applications, which were the motivation for much of his methodological work.

Although much of Marvin's work since the 1960's has been in clinical trials and screening studies, he also had ventured into other fields. He and I became involved in a health study in the 1980s concerning the health effects of contaminated water in a community called Woburn near Boston. This involvement began following a seminar given at Harvard by a minister—Bruce Young--who was counseling several the families that had recently had children diagnosed with leukemia. The minister's talk described the community's concerns and their dissatisfaction with the help they had received from the government agencies that had noted the elevated leukemia rates and contaminated water in the area but were unable to assess whether there was any connection. Subsequently, we worked very closely with a community-based organization in the town to help in the design and conduct of a health study. The scientific challenges in this area were enormous, as were the politics, yet our close involvement with the community was a unique and enormously gratifying experience for both of us, and reflect Marvin's lifelong commitment to help others.

Marvin's energy level is remarkable. For several years during his early years at Harvard, Marvin simultaneously held three positions: heading the DFCI Dept of Biostatistical Science, Group Statistician for the Eastern Cooperative Oncology Group, and Chair of the Department of Biostatistics at Harvard School of Public Health. Today, each of these positions is headed by a different person: Dave Harrington, Bob Gray, and LJ Wei, and each would tell you that they almost always are overwhelmed by the responsibilities of their jobs. It's amazing that Marvin handled all three simultaneously.

Beyond his research, Marvin, as much as anyone I have known, has made incredible contributions to the promotion of biostatistics, though his mentoring and training of younger persons and through his promotion of the field organizationally. The number of people he has mentored or advised is very large, and includes his doctoral students, junior faculty, postdocs, and persons around the world who sought his advice, or in some cases heard him give a lecture.

- Just focusing on faculty, the list is quite long, and from just his time at SUNY/Buffalo includes: Ross Prentice, Jack Kalbfleisch, Stuart Pocock, Jim Hanley, Steve Lagakos, Rich Gelber, Ken Stanley, Colin Begg, and David Schoenfeld.

- Marvin has also mentored others, including Catherine Hill, Yu Shen and Ping Hu, who are in attendance at this conference.

- I have also heard other persons indicate that Marvin was an important influence in their decision to become biostatisticians. This list includes Cyrus Mehta, Brent James, John Simes, Richard Landis, and Giovanni Parmigiani.

Marvin, through his activities as a statistical scientist, also has had a big influence on non-statisticians with whom he has collaborated. Two were very eminent physicians—Tom Frei and Paul Carbone. Tom Frei and Marvin were together at the National Cancer Institute, and Tom subsequently helped to recruit Marvin to Harvard and the DFCI. Paul Carbone headed the Eastern Cooperative Oncology Group for many years, and together with Marvin implemented many modern and novel methodologies for clinical trials into their studies. Both Tom and Paul benefited enormously from their interactions with Marvin, and give him enormous credit for some of their successes.

Marvin also has facilitated the development of young biostatistical scientists through his promotion of the discipline. He was very instrumental in the development of the data center for the Eastern Organization for Research on the Treatment of Cancer, and sent several of us (then) young biostatisticians to Brussels to help them during their early years. More recently, Marvin has been key in the formation and conduct of a summer institute in biostatistics in Taiwan.

In summary, Marvin has done a great deal for our field through his research, his leadership in fostering biostatistical science, and his mentoring, advising, and encouragement of many young members of our profession.

In most things we do professionally, it is hard to know the eventual impact—if any—on future generations or the field. Today, biostatistics is a thriving profession and the importance of the discipline in biomedical research is widely recognized. It was not always this way, and I can recall early debates with surgeons and clinicians on such fundamental concepts as the value of randomization. Bringing 'science' into medical research has been cited as one of the top 10 contributions to medicine in the last millennium. No doubt many people have contributed to this, and as I noted above, it is hard to judge the contributions of any one person. What is clear, however, is that Marvin Zelen—through his research, his promotion of our discipline, and his mentoring of younger generations—has played an important role in fostering an environment for the healthy advance of biomedical research.

Stephen Lagakos, Ph.D. *is the Chairman of the Biostatistics Department, Harvard School of Public Health, Boston, MA, USA.*
Email: lagakos@hsph.harvard.edu

RISK MANAGEMENT VIA VALUE AT RISK

Kao-Tai Tsai
Aventis Pharmaceuticals, Bridgewater, New Jersey, USA

SUMMARY

Risk management is an integral part of the financial investment activities to gird investment decisions and manage risks of an entire institution. Value at Risk has become an indispensable tool for this purpose. The objective of a VaR framework is to provide a flexible and unbiased assessment of the risks of the portfolio over time so that the senior management will have the ability to analyze all business risks, to communicate the risks of the portfolio, and to make strategic risk-based decisions. This article describes a statistical framework for the derivation of the VaR via the relationship of the risk factors of a portfolio with the pricing functions and the correlations among the components of the portfolio. The methodologies of estimations, the stress test, model backtesting, and the impact of position changes to the value of the portfolio are also discussed.

1. Introduction

Risk management is an integral part of the financial investment activities. It covers a broad scope of frameworks and methodologies to define, measure, and evaluate all types of risks in order to allocate capital to cover various risks for a mix of portfolios. To put it in the most general term, risk management is a philosophy to gird investment decisions or a process to manage risks of an entire institution. One of the popular tools of risk management in recent years is Value at Risk (VaR), which is defined as the loss that will be expected with a given probability over the period of time during which a portfolio is held.

A primary thrust for its popularity is the 1996 amendments to the Basle Capital Accord to incorporate market risk into the regulatory capital framework. It permits supervisory agencies to allow banks to use their internal VaR models to calculate their regulatory capital requirements as an alternative to the conventional methodology. Another reason that VaR was embraced by bankers and their regulators is that VaR offers a measure of a firm's overall exposure to market risk. And it is indeed useful for understanding how much a given instrument or portfolio might lose over a period of time. The systematic use of the VaR for different portfolios has had a positive effect on the handling of risk within financial institutions due to the increasing degree of complexity investment instruments have had assumed over the last few years.

However, the popularity of the VaR also carries its criticism. From a pure data analytical point of view, to reduce all the information down to a single number means the loss of important information. That may in turn create misleading interpretations of the analytical results. Furthermore, VaR is of limited use at the strategic level because it is difficult to allow meaningful comparisons across various financial markets by reducing everything to a

single number. To measure performance across markets, benchmarks are typically needed to reflect the peculiarity of the various markets. If a firm chooses a benchmark trade with an above average expected return, then a trader must assume more risk for that market if the investment is expected to perform in line. However, a single number of the VaR does not usually facilitate that kind of comparison.

The objective of a VaR framework is to provide an unbiased assessment of the portfolio risks over time. To achieve this objective, the framework chosen must be flexible and general enough to accommodate changes to the corporation's portfolio due to the changes in the associated risk factors. A great effort has been devotèd to the discussion of VaR under various scenarios and different sampling scheme as well as the distribution assumptions (e.g., Hendricks, 1996; MeVay & Turner, 1995; Fong & Vasicek, 1997; Lo, 1999). However, a systematic quantitative framework is needed to fully understand the merits and drawbacks of the derivation of the VaR.

2. General Framework of VaR Estimation

Let R_i be the return of the ith asset class of the entire portfolio during a specific time interval, also assume the risk factors associated with the entire portfolio be $F_1, F_2, \cdots, F_k$. Since the return of an asset class is often caused by the change of the underlying risk factors, one can write the return of the ith asset class as

$$R_i = \sum_{j=1}^{k} a_{ij} f_{ij}(\Delta F_j) + \epsilon_i = A_i' \mathcal{F}_i + \epsilon_i \tag{1}$$

where $A_i = (a_{i1}, a_{i2}, \cdots, a_{ik})'$ is a vector of constant coefficients for the ith asset class and $\mathcal{F}_i = (f_{i1}(\Delta F_1), f_{i2}(\Delta F_2), \cdots, f_{ik}(\Delta F_k))'$ is a random vector of functions, which relate the return with the change of the risk factors. These functions can be either smooth or discrete. They can even be a set of differential equations, which relate the asset pricing to the underlying parameters. The random error ϵ_i with mean 0 accounts for the information not captured by the risk factors. A simple example of equation (1) is the CAPM, which relates the return of a security with the market return, namely,

$$R_s = \alpha + \beta(R_m - R_f) + \epsilon_s = a_{s1} f_{s1}(\Delta F_1) + \epsilon_s$$

with $a_{s1} = \alpha$ and $f_{s1} = (\beta/\alpha)(R_m - R_f)$. A more complicated example of equation (1) is the return of options such as the Black-Scholes equation or the Taylor series expansion relating the return of an option to the change of gamma, theta, and vega, etc. Following equation (1), the mean of R_i is $E(R_i) = A_i' E(\mathcal{F}_i)$ and the variance of R_i is $\mathrm{Var}(R_i) = A_i' \Sigma_{\mathcal{F}_i} A_i + \sigma_\epsilon^2$, where $\Sigma_{\mathcal{F}_i}$ and σ_ϵ^2 are the covariance matrices of $\mathcal{F}_i$ and ϵ_i, respectively. Note that the components a_{ij} of the coefficient vector A_i will be zero if the ith asset class is not exposed to the jth risk factor.

If a portfolio consists of γ asset classes, then the return of the portfolio can be written

as

$$R_P = \sum_{i=1}^{\gamma} w_i R_i = W'R. \tag{2}$$

Combing equations (1) and (2), the return of the entire portfolio can be expressed as

$$R_P = \sum_{i=1}^{\gamma} w_i \Big(\sum_{j=1}^{k} a_{ij} f_{ij}(\Delta F_j) + \epsilon_i \Big) = \sum_{i=1}^{\gamma} \sum_{j=1}^{k} w_i a_{ij} f_{ij}(\Delta F_j) + \Upsilon, \tag{3}$$

where Υ is the aggregated random error term from all the asset classes of the portfolio.

An estimate of the VaR at the $\alpha\%$ level for the ith asset class of the portfolio is defined as

$$\text{VaR}_i = G_i^{-1}(\alpha\%) \times V_i \tag{4}$$

where G_i^{-1} is the quantile function of R_i and V_i is the market value of the ith asset class. Generally, G_i^{-1} is unknown and needs to be estimated from the existing data. If R_i follows a normal distribution, then

$$\text{VaR}_i = \Phi^{-1}((x - E(R_i))/\sigma_{R_i})(\alpha\%) \times V_i. \tag{5}$$

Similarly, an estimate of the VaR at the $\alpha\%$ level for the entire portfolio can be defined as

$$\text{VaR}_P = G_P^{-1}(\alpha\%) \times V_P, \tag{6}$$

where G_P^{-1} is the quantile function of R_P and V_P is the market value of the entire portfolio. Generally, the VaR of the whole portfolio is not necessary equal to the summation of the VaRs of the asset class components because the inter-correlation among all the asset classes of the portfolio.

3. General Issues of VaR Estimation

The estimate of an unknown parameter depends on the quality of the data and the methodology utilized in the estimation process. This also applies to the estimation of the VaR. A VaR estimate is useful only if the data used to derive the estimate appropriately reflects the market conditions under consideration. However, since forecasting the future condition is usually not straightforward, risk managers need to be flexible in approximating the targeted market condition with appropriate information in order to make the VaR estimate useful. In the following, we describe a few important issues in estimating VaR.

3.1 Data Collection and Data Processing

The VaR estimate depends on the quality and timeliness of input data. An important issue is the frequency and the duration of data. High frequency data is not necessary better than the low frequency data due to the noise it may be associated with and the cost of

storage. On the other hand, low frequency data has the possibility of missing some critical events. One example is the daily trading data verses the weekly data. Since the VaR estimate assumes that the portfolio's composition does not change over the holding period, the solution of this issue very much depends on the investment strategies.

A related issue in data collection is the time span of data coverage. Short duration data may not cover some important market cycles. On the other hand, long duration data has the potential of including the market cycles which are not of interest and therefore contaminate the primary focus. Since the return distribution is most important for the historical simulation and Monte Carlo simulation approaches, the validity of these approaches depends on the extent of sample selection by choosing an appropriate period and frequency of the historical data.

However, data collection is just the beginning of the modeling process. Special attentions are needed to ensure the high quality of the data collected. A simple visual inspection of the data is often insufficient, especially for large volumes of multivariate data in high dimension. Therefore, in addition to descriptive statistics, graphical methods and exploratory data analysis techniques should be utilized to better understand the data with large volume.

3.2 Estimation of Quantiles of Distribution

Several methodologies have been proposed to estimate the quantiles of the return distribution of individual asset class or the entire portfolio. Since the return distribution does not always follow the normal distribution, the approximated quantile using a multiple of the standard deviation of the normal distribution usually produces substantial bias. The bias can become much worse if the distribution is highly skew or with substantial fat tails.

Alternative methods have been proposed to reduce the potential bias of the estimation. For example, one can estimate the pth quantile (a) by the $([np]+1)$th smallest data point of the return distribution, where $[np]$ is the largest integer less than np; (b) by the weighted sum of the $[np]$th and the $([np]+1)$th smallest data points of the return distribution; or (c) with all the data weighted by incomplete beta function (Harrell and Davis, 1982). Even though these methods can sometimes result in biased estimates, generally speaking, method (c) seems to be a better choice in estimating the required quantile even though it has a higher degree of complexity in computations.

Since the estimated quantile is a random variable associated with estimation error, the precision of the estimate, by the construction of the $100(1-\alpha)\%$ confidence interval, can be quite informative. A similar confidence interval of VaR can also be provided to the risk managers for better strategic decision makings.

3.3 Issues of Covariance Matrix Estimation

When the portfolio consists of more than one asset class and the return distributions are

assumed to be inter-correlated, the covariance matrix of the returns of these asset classes is needed before estimating the VaR. As in the estimation of quantiles of the return distributions, the appropriateness of the input data and the methodology utilized in the estimation need to be considered carefully in order to yield useful results. Many approaches to estimate the covariance matrix are available in the literature (e.g., Alexander & Leigh, 1997; Boudoukh, Richardson, & Whitelaw, 1997). The classical least squares estimate of the covariance matrix is efficient if the return distributions are multivariate normal. However, it is well known, the return distributions of the market data often consist of outliers or some degree of deviation from normality, which make the least squares estimation less optimal.

Alternatives have been proposed to improve the problematic scenarios by using nonparametric method (e.g., Hull & White, 1998), semiparametric method (Fan & Gu, 2003), or robust estimate in order to reduce the sensitivity of outliers. However, even though less sensitive to outliers, nonparametric or robust approaches could eliminate the rare but important information of the market activities. Therefore, proper compromises by the risk managers between these two classes of approaches are often needed to provide a sensible estimate of the covariance matrix for the estimation of the VaR.

Furthermore, the stability of the estimated covariance matrix also needs to be considered. A simple simulation can be conducted to reveal the degree of uncertainty of the estimated covariance matrix. Combing the variation of the estimated covariance matrix with the variation of the estimated quantile, one can easily have a wide range of possible values for the VaR. Some of them may even be different significantly from each other. This phenomenon is similar to the portfolio optimization as both cases involve the estimated covariance matrix of the return distributions of the portfolio components (e.g., Jorion, 1992; Michaud, 1997). The conventional Delta-Normal or Delta-Gamma method (e.g., Wilson, 1996; RiskMetrics, 1996) which heavily depends on the multivariate normality can produce substantial bias if the distributional assumption is violated significantly.

4. Components of VaR

VaR is an aggregated summary of the risk for the entire portfolio. It is critical to understand the risk exposure of each component of the portfolio in order to achieve effective hedging and reach the total risk management. Garman (1996) proposed the concept of VaRdelta (∇VaR), which is defined as the gradient of the VaR. Even though straightforward, ∇VaR can be used as a tool to assess the componentwise impact on the risk of the portfolio.

One should notice that the estimation of ∇VaR is valid only for a small amount of change in the value of the components and may be misleading if the change is substantial in magnitude due to the nonlinear nature of the VaR estimate. One can extend the results to take into account the functional non-linearity by a higher order approximation. Furthermore, the variation of ∇VaR can also be estimated by simulation as the variation is difficult to be estimated from explicit analytical calculations.

5. Stress Test and General Issues

The general practice of stress test is to subjectively assume the occurrence of some unusual events related to the risk factors and to estimate the impact of these events to the value of the portfolio. One example could be the impact of a substantial change in interest rate. Another example could be the impact of the sudden decrease of the oil supply. Since the events are unusual and sometimes extreme, the related data is either not existing or very rarely collected. The estimates of the VaR or other risk measures could become very difficult or unreliable. To proceed with the VaR estimation based on the information of normal market condition except for the stressed risk factors is usually insufficient due to the fact that most of the risk factors are correlated. To stress one risk factor without simultaneous considering other factors can cause an erroneous estimate of the VaR. Therefore, it is essential to consider how the changes of the risk factors can affect other risk factors.

Other complications, such as the distortion of the pricing relationship between the investment instruments and the risk factors in the normal market conditions, can occur under the stressed market conditions. Specifically, the relationship as shown in equation (1) under the normal market condition could possibly become

$$R_i = \sum_{j=1}^{k} a_{ij}^* f_{ij}^*(\Delta F_j) + \epsilon_i^* \tag{7}$$

for a new set of functions f_{ij}^* and coefficients a_{ij}^*, and the relationship as shown in equation (3) would consequently become

$$R_P = \sum_{i=1}^{\gamma} \sum_{j=1}^{k} w_i a_{ij}^* f_{ij}^*(\Delta F_j) + \Upsilon^*, \tag{8}$$

where f_{ij}^* and a_{ij}^* may be different from f_{ij} and a_{ij}.

The expected impact to the portfolio return due to the change of the jth risk factor under the stressed conditions can then be estimated by the following equation:

$$E\Big(\frac{\partial R_P}{\partial F_j}\Big) = \sum_{i=1}^{\gamma} \sum_{j=1}^{k} w_i a_{ij}^* E\Big(f_{ij}^{*\prime}(\Delta F_j)\Big), \tag{9}$$

and the impact to the VaR can be estimated accordingly even though it could be a tedious process. Mathematical and statistical modeling are needed to derive the new functional relationship of equation (8) between the asset classes and the risk factors for the purposes of stress tests.

Even though the extreme conditions are usually subjective with a small probability of occurring, an estimation of the probability of these conditions actually occurring is highly desirable. Methodologies similar to the low dose extrapolation in toxicology or procedures to estimate the probability of extreme events can be very useful for the estimation of the

extreme probability. With the estimated probability of the stressed events and the estimated VaR under this condition, one can estimate the composite confidence level of the VaR under the stressed market condition.

6. Backtest & Model Validation

Most of the modeling, estimation, and prediction processes are based on the historical data. It is crucial to know how well the results based on the derived model confirm with the empirical results in the future. In the context of risk management, this implies the institutional effort to check whether the model has an accurate, complete, and consistent description of the risks in the portfolios; whether the VaR methodology aggregates and measures the risks of the portfolios appropriately when compared with the actual P&L of the portfolios; and whether the allocation of risk capital can appropriately hedge the risks predicted by the models and simultaneously maximize the returns predicted by other related models. This is important not only for the internal institutional risk management but also for the compliance of regulatory requirements. We discuss a few related issues in the following.

For the validation of the VaR methodology, the distributional assumptions need to be examined first. The quantile-quantile plot of the distributions with the confidence band of the plot can provide useful detection of any deviations. Any significant deviations would indicate an inconsistent distributional assumption. Special attention should be paid at the tails of the distributions. A simple comparison of means and standard deviations of the distributions is generally insufficient for the detection of distributional differences.

Secondly, the consistency of the covariance matrices needs to be investigated. Statistical methods can be used to perform the test of equality of covariance matrices (e.g., Anderson, 1984). Alternatively, one can test the equality of the Mahalanobis distance of covariance matrices.

Functional relationship between the returns of an asset class and the risk factors as described in equation (1) needs to be examined too. Consistent over-estimation or under-estimation with the updated data could indicate the under-parameterization or misspecification of the underlying model. Therefore, interactions between the modeling effort and the risk managers to understand the subtlety of the market behavior and the effects of the risk factors is essential in order to make the model useful for the risk management purpose.

7. Other Measures of VaR

VaR as defined above only consider the results at the end of the holding period; however, it ignores what might happen during the holding period. Therefore, the concept of "Continuous VaR" has been proposed by some researchers (e.g., Kritzman & Rich, 2002). The interim risk is particularly important for investors with some predefined thresholds, which cannot

be breached if the investment is to survive to the end of the holding period. Examples can easily be found in, for instance, asset management, hedge fund solvency, loan agreement, or securities lending, etc. The calculation of Continuous VaR involves the reflection principle of stochastic processes and usually is larger than the traditional VaR. Consequently, this may result in the requirement of a larger amount of reserve for the institutions to cushion the unusual losses.

Another shortcoming of VaR is that it provides no information on the extent of the losses that might be suffered beyond the threshold amount indicated by the measure of VaR. During dramatic market conditions, portfolios break same measure of VaR may incur very different losses due to the nature of the holdings. Therefore, alternative measures such as "Conditional VaR" has been proposed for this purpose (Longin, 2001; Rockafellar & Uryasev, 2002). The Conditional VaR of the loss at the α-level is the conditional mean of the α-tail distribution of the portfolio's returns. As a tool in optimization modeling, Conditional VaR maintains consistency with the regular VaR in the settings where VaR computations are tractable. Furthermore, Conditional VaR can be expressed in minimization formula and be incorporated into the optimization problems.

8. Conclusion

Risk is a multi-dimensional event and no single algorithm or measurement will serve the purpose of including all possible dimensions of risk in the market place. Financial institutes have established independent risk management units because the boards of those firms want to understand and control the risk taken. During most market conditions, risk management can provide measurement and analysis that aids the understanding of financial risks taken. Risk management is more than just a regulatory reporting exercise or a defensive activity, if implemented effectively, it can also increase business profitability by focusing capital and attention on the areas with higher risk/return ratio.

Effective risk management function requires commitment from the highest levels of the firm to develop a proactive risk culture by setting the appropriate policies and guidelines. Many methods and tools are fundamental to support and analyze risk, yet, they alone do not assure a firm's ability to effectively manage risk. Even though highly popular, the VaR is just one tool for communicating the risks and should be incorporated with sound risk management judgment and appropriate risk controls. Senior management should understand the limitations of the VaR such as those discussed in the previous sections and complement with other risk management techniques such as sensitivity analysis, stress test, and scenario test to develop a portfolio approach for risk measurement. It is important to know that portfolios with the same VaR do not necessary have equal risk when the market makes its next turn. The market can often defy the model no matter how advanced the model was constructed. Having a VaR system in place does not preclude the possibility of greater losses than expected.

Notes

1. The concept of ∇VaR is similar to duration on the value of a bond when the bond yield is changed. Specifically, ∇VaR is defined as

$$\nabla \text{VaR} = \frac{\partial}{\partial \mathbf{P}} \text{VaR} = \frac{\partial}{\partial \mathbf{P}} \sqrt{\mathbf{P}' \Sigma_{\mathbf{P}} \mathbf{P}} = \frac{\Sigma_{\mathbf{P}} \mathbf{P}}{\sqrt{\mathbf{P}' \Sigma_{\mathbf{P}} \mathbf{P}}} = \frac{\Sigma_{\mathbf{P}} \mathbf{P}}{\text{VaR}},$$

where $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_\gamma)'$ is a vector of portfolio components and $\Sigma_{\mathbf{P}}$ is the covariance of the components of the portfolio. It is easy to notice from above equation the following relationship between the VaR of a portfolio and ∇VaR

$$\text{VaR} = \mathbf{P}' \ \nabla \text{VaR} = \sum_{i=1}^{\gamma} \mathbf{p}_i \nabla \text{VaR}_i,$$

where $\mathbf{p}_i$ and ∇VaR_i are the ith component of the portfolio and ∇VaR, respectively.

2. The following equation extends the first order result to the second order expansion:

$$\frac{\partial}{\partial \mathbf{p}_i} \nabla \text{VaR} = \Sigma_{\mathbf{P}} \mathbf{J}_i' (\mathbf{P}' \Sigma_{\mathbf{P}} \mathbf{P})^{-1/2} + \Sigma_{\mathbf{P}} (\Sigma_{\mathbf{P}} \mathbf{P})_i \times (\mathbf{P}' \Sigma_{\mathbf{P}} \mathbf{P})^{-3/2},$$

where $\mathbf{J}_i$ is a vector of zeros with 1 in the ith place and $(\Sigma_{\mathbf{P}} \mathbf{P})_i$ is the ith component of the matrix $\Sigma_{\mathbf{P}} \mathbf{P}$. Therefore, when the position of the ith component in a portfolio is changed from the original position $\mathbf{P}_0$ by a small amount ϵ_i, the change of the VaR can be expressed by the following relationship:

$$\frac{\partial}{\partial \mathbf{p}_i} \text{VaR}(\mathbf{P}_0 + \epsilon_i) = \text{VaR}(\mathbf{P}_0) + \frac{(\Sigma_{\mathbf{P}} \mathbf{P})_i}{\text{VaR}(\mathbf{P}_0)} \times \epsilon_i + \frac{1}{2} \Sigma_{\mathbf{P}} \mathbf{P}_0 (\Sigma_{\mathbf{P}} \mathbf{P}_0)_i \times (\mathbf{P}_0' \Sigma_{\mathbf{P}} \mathbf{P}_0)^{-3/2} \times \epsilon_i^2 + \text{higher order terms},$$

where $\text{VaR}(\mathbf{P}_0)$ is the original VaR before any change of the position.

3. For the portfolio $\mathbf{R}_{\mathbf{P}}$, the Mahalanobis distance can be calculated by

$$\left\{ (R_1, R_2, \cdots, R_\gamma)' (W' \Gamma \Sigma_{\vec{x}} \Gamma' W + W' \Sigma_\Theta W)^{-1} (R_1, R_2, \cdots, R_\gamma) \right\}^{1/2}.$$

This quantity has a chi-square distribution with γ degrees of freedom if the joint distribution of the returns is multivariate normal.

References

[1] Alexander, C.O. & Leigh, C.T. (1997). "On the covariance matrices used in value at risk models," The Journal of Derivatives, 50-62.

[2] Anderson, T.W. (1984). *An introduction to multivariate statistical analysis*, 2nd edition, John Wiley & Sons, New York.

[3] Boudoukh, J., Richardson, M., & Whitelaw, R. (1997). "Investigation of a class of volatility estimators," The Journal of Derivatives, 63-71.

[4] Box, G.E.P. & Jenkins, G.M. (1976). *Time Series Analysis: forecasting and control.* Holden-Day, Oakland, California.

[5] Fan, J. & Gu, J. (2003). "Semiparametric estimate of Value at Risk," Econometrics Journal, v.6, 261-290.

[6] Fong, G. & Vasicek, O. (1997). "A multidimensional framework for risk analysis," Financial Analysts Journal, July/August, 51-57.

[7] Garman, M. (1996). "Improving on VaR," Risk, May 1996, vol.9, no.5, pp.61-63.

[8] Harrell, F & Davis, C. (1982). "A new distribution-free quantile estimator," Biometrika, 69, 635-40.

[9] Hendricks, D. (1996). "Evaluation of value-at-risk models using historical data," FRBNY Economic Policy Review, Apr., 39-69.

[10] Hull, J. & White, A. (1998). "Value at risk when daily changes in market variables are not normally distributed," Journal of Derivatives, 9-19.

[11] Jorion, P. (1992). "Portfolio optimization in practice,' Financial Analysts Journal, 48(1), p.68-74.

[12] MeVay, J. & Turner, C. (1995). "Could companies use value-at-risk ?" Euromoney, Oct., 84-86.

[13] Lo, A. (1999). "The three P's of total risk management," Financial Analysts Journal, 13-26.

[14] Michaud, R. (1997). "Efficient asset management,' Harvard Business School Press.

[15] RiskMetrics^{TM} - Technical Document (1996), J.P. Morgan, New York.

[16] Wilson, T. (1996). "Calculating risk capital," in *The handbook of risk management and analysis*, ed. C. Alexander, Wiley & Sons, New York.

[17] Kritzman, M. and Rich, D. (2002). "The mismeasurement of risk," Financial Analysts Journal, v.58, p.91-99.

[18] Longin, F.M. (2001). "Beyond the VaR," The Journal of Derivatives, p.36-48.

[19] Rockafellar, R.T. and Uryasev, S. (2002). "Conditional value-at-risk for general loss distributions," Journal of Banking & Finance, v.26, p.1443-1471.

Controversial Statistical Issues: Surrogate Endpoint, Biomarker, & Imaging Test

Special Topic Editor: Sue-Jane Wang, Ph.D.

The Use of Receiver-Operating Characteristic Curve to Identify Biomarkers: Have They Been Properly Validated?

By: Sue-Jane Wang, Ph.D.

Recent developments in microarrays and protein arrays offer new approaches to identify gene expression signatures, proteomics or DNA-based biomarker panels that could potentially serve various purposes in drug discovery and development, clinical trials and therapeutic assessment strategies. In view of a needed consensus in the use of terminology that facilitates the interaction to advance science and policy, the Biomarkers Definitions Working Group of NIH (2001) proposed "the biological marker (biomarker) is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention", also see Baker and Kramer; Feng; Johann, Liotta and Petricoin for definitions.

Surrogate endpoints are considered a subset of biomarkers. Defined by the same group, a surrogate endpoint is a biomarker that is intended to substitute for a clinical endpoint. It is expected to predict clinical outcome(s) based on epidemiologic therapeutic, pathophysiologic or other scientific evidences requiring shorter study duration than the clinical endpoint studies. The controversy in the statistical evaluation of surrogate endpoints is addressed by Molenberghs, Buyse and Burzykowski. Imaging test is sometimes considered an early detection biomarker to trigger early intervention involving a subjective nature in reading the image (see Baker & Kramer.)

Researchers have proposed new or sophisticated statistical prediction algorithms to study genomics and proteomics data for prognostic or diagnostic classification. In class prediction, concerns over the unsupervised cluster analysis have been raised. Not only does the cluster analysis make no use of sample grouping, it does not provide statistically valid measure on which genes are altered in different known classes of tissue. Various ranking procedures by way of parametric or rank statistics are also advocated to explore potential molecular biomarker discovery. The identified biomarker panels, analogous to identification of differentially expressed genes, are often reliant on the particular statistical procedure (supervised or unsupervised) adopted, the number of genes (or probes) studied, the number of disease and control tissue samples studied and the threshold value used to make the individual gene or multiple gene sets significant.

The Receiver-Operating Characteristic (ROC) plot is a fundamental evaluation tool in clinical medicine. Given the gold standard and a new test that classifies a subject based on laboratory or clinical continuous measure, the ROC curve measures a test's ability to discriminate between two subclasses of subjects, e.g., non-diseased vs. diseased, responder vs. non-responder to therapy. The concept of the diagnostic accuracy in the use of ROC plots has increasingly been exploited for identification of genomic or proteomic biomarkers gene-by-gene, protein-by-protein, or simultaneous assessment of tens of thousands of genes or proteins. These potential molecular biomarkers are mostly used for early detection of cancer or for cancer stage prognostication. The ROC curve can be viewed as a plot of true positive rate versus false positive rate associated with all possible thresholds for classifying a tissue as poor or good signatures based on the relative expression or protein levels with an integrated ranking score. Intuitively, a highly predictive biomarker possesses high sensitivity. To minimize unnecessarily costly medical intervention following disease state detection, it is desirable to have high specificity. In cancer screening, this translates into a very small false positive rate due to its cost and consequence of errors.

To gain insight into the statistical vigor in the assessment of predictive performance of gene expression signature for diagnostic or prognostic classification, a discovery-based biomarker test, Simon (2003) pointed out the limitation of small studies to accurately build a predictor set that have high prediction accuracy. He also showed via simulation studies the importance of cross validating all steps of predictor selection in estimating a true error rate when an independent validation is not feasible. It goes without saying that a large independent dataset is preferred for validation purpose. It is to confirm the robustness of the predictive power in a similarly studied population. The above critiques were later demonstrated via an empirical assessment of 30 microarray studies that contain major clinical outcomes (death, metastasis, recurrence, response to therapy), which were selected via MEDLINE search between 1995 to 2003 (Ntzani and Ioannidis). Below is an example study used in Ntzani and Ioannidis that performed a cross validation and two independent validations to illustrate the controversial nature of 'validation.' The pertinent questions are "Is the predictive power acceptable?" "Is the false positive rate too high?" "Is the study reasonably validated?" and "Is the validation unbiased?"

Approach	n	Sensitivity	FPR*
Cross-Validation	78	91%	27%
Independent V.	19	100%	29%
Independent V.	180	93%	47%

* False Positive Rate

Thus, convincing reports of achieving more than 90% or close to 100% sensitivity might be the consequence of an improper cross validation resulted in serious underestimation of error rate. Alternatively, the inaccurate or misleading reports could be at the expense of independent validation but on small studies. It is critically important that the potential valid biomarker be independently validated or completely cross-validated. A cautionary note: it is not rare to observe high sensitivity in small studies as also illustrated by Ntzani and Ioannidis. From the existing literature, it appears that seldom have high predictive power results being conclusively shown to hold in larger samples. A possible speculation may be that of overfitting the predictor in the sense that the number of potentially correlated parameters of the model is too big relative to the number of tissues studied. Until statistically and biologically validated and further confirmed in randomized clinical trials, many of the early preclinical or clinical phase retrospective biomarker studies should be viewed as plausible exploratory exercise.

Acknowledgement: The author thanks Dr. Greg Campbell for insightful discussion on ROC curve.

Sue-Jane Wang*, The Statistics Representative, FDA Inter-Center Pharmacogenomics / Pharmacogenetics Working Group, U.S. FDA.

*The views presented in this article do not represent those of the U.S. Food and Drug Administration. Email: WANGS@cder.fda.gov

Biomarker, Surrogate Endpoints, and Early Detection Imaging Tests: Reducing Confusion

Stuart G. Baker, Sc.D.
Barnett S. Kramer, M. D., M.P.H.

Sometimes discussions of surrogate endpoints, biomarkers, and imaging tests start off "on the wrong foot" because different investigators use different definitions of these terms. Therefore we begin this article by presenting our definitions. The most general term of the three is biomarker, which we define as a measure or indicator of a biological process. Some biomarkers are potential

indicators of a health outcome. If a biomarker for the true endpoint of the study is collected earlier than the true endpoint (health outcome) and is intended to provide information about the true endpoint, we call it a potential surrogate endpoint. (Here the word potential refers to the usual usage not a counterfactual variable, as mentioned later). If the potential surrogate endpoint provides valid inference about the true endpoint, we call it a surrogate endpoint. We call a biomarker for predicting later clinically apparent or symptomatic disease, an early-detection biomarker. (Because of our work in cancer prevention, the focus in this article is early detection of cancer). Early detection biomarkers can either be molecular markers, such as PSA for the early detection of prostate cancer, or imaging tests, such as mammography for breast cancer. Another common use of biomarkers in cancer prevention (not covered here) is to identify subjects who might be most likely to benefit from an intervention. The statistical issues associated with each type of biomarker differ and depend on the purpose of the biomarker.

1. Potential surrogate endpoints

The purpose of a potential surrogate endpoint in a randomized trial is to obtain correct inference about the effect of an intervention on a true endpoint without waiting to observe the true endpoint. Potential surrogate endpoints either replace or predict the true endpoint. In the replacement mode, inference focuses on hypothesis testing. In a landmark paper, Prentice (1989) formulated criteria for valid hypothesis testing based on a potential surrogate endpoint. The key criterion, later called the Prentice Criterion, is that the distribution of true endpoint conditional on potential surrogate endpoint does not depend on randomization group. In other words the association between the potential surrogate endpoint and the true endpoint is the same, irrespective of the intervention. Given the complexity and multiple pathways of pathogenesis for most chronic diseases such as cancer, this criterion is often difficult to meet. A simple model can test if the Prentice Criterion can be rejected (Freedman et al, 1998). The controversy is what to report when the Prentice Criterion cannot be rejected. Although various summary measures have been proposed (e.g. Freedman et al 1998, Molenberghs et al, 2002), we do not find them too appealing, as it is difficult to specify levels that indicate acceptable (but not perfect) potential surrogate endpoints.

When a potential surrogate endpoint is used to predict the true endpoint, the basic idea is to (1) estimate parameters modeling the relationship between potential surrogate and true endpoints in previous randomized trials, and (2) predict true endpoint using the estimated parameters and the surrogate endpoint in a new randomized trial The models involve either individual level data with observed potential surrogate endpoints (e.g. Day and Duffy, 1996), counterfactual potential surrogate endpoints (Frangakis and Rubin, 2002), or trial-level summary measures of potential surrogate endpoints (e.g. Gail et al, 2000, Molenberghs et al 2002). All the models rest on the critical assumption that parameters estimated from the previous trials are appropriate for the new trial.

There is some confusion in the use of these models, which we briefly mention. First, with individual-level data models for observed potential surrogate endpoints, it is sometimes thought that the Prentice Criterion is necessary for inference, when in fact, it is just an optional assumption (Baker and Kramer, 2003). Second, with the counterfactual model for potential surrogate endpoints, it may not be apparent that some strong assumptions are necessary for estimation. Consider a binary potential surrogate endpoint, such as "low" and "high" values for PSA doubling time, which is a potential surrogate endpoint for the recurrence of prostate cancer. By analogy with all-or-none compliance models (e.g. Baker and Lindeman, 1994), one needs to assume that there is no subject for whom the counterfactual potential surrogate endpoint is "low" if randomized to A and "high" if randomized to B, or vice versa. In many situations this is not likely to hold due to multiple pathways of cancer pathogenesis. Third, with models involving summary measures of potential surrogate endpoints, it is not clear how much confidence one can place in predictions of true endpoint based on trends in summary measures when individual level data do not show a trend or trend in a different way.

While there has been much work on statistical theory for predicting true endpoints, the practical validation has lagged, primarily because of few data on both surrogate and true endpoints in multiple previous trials. An important issue for more practical investigation is how much between-study variability affects the precision of predictions. A final word of caution: even if a potential surrogate is validated, it might not be desirable, as it provides no information on any harmful side effects of the intervention that might arise after observation of the potential surrogate endpoint, but prior to the true endpoint (Baker and Kramer, 2003). In such a case, a statistically validated surrogate endpoint could correctly predict a beneficial true endpoint, but miss the critical fact that the harms outweigh the benefits, and thus lead to an incorrect conclusion as to the value of the intervention.

2. Early detection markers for cancer

Early detection of cancer is not an end in itself. To be beneficial, it must lead to early intervention that reduces mortality from cancer. In addition, the harms must be small because so few of the usually targeted healthy population will ever die of the specific cancer. So only a small minority can theoretically benefit from the early detection test, but all are at risk for harm. This raises sticky ethical issues in promoting screening tests before a net benefit has been firmly established. With this idea in mind, the statistical issues become clearer. The main question is "What are the performance criteria for early detection biomarkers that would lead to the next phase in evaluation, namely testing the biomarker as a trigger for early intervention?" Because the markers are tested in asymptomatic persons for whom cancer is rare, it is important to minimize false positives and hence unnecessary biopsies. Therefore one criterion for an early detection biomarker is that the false positive rate (FPR), the probability of cancer detection in subjects without cancer, should be very small. A second criterion is that the true positive rate (TPR), the probability of detection given cancer, should be fairly high. For breast cancer detection reasonable target values based on costs and benefits are FPR=.01 and TPR=.80 (Baker, 1998; Baker et al, 2002). For other cancers the target values may differ due to differing "costs," both human and economic, of FPR's. If the marker is not binary, one obtains a series of (FPR, TPR) which form an ROC curve, and one should select the values closest to the target (Baker, 2003).

Data for early detection markers often come from a nested case-control study with biological specimens collected prospectively. At the end of follow-up, the marker is tested in specimens from all subjects with the cancer of interest and a random sample of subjects without cancer. See Baker et al (2002) for issues of quality control. If many markers are involved, one should fit the model in a training sample and get definitive results in a test sample. In this case, the ROC (Receiver Operating Characteristic) curve generalizes to the envelope of a cloud of points (Baker, 2000).

3. Early Detection Imaging Tests

Imaging tests for early detection of cancer are a special type of early detection marker. Unlike most molecular markers, their classification often involves subjective ratings. This makes specific classifications somewhat arbitrary so that FPR and TPR for a specific classification may be difficult to interpret. To circumvent this difficulty, interest focuses on the area under the ROC curve. Though not universally accepted, we believe that one should only report (or at least focus primarily on) the area under the part of the ROC curve corresponding to small FPR's rather than the entire area under the ROC curve, as is sometimes done. See Baker (2003). Another issue is how best to compare the performance of two imaging tests. A major complication is the missing information on cancer status when subjects do not receive a biopsy (Baker and Pinsky, 2003).

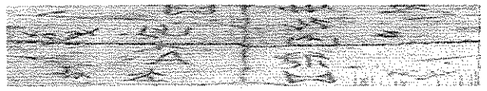
The above is only a brief outline of some of the controversial issues in this field. More details can be found in the references.

References

1. Baker SG (2000). Identifying combinations of cancer biomarkers for further study as triggers

of early intervention, Biometrics 56:1082-1087.
2. Baker SG. (2003) The central role of Receiver Operating Characteristic (ROC) curves in evaluating tests for the early detection of cancer, Journal of the National Cancer Institute 95: 511-5.
3. Baker S. G. and Kramer B. S. (2003) A perfect correlate does not a surrogate make. BMC Medical Research Methodology 3:16.
4. Baker S. G., Kramer, B. S. and Srivastava, S.(2002) Markers for early detection of cancer: Statistical issues for nested case-control studies. BMC Medical Research Methodology 2002, 2:4
5. Baker S. G., Lindeman K. S. (1994) The paired availability design: a proposal for evaluating epidural analgesia during labor. Statistics in Medicine 13:2269-2278.
6. Baker S. G. and Pinsky P. (2001) A proposed design and analysis for comparing digital and analog mammography: special ROC methods for cancer screening. Journal of the American Statistical Association, 96: 421-428.
7. Day, N. E. and Duffy, S. W. (1996), Trial design based on potential surrogate end points-application to comparison of different breast screening frequencies. Journal of the Royal Statistical Society A, 159, Part 1, 49-60.
8. Frangakis C. E. and Rubin D. B. (2002) Principal stratification in casual inference. Biometrics, 58, 21-29.
9. Freedman L. S., Graubard B. I., and Schatzkin A. (1992). Statistical validation of intermediate endpoints for chronic disease. Statistics in Medicine 11: 167-178.
10. Gail, M H., Pfeiffer, R, Houwelingen, H. C., Carroll, R. J. (2001). On meta-analytic assessment of surrogate outcomes. Biostatistics 2001, 3: 231-246.
11. Molenberghs, G., Buyse M.. Geys H., Renard D., Burzykowski T., and Alonso A. (2002) Statistical challenges in the evaluation of surrogate endpoints in randomized trials. Controlled Clinical Trials 23:607-625.
12. Prentice, R. L. (1989) Potential surrogate endpoints in clinical trials: Definitions and operational criteria. Statistics in Medicine, 8, 431-430.

Stuart G. Baker, Sc.D.
Biometry Research Group, Division of Cancer Prevention, National Cancer Institute,
Email: sb16i@nih.gov
Barnett S. Kramer, M. D., M. P. H.
Office of Disease Prevention, NIH.



Surrogate Endpoints

Geert Molenberghs, Ph.D.
Marc Buyse, Sc.D.
Tomasz Burzykowski, Ph.D.

1. The Concept of a Surrogate Endpoint

One of the most important factors influencing the duration and complexity of the process of developing new treatments is the choice of endpoint used to assess treatment efficacy. Two main criteria to select the endpoint are (1) its sensitivity to detect treatment effects and (2) its clinical relevance to the goals of the study (Fleming 1996). The relevance depends on, for example, whether evidence for biological activity of a drug is sought (Phase II trials) or whether a definitive evaluation of clinical benefit to patients has to be made (Phase III trials). For example, in life threatening diseases, such as cardiovascular diseases or cancer, the endpoint relevant for definitive evaluation of a treatment typically is survival.

It often appears, however, that the most sensitive and relevant clinical endpoint (the 'true' endpoint), might be difficult to use in a clinical trial. This can happen if measurement of the true endpoint: (1) is costly (e.g., to diagnose 'cachexia', a condition associated with malnutrition and involving loss of muscle and fat tissue, expensive equipment measuring content of nitrogen, potassium and water in patient's body is required); (2) is difficult (e.g., involving compound measures such as typically is the case in quality of life or pain

assessment); (3) requires a long follow-up time (e.g., survival in early stage cancers); or (4) requires a large sample size due to a low incidence of the event (e.g., short-term mortality in patients with suspected acute myocardial infarction). In such cases, use of the true endpoint increases the complexity and/or the duration of research. To overcome these problems, a seemingly attractive solution is to replace true endpoint by another one, which is measured earlier, more conveniently, or more frequently. Such 'replacement' endpoints are termed 'surrogate' endpoints (Ellenberg and Hamilton 1989).

2. Why Is There Reservation Towards the Use of Surrogate Endpoints?

Because of the possible benefits for the duration of a clinical trial, surrogate endpoints have been used in medical research for a long time (Ellenberg and Hamilton 1989, Fleming and DeMets 1996). Table 1 presents several examples. The use of the surrogate endpoints presented in Table 1 was based on an established *association* between them on the one hand and the corresponding true endpoints on the other hand. However, the mere existence of an association between a candidate surrogate endpoint and the true endpoint is not sufficient for using the former as a surrogate: 'a correlate does not make a surrogate' (Fleming and DeMets 1996). What is required is that the effect of the treatment on the surrogate endpoint reliably predicts the effect on the true endpoint.

Unfortunately, partly due to the lack of appropriate methodology, this condition was not checked in the early attempts to use surrogates. Consequently, for most of the surrogates mentioned in Table 1, it was found that their use, at least in some applications, led to erroneous, or even harmful, conclusions. A review of several such examples is given by Fleming and DeMets (1996). Probably the best known case is the approval by the Food and Drug Administration (FDA) in the United States of the use of three drugs: encainide, flecainide, and moricizine. The drugs were approved based on the fact that they were shown to effectively suppress arrhytmias. It was believed that, since arrhythmia is associated with an almost fourfold increase in the rate of cardiac-complication-related death, the drugs would reduce the death rate. However, a clinical trial conducted after the drugs had been approved by FDA and introduced into clinical practice showed that in fact the death rate among patients treated with encainide and flecainide was more than twice the one among patients treated with placebo (The Cardiac Arrhythmia Suppression Trial (CAST) Investigators 1989). An increase of the risk was also detected for moricizine. Subsequently, negative opinions about the use of surrogates in the evaluation of treatment efficacy have been voiced (Fleming 1996, Fleming and DeMets 1996, DeGruttola *et al.* 1997).

3. Why the Use of Surrogate Endpoints Is Still Being Considered

Thus, it will be clear that the very mention of surrogate endpoints has always been very controversial. However, not all early applications were failures. For example, the dramatic surge of the AIDS epidemic, the impressive therapeutic results obtained early on with zidovudine, and the pressure for an accelerated evaluation of new therapies, have all led to first the use of CD4 blood count and, with the advent of HAART, viral load, as endpoints that replaced time to clinical events and overall survival (DeGruttola and Tu 1995), in spite of some concerns about their limitations as surrogate markers for clinically relevant endpoints (Lagakos and Hoth 1992).

Currently, the number of candidate biomarkers and ultimately the number of surrogate endpoints based upon them is increasing dramatically. Indeed, an increasing number of new drugs have a well-defined mechanism of action at the molecular level, allowing drug developers to measure the effect of these drugs on the relevant biomarkers (Ferentz 2002). There is also increasing public pressure for new, promising drugs to be approved for marketing as rapidly as possible and such approval will have to be based on biomarkers rather than on some long-term clinical endpoint (Lesko and Atkinson 2001). The pressure can become especially high in a situation of a rapidly increasing incidence of a disease can become a serious threat to public health. This trend towards early decision-making is seen through recently proposed clinical trial designs, using treatment

effects on a surrogate endpoint to screen for treatments that show insufficient promise to have a sizeable impact on survival (Royston, Parmar, and Qian 2003). Last but not least, if the approval process is shortened, there will be a corresponding need for earlier detection of safety signals that could point to toxic problems with new drugs. It is a safe bet, therefore, that the evaluation of tomorrow's drugs will be based primarily on biomarkers, rather than on the longer-term, harder clinical endpoints that have dominated the development of new drugs until now. Another reason to shorten the duration of the process of testing new therapies may be related to new discoveries in medicine and biology, which create a possibility for development of many potentially effective treatments for a particular disease. In such a situation, the need to cope with a large number of new promising treatments that should be quickly evaluated with respect to their efficacy might appear. This can already be observed in oncology, as the increased knowledge about the genetic mechanisms operating in cancer cells led to the proposal of qualitatively new approaches to treat cancer. An example is found in the use of a genetically-modified virus that selectively attacks p53-deficient cells, sparing normal cells (Heise *et al.* 1997).

Further, shortening the duration of a clinical trial also limits possible problems with non-compliance and missing data, which are more likely in longer studies, and therefore increases effectiveness and reliability of the research.
Finally, an important area of potential application of surrogate endpoints is the assessment of safety of new treatments. Duration and sample size of clinical trials aimed at development of new drugs are usually insufficient to detect rare or late adverse effects of the treatment (Jones 2001, Dunn and Mann 1999). The use of surrogate endpoints (for toxicity-related clinical endpoints) might allow one to obtain information about such effects even during the clinical testing phase.

4. Statistical Evaluation of Surrogate Endpoints

Thus, while avoiding surrogate markers and surrogate endpoints does not seem to be an option, the failed past attempts to use surrogate endpoints make it clear that, before deciding on the use of a candidate surrogate endpoint, it is of the utmost importance to evaluate its validity and hence formal methods allowing for validation are required. Such methods have become the subject of intensive research over the past couple of decades. Such methods should be seen as a factor in the decision process, not as a substitute for it.

Prentice (1989) proposed a formal definition of surrogate endpoints and outlined a set of criteria, all within the framework of hypothesis testing. Much debate ensued, for the criteria set out by Prentice are not straightforward to verify (Freedman *et al.*, 1992, Fleming *et al.*, 1994). In addition, Prentice's criteria are only equivalent to the definition he proposed in the case of binary endpoints (Buyse and Molenberghs 1998). Freedman *et al.* (1992) supplemented Prentice's approach by introducing the *proportion explained* (PE), aimed at measuring the proportion of the treatment effect mediated by the surrogate. This proposal was important in that it shifted the interest in the validation of surrogate endpoints from significance testing to estimation. However, it is itself surrounded with difficulties. For example, it is not restricted to the unit interval and violations do occur frequently in practice (Molenberghs *et al.*, 2003). Consequently, Buyse and Molenberghs (1998) proposed to replace it by two new measures. The first one, defined at the population level and termed *relative effect* (RE), is the ratio of the overall treatment effect on the true endpoint over that on the surrogate endpoint. The second one is the individual-level association between both endpoints, after accounting for the effect of treatment, and referred to as *adjusted association*. These proposals are not free from problems either. Indeed, the framework sketched above required data from a single trial only, its main drawback being the lack of replication of the treatment effect. Therefore, Daniels and Hughes (1997) and Buyse *et al.* (2000a) proposed to undertake validation in a meta-analytic or multi-centric context, thereby allowing replication not only at the patient-level, but also at the trial level. These authors concentrated on continuous responses and employed linear mixed-effects models (Verbeke and Molenberghs 2000). The quality of a surrogate is quantified using R^2 type measures. In addition, one can derive prediction equations for the true treatment effect in a new trial, based on the treatment effect of the surrogate endpoint.

As a consequence of this paradigm shift, for different types of endpoints, a different joint modelling strategy is obviously necessary. For two binary outcomes, a joint probit model is proposed (Renard *et al.*, 2002a). Molenberghs *et al.* (2001) considered the combination of continuous and discrete endpoints.
A copula model for joint survival data is used for censored time-to-event outcomes and for a combination of time-to-event and categorical endpoints (Burzykowski *et al.*, 2001, 2003, Buyse *et al.*, 2000b).

When the surrogate is measured repeatedly over time and the true endpoint is of a time-to-event nature, one is confronted with the joint modelling of longitudinal and survival data. A considerable literature exists in this area (Boscardin *et al.*, 1998). Renard *et al.* (2002b) show how such a joint model can be implemented in the context of surrogate marker validation.

5. Where Do We Go From Here?

The approach outlined above combines the concepts initially formalized by Prentice (1989) with a hierarchical view, using evidence from a multitude of trials rather than from a single one. The intuitive appeal of this approach is the requirement of a good surrogate to satisfy two distinct but similarly looking properties. First, the surrogate endpoint must predict the true endpoint in individual patients. This condition is in fact strongly related to Prentice's definition. Second, the effect of a treatment on the surrogate endpoint must predict the effect of that treatment on the true endpoint. While different, at face value, from Prentice's fourth criterion, this condition is consistent with its spirit, properly exploiting replication across trials.

While we like to underscore the elegance of such a meta-analytic framework, important questions remain open. First, the hierarchical framework is computationally more involved, and requires the number of trials and the number of patients per trials to be sufficiently large. Tibaldi *et al.* (2003) have investigated a number of simplified approaches, where a fully hierarchical analysis is replaced by a two-stage approach and/or the two endpoints are analyzed separately. The latter is convenient when only the trial-level is of interest. Second, in several of the examples we analyzed, we used, by way of poor man's choice, 'center' or 'investigator' as sub-unit, rather than trial. Cortinas *et al.* (2004) have investigated the impact of either ignoring or shifting between hierarchical levels. Thus, it is of great interest, both to the public and to the scientific community, that data be shared to undertake the widest possible meta-analytic evaluations, rather than being considered the sole propriety of pharmaceutical companies. Third, the use of complex hierarchical models implies that different surrogacy measures are proposed for different types of outcomes, especially at the individual level. Indeed, while R^2 measures are used throughout at the trial level, individual-level measures include R^2, the odds ratio, Kendall's □, etc. Alonso *et al.* (2003) initiated the investigation to unify the various approaches. Fourth, the models considered so far reflect practice within later phase clinical trials, in the sense that, apart from treatment assignment, no other explanatory information is used. It is conceivable, especially in earlier phase trials and in preclinical research, that more elaborate models be used, incorporating explanatory (baseline) covariates, (molecular) biological, pharmacokinetic, or pharmacodynamic information. Fifth, Gail *et al.* (2000) have indicated that not properly accounting for measurement error may paint too optimistic a picture about surrogacy. It remains to be explored, theoretically and empirically, how useful surrogate markers are when all sources of measurement error are taken into account. Sixth, and linked to the previous issue, is the question how a properly evaluated surrogate endpoint could be used when designing a new trial. For example, one may want to determine the sample size to allow prediction of a significant effect on the true endpoint, without actually measuring it. Seventh, a properly evaluated surrogate endpoints will rarely be universally valid. The difficult question remains as to how broad the class of drugs is within which it can be used.

Table 1: Surrogate endpoints used in medical research.

Disease	Endpoints	
	Surrogate	True
Early stage cancer	Time to progression	Survival time
Advanced cancer	Tumor response	Survival time
Osteoporosis	Bone mineral density	Bone fracture
Ophthalmology (glaucoma)	Intraocular pressure	Long-term visual acuity
Chronic granulomatous disease	Superoxide production	Serious infection
Cardiovascular disease	Ability to kill bacteria	Serious infection
	Ejection fraction	Myocardial infarction
	Blood pressure	Stroke, survival time
HIV infection	Arrhytmias	Survival time
	CD4 counts; viral load	Development of AIDS, Survival time

References

1. Alonso, A., Molenberghs, G., Burzykowski, T., Renard, D., Geys, H., Shkedy, Z.,Tibaldi, F., Cortinas, J., and Buyse, M. (2003). Prentice's approach and the meta-analytic paradigm: a reflection on the role of statistics in the evaluation of surrogate endpoints. Submitted for publication.
2. Boscardin, W.J., Taylor, J.M.G. and Law Ngayee (1998). Longitudinal models for AIDS marker data. Stat. Methods in Med. Res. 7, 13-27.
3. Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H., and Renard, D. (2001). Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. Applied Statistics 50, 405-422.
4. Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H., and Renard, D. (2003). The validation of surrogate endpoints using data from randomized clinical trials: a case study in advanced colorectal cancer. J. of the Royal Stat. Society, Series A.
5. Buyse, M. and Molenberghs, G. (1998). The validation of surrogate endpoints in randomized experiments. Biometrics 54, 1014-1029.
6. Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., and Geys, H. (2000a). The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics 1, 49-67.
7. Buyse, M., Thirion, P., Carlson, R.W., Burzykowski, T., Molenberghs, G., and Piedbois, P., for the Meta-Analysis Group In Cancer (2000b). Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Lancet 356, 373-378.
8. Cardiac Arrhythmia Suppression Trial (CAST) Investigators (1989). Preliminary Report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infraction. NEJM 321, 406-412.
9. Cortinas Abrahantes, J., Molenberghs, G., Burzykowski, T., Shkedy, Z., and Renard, D. (2004). Choice of units of analysis and modeling strategies in multilevel hierarchical models. Computational Statistics and Data Analysis.
10. Daniels, M.J. and Hughes, M.D. (1997). Meta-analysis for the evaluation of potential surrogate markers. Stat. in Med. 16, 1515-1527.
11. De Gruttola, V., Fleming, T.R., Lin, D.Y., and Coombs, R. (1997). Validating surrogate markers - are we being naive? J. of Inf. Dx. 175, 237-246.
12. DeGruttola, V. and Tu, X.M. (1995). Modelling progression of CD-4 lymphocyte count and its relationship to survival time. Biometrics 50, 1003-1014.
13. Dunn, N. and Mann, R.D. (1999). Prescription-event and other forms of epidemiological monitoring of side-effects in the UK. Clinical and Experimental Allergy 29, 217-239.
14. Ellenberg, S.S. and Hamilton J.M. (1989). Surrogate endpoints in clinical trials: cancer. Statistics in Medicine 8, 405-413.
15. Ferentz, A.E. (2002). Integrating pharmacogenomics into drug development. Pharmacogenomics 3, 453-467.
16. Fleming, T.R. (1996). Surrogate endpoints in clinical trials. Drug Info. J. 30, 545-551.
17. Fleming, T.R. and DeMets, D.L. (1996). Surrogate endpoints in clinical trials: are we being misled? Annals of Internal Med. 125, 605-613.
18. Fleming, T.R., Prentice, R.L., Pepe, M.S., and Glidden, D. (1994). Surrogate and auxiliary endpoints in clinical trials, with potential applications in cancer and AIDS research. Statistics in Medicine 13, 955-968.
19. Freedman, L.S., Graubard, B.I., and Schatzkin, A. (1992). Statistical validation of intermediate endpoints for chronic diseases. Statistics in Medicine 11, 167-178.
20. Gail, M.H., Pfeiffer, R., van Houwelingen, H.C., Carroll, R.J. (2000). On meta-analytic assessment of surrogate outcomes. Biostatistics 1, 231-246.
21. Heise, C., Sampson-Johannes, A., Williams, A., McCormick, F., Von Hoff, D.D., and Kirn, D.H. (1997). ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemo-therapeutic agents. Nature Medicine 3, 639-645.
22. Jones, T.C. (2001). Call for a new approach to the process of clinical trials and drug registration. British Medical Journal 322, 920-923.
23. Lagakos, S.W. and Hoth, D.F. (1992). Surrogate markers in AIDS: Where are we? Where are we going? Annals of Int. Med. 116, 599-601.
24. Lesko, L.J. and Atkinson, A.J. (2001). Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. Annual Review of Pharmacological Toxicology 41, 347-366.
25. Molenberghs, G., Geys, H., and Buyse, M. (2001). Evaluation of surrogate endpoints in randomized experiments with mixed discrete and continuous outcomes. Stat. in Med. 20, 3023-3038.
26. Molenberghs, G., Burzykowski, T., Alonso, A., and Buyse, M. (2003). A perspective on surrogate endpoints in controlled clinical trials. Statistical Methods in Medical Research.
27. Prentice, R.L. (1989). Surrogate endpoints in clinical trials: definitions and operational criteria. Statistics in Medicine 8, 431-440.
28. Renard, D., Geys, H., Molenberghs, G., Burzykowski, T., and Buyse, M. (2002a). Validation of surrogate endpoints in multiple randomized clinical trials with discrete outcomes. Biometrical Journal 44, 921-935.
29. Renard, D., Geys, H., Molenberghs, G., Burzykowski, T., Buyse, M., Vangeneugden, T., and Bijnens, L. (2002b). Validation of a longitudinally measured surrogate marker for a time-to-event endpoint. J. of App Stat 30, 235-247.
30. Royston, P., Parmar, M.K.B., and Qian, W. (2003). Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer. Stat. in Medicine 22, 2239-2256.
31. Tibaldi, F.S, Cortinas Abrahantes, J., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., Stijnen, T., and Wolfinger, R. (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. J. of Statistical Comp. and Simulation 73, 643-658.
32. Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. New York: Springer-Verlag.

Geert Molenberghs and Tomasz Burzykowski are Professors of Biostatistics at CenStat, Limburgs Universitair Centrum, Diepenbeek, Belgium.

Marc Buyse is Director of the International Drug Development Institute, Brussels, Belgium.

Email: geert.molenberghs@luc.ac.be, tomasz.burzykowski@luc.ac.be, marc.buyse@iddi.com.

Using Genomics and Proteomics Biomarkers for Disease Classification

By: Ziding Feng, Ph.D.

1. Introduction

The recent advances in genomics (gene expression arrays and SNPs), proteomics (protein expression using mass spectrometry or antibody arrays) opened the door for biomedical researchers to combine multiple biomarkers measured using noninvasive procedures (e.g., samples from serum, urine, stool) for disease classifications (detection, diagnosis, and prognosis). The hope is high because we suddenly have at our disposal thousands of biomarkers from which we believe there should be a signal for disease and there seem to be some promising findings in which the microarray biomarker panels (1) or protein fingerprints (2) correlated well with diseases.

However, this great promise comes with great challenges and dangers. First, many people, encouraged by the reported high sensitivities and specificities of these biomarker panels, are jumping into these types of studies without fully understanding the caveats and the requirements for rigorous evaluation. Second, there are many data mining methods out there and lacks guidance on how to choose them to construct a biomarker panel. The number of ways of combining multiple biomarkers increases exponentially with the number of biomarkers. As a consequence, searching for the optimal combination of biomarkers is not only computationally demanding but also run the high risk of false positive findings. False positive findings waste precious resources used for further investigations, or potentially cause harm to patients if they are put for clinical use without firm validations.

In this article I want to discuss three issues related to the concerns raised above: a. "the curse of dimensionality"; b. Likelihood principle; and c. the use of protein "fingerprints" as biomarker.

2. Curse of dimensionality

The term of "Curse of dimensionality" first came from Bellman (3). It basically says that the required number of samples to uncover truth grows exponentially fast with the dimension of data (e.g., the number of biomarkers). For example, one would feel uncomfortable to study a relationship between Y, disease status (1 for diseased, 0 for non-diseased), and X, the value of a single biomarker, if only three observations are given. One would probably feel comfortable if information about Y and the values of 10 biomarkers are given on 59,000 observations. However, the data sparseness for these two settings is about the same because the data density is proportional to $N^{1/p}$, where N is the number of observations and p is the number of biomarkers. With 10 biomarkers, numerous complex combinations of them exist. The reason we feel comfortable in the later case is our assumption of no high order interactions among 10 biomarkers. This may not hold as it was discovered that even in simple organism like yeast there are complicated mechanisms, i.e., complicated interactions among genotypes, to regulate its functions. However, in order to uncover the complicated interactions, "the curse of dimensionality" is in the way. This leads to three important messages. First, whenever possible, simple classification rules using small number of biomarkers guided by subject knowledge are preferable to more complex ones. Second, findings from such sparse data should be treated as exploratory and require further confirmations. Third, the models (classification rules) built from such high dimensional sparse data have a great risk of *overfit*, a phenomenon that a model has very low prediction error in the data used for model building but has poor performance (generalizability) on new independent data.

To avoid an overfit, we need to assess the model very carefully. The best approach is to set aside samples that are not used for model construction as a *test set* and evaluate the classification performance in this test set after the model (disease classifier) is finalized from the *training set* of samples. This test set should be used only once after the model is fixed. Repeatedly update the model based on the test set performance will render the value of the test set to a training set, and could not provide an unbiased estimate of the prediction error on new data. Also, fixing more than one final models from the training set and taking the best test-set performance among them is not a proper use of the test set: this scheme uses the test set as part of the training set.

Note that assessments of prediction errors serve two purposes: for selecting a model with the minimum estimated future prediction error, and also for estimating future prediction error of the model. The test set is useful for the latter purpose. The former purpose could be achieved by Cross-Validation (CV) procedure (4).
When we cannot have enough samples to split into training and test sets, we need to assess the future prediction error for the final chosen model. This can be achieved by using bootstrap (5). We randomly draw, with replacement, N observations from the original N observations we have in the observed data to form a bootstrap set of N samples. For each bootstrap set of N samples, we repeat the above Cross-Validation model selection process to get a final model from that bootstrap set and use this model to predict observations in the original samples that were not selected in this bootstrap sample, and compute the prediction error among them. Repeat this process B times (at least 100) and average the prediction errors. We call this quantity Validation Prediction Error. This is our estimate of the true prediction error in the absence of a test set. The best Cross-Validation Error from a given dataset is likely to be somewhat too optimistic (smaller) than the Validation Prediction Error of that dataset, or true future prediction error, because the model selection used the same data as those used for calculating the Cross-Validation Error of the final model.

3. Likelihood principle

An application of Neyman-Pearson Lemma to diagnostic tests has led to remarkable insights about combining multiple tests (6-10) that we elaborate here. It says that, with a biomarker covariate vector $X=(X_1, \ldots, X_p)$ and a disease status indicator Y (1 for diseased, 0 for non-diseased), the optimal classification rule is to predict Y=1 for a sample if $LR(X) > c$, where $LR(X) = P(Y=1|X) / P(Y=0|X)$. $LR(X)$ is called a likelihood ratio for Y=1 vs. Y=0 given X. The threshold c is determined by the requirement of sensitivity or specificity. The optimality here means that, among all classification rules, the optimal rule has the highest sensitivity for a fixed specificity, the highest specificity for a fixed sensitivity, minimize the overall misclassification probability, and minimize the expected cost regardless of the cost balance between false positive and false negative results.

An equivalent, but more practical, way of applying this optimal rule is to employ the following Risk Score rule: $RS(X) > c^*$, where $RS(X) = P(Y=1|X)$. $RS(X)$ is a monotone increasing function of $LR(X)$ so this amounts to choose a c^* in $RS(X)$ for a corresponding c in $LR(X)$.

Both $LR(X)$ and $RS(X)$ require a correct probability model for the relationship between multivariate biomarkers X and the disease status Y in order to obtain the optimality. In reality, we even don't know which subset of biomarkers is related to disease, even less on how they are related. Therefore, the optimality is rarely achievable. However, understanding this principle will guide us in choosing procedures to combine multiple biomarkers. The procedures that approximate the principle are likely to have better performance than that violate this principle.

One application of the likelihood principle explains why logistic regression model performs well under very wide settings. logit(E[Y]) is a monotone function of RS(Y|X) so it is optimal as long as the logistic model we use with X is monotone function of the unknown true model. This makes the logistic regression robust for model mis-specification. Adding interaction terms are easy in logistic regression models, although they need to be specified. Another interesting example of using LR to combining biomarkers without the knowledge of the joint distribution of biomarkers is due to Baker (8). He estimated risk scores non-parametrically.

4. Use of protein fingerprints as biomarker.

For serum protein/peptide spectra from a time-of-flight mass spectrometry, for example, the peptide or protein identities are unknown without performing further intensive protein/peptide

identification steps in laboratories. A number of studies attempted to construct classifiers for diseases using protein fingerprints (2, 11, 12). Though the findings seem promising, it is necessary to understand the limitations. One major difficulty is that there are measurement errors both in mass points and intensities. The first one is more problematic because it causes mislabeling of the "biomarker" candidates. This leads to the needs for protein peak identification and alignment. They are analytically challenging but not discussed here: see, for example, (13). Instead we discuss the statistical considerations after the peaks are identified and aligned. Analytically we can treat these peaks as "biomarkers" and then apply your favorite analysis method. However, to push this type of fingerprint biomarker for clinical use, more verification is needed.

First, after initial findings of "fingerprints", their reproducibility must be established firmly before further confirmatory studies are conducted. If biomarkers are known such as the cases in microarray gene expression and antibody array experiments, after an initial positive finding was made for a biomarker panel for disease classification, a natural step is to confirm the finding in a larger sample, a sample from more heterogeneous population, and/or confirm the panel against specified sensitivity and specificity that are of clinical significance. Reproducibility issues are still important but are more for the consideration of assay scale up for routine clinical use. However, with only "fingerprints" at hand, the first step is to show that the fingerprints identified are robust: reproducible and accurate. Protein fingerprints must have mass and intensity measurements that are reproducible and accurate enough such that they will lead to the same classification when the diagnostic model is applied on their spectra obtained from repeated runs, preferably from different instruments at different labs. Existence of the fingerprints observed in repeated assays of a single sample, for example, is not adequate to claim reproducibility and accuracy. The ability to generate another panel/ model with good performance using a new set of data is not adequate, either: it has to be the same panel/model with good performance on the new data. Otherwise, the model has no clinical use.

Second, the reproducibility and accuracy have to be examined on samples from multiple studies/populations than the study/population where the initial findings were made. The objective of this step, when biomarkers are known, is to examine whether the same biomarker panel works for a more general population and fine-tune the model if necessary. With "fingerprints" data, however, the main objective is more basic: mainly to find out whether the "fingerprints" of the initial findings represent artifacts or they are likely to be real biological signals. For example, if there were differences between disease and non-disease groups in the methods for sample collection, preparation, and storage in a study, the "fingerprints" we identify from this study would reflect these experimental differences between disease and non-disease, an artifact, not a real biological signal of the disease. The distinction between artifacts and real biological signals cannot be made by any analytical approach without data from different studies/populations.

Basically, the hurdle is higher for a biomarker panel to pass to the next validation phase if the identities of these biomarkers are unknown. There are more to prove and to rule out. That is the price to pay when the identities are unknown.

One fruitful use of the "fingerprints" approach is the use of fingerprints to narrow down the region for further studying of protein/peptide identification/ function. This is a very challenging problem because the high dimensionality of proteomic data often allows a large number of models to have "good" performance in a given dataset so one needs to narrow down from them to make the phase of protein/peptide identification/function study practical. Creative experimental design and analysis are needed.

5. Concluding remarks

Several messages I want to emphasize here: First, understand the exploratory nature of high dimensional genomics and proteomics data analyses and do not over-interpret the findings. Second, work closely with biologists and apply the subject knowledge to reduce dimensionality of the biomarker candidates to a subset small enough to apply likelihood principles. Third, when biology basis for biomarker is lacking, verify, verify, and verify, and remember that there is no free lunch.

Reference

1. Van De Vijver MJ, et al. A gene-expression signature as a predictor of survival in breast cancer. NEJM 2002; 347:1999-2009.
2. Petricoin EF, et al. Use of proteomic patterns in serum to identify ovarian cancer. The Lanet 2002; 359:572-577.
3. Bellman RE. Adaptive Control Processes. Princeton University Press. 1961.
4. tone M. Cross-validatory choice and assessment of statistical predictions. J. Royal Statistical Society; 36: 111-147.
5. Efron B, Tibshirani R. An Introduction to the Bootstrap. New York: Chapman & Hall, 1993.
6. Green DM, Swets JA. Signal detection theory and psychophysics. New York: Wiley, 1966.
7. Egan JP. Signal detection theory and ROC analysis. New York: Academic Press, 1975.
8. Baker SG. Identifying combinations of cancer markers for further study as triggers of early intervention. Biometrics 2000; 56:1082-1087.
9. McIntosh M, Pepe MS. Combining several screening tests: Optimality of the risk score. Biometrics 2002; 58:657-664.
10. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. New York: Oxford University Press, 2003.
11. Adams B, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from Benign Prostate Hyperplasia and healthy men. Cancer Research 2002; 62:3609-3614.
12. Rai AJ, et al. Proteomic approaches to tumor marker discovery. Arch Pathol Lab Med 2002; 126:1518-1526.
13. Yasui Y, McLerran D, Adam BL, Winget M, Thornquist M, Feng Z. An automated peak-identification / calibration procedure for high-dimensional protein measures from mass spectrometers. Journal of Biomedicine and Biotechnology 2003; Special Issue: 1-7.

Ziding Feng, Ph.D. Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109
Email: zfeng@fhcrc.org

Some Statistical Issues and Utility of Surrogate Biomarker Endpoint?

By: Donald J. Johann, Jr., M.D.
Lance A. Liotta, M.D., Ph.D.
Emanuel F. Petricoin III, Ph.D.

New types of proteomic biomarker technologies that correlate with clinically meaningful outcomes are active research areas in Surrogate Biomarker Endpoint (SBE) scientific field. Clinical trials incorporating advanced biomarker techniques are scheduled to begin in the near future at the National Cancer Institute (NCI), Center for Cancer Research. The first trial will be to clinically evaluate a multiplexed biomarker "pattern" obtained from mass spectrometer analysis of the low molecular weight proteome present in blood for monitoring the recurrence of ovarian cancer. The pattern is comprised of ions whose combined relative intensities serve as a discriminator for a disease phenotype. The resultant pattern is derived using algorithmic methods employing machine learning and pattern matching techniques in which a training set and a blinded separate test set of controls and cases are analyzed.

Several feasibility studies using new biomarker discovery techniques based on high throughput mass spectrometer analysis of serum have shown predictive abilities in the identification of previously difficult to detect disease. Fundamental aims of these new classes of biomarkers include: improving the early diagnosis of disease, developing new methods to measure efficacy of therapy, and advancing the existing methods to detect therapy-induced toxicity. But, the high throughput technologies of molecular medicine have significant challenges. First, these new techniques involving SBE approaches need to be effectively integrated into clinical trial design. Secondly, the resultant data streams are massive and multidimensional (upwards of 1-2 million data points per sample) with a relative paucity of observations needing innovative analyses methods.

Since clinical trials are time consuming and expensive, there is great interest to quicken trial times and reduce costs. SBE methods are being evaluated on a case-by-case basis to add efficiency to trail design and help speed the advances of molecular medicine to the public. However, it is known surrogate techniques can be misleading. For instance, at first glance it may appear CD4 counts could be used as an effective screen for HIV/AIDS. Then again, CD4 counts can be affected by other infections (colds, flu), can have a diurnal variation, and exhibit a wide range of results when HIV infection is present. Therefore, vigilance is critical in the group analyses of outcome measures, SBEs, and potential covariate or confounding factors. This is especially true as the heterogeneity of disease processes, patient populations, and sample collection and handling methods is further explored and unraveled.

Notwithstanding these concerns, aggressive pursuit of new SBE methods is imperative because new classes of biomarkers are needed. Only ten new protein analytes have been approved by the FDA under CLIA regulation since 1993 [12]. Current biomarker methods based on solitary proteins are unreliable and often fail due to poor specificity in the general population. Many diseases have no reliable method of early detection and thus medical outcomes are generally poor because the disease has become advanced before treatment commences. Consequently, physicians have to accept inaccurate biomarkers as a "better than nothing" type strategy. For instance, ovarian cancer has a poor outcome because it is usually diagnosed after it has already reached a disseminated stage. Cancer Antigen 125 (CA-125) is a solitary protein and is used in FDA approved tests to evaluate for ovarian cancer recurrence. However, in the community this test is frequently used as a screening method, although it is not FDA approved for this role. Community physicians are well aware CA-125 has no real clinical utility at detecting early stage ovarian cancer, but use it as an "off-label" test, as there is no other marker to measure this disease. A new diagnostic test that could effectively identify early stage cancer could have a dramatic impact on this disease today since stage I ovarian cancer has a 5 year survival rate of over 95%. A second example concerns pancreatic cancer, which on the whole, is an insidious and deadly disease. It does not have any effective screening test, is normally detected at a very advanced stage, and thus has a very poor outcome. Several clinical pilot studies have shown that new types of biomarkers involving a proteomic pattern comprised of peptides/protein fragments will greatly improve the ability to discriminate the absence or presence of ovarian and pancreatic cancer and thus improve risk stratification and outcome[1-5, 9].

Serum proteomic tests represent a new form of SBE. The low molecular weight serum proteome is an unexplored archive of metabolic and physiologic information that can reflect the changes within an organ system. Within defined feasibility study sets, experimentalists have used serum proteomic information content to detect the onset of disease at a very early stage, thus maximizing the likelihood of successful outcomes. However, this new source of information has its challenges and a very significant one involves the analysis and management of this data stream. The datasets are of massive dimensions with relatively few observations. Many standard statistical methods do not work well with small sample sizes. Additionally, attempting analysis of massive datasets with traditional spreadsheet programs and plotting tools is impractical. New SBE methods will therefore involve FDA filing of innovative software tools for the management and analysis of this challenging data [10].

Visual data mining methods in combination with machine learning techniques have been successfully used to explore and classify with the goal to provide a means to develop an intuitive understanding of this data that goes beyond the numbers to underlying biologic truths. Massive multi-dimensional datasets containing diagnostic information bring a unique challenge in regard to the perceived clinical utility of a proteomic pattern based diagnostic test. Pattern hunting occurs in data streams composed of 300,000 to a few million records per patient sample. Discriminating information has been found to be contained within specific regions of the dataset, in the form of a greater or lesser abundance of a peptide/protein fragment in a diseased state compared to the unaffected state. Aggregating patient samples allows for the identification of global trends, isolation and removal of defective data, and the

formation of reduced datasets for additional analyses leading to a predictive model [6].

Concordance of multiple tools using different algorithmic methods has been found to lend significant support to the findings of visualization methods and machine learning techniques. Visualization tools combined with machine learning techniques have fostered both hypothesis generation and discovery. A significant discovery is the discriminating ability of groups of low abundant ions for the ovarian cancer group verses unaffected across a multitude of m/z values. Analytical methods utilizing *a priori* spectral filtering with subsequent peak picking strategies would have most likely excluded the enriched low abundance regions of spectra.

The heterogeneity of disease and populations may be best modeled through a multiplexed pattern based biomarker. When issues of diversity are considered, the unreliability (or absence) of disease markers based on solitary proteins is more understandable [7]. New synergistic SBE techniques employing the combined abilities of conventional or novel medical imaging coupled with a serum proteomic test offer the potential for a less invasive method in the evaluation of perplexing clinical problems such as: tumor re-growth verses necrosis, tumor verses infarct, assessment of solitary lung nodules, etc. Central to such breakthroughs are advanced statistical methods enabling robustness in both discovery and validation of a component-based SBE, utilizing a pattern based technique, and possibly linked to a second modality.

Public funding of the Human Genome Project and the NCI-FDA Clinical Proteomics Program are contingent on goals to speed molecular medicine advances and revolutionize clinical medicine especially in the areas of early detection, prevention, and prediction. Both genomics and proteomics employ multiplexed biomarker techniques. Specific standards and criteria are needed for high throughput molecular medicine technologies to enter into clinical practice. These issues are now beginning to be more formally addressed [11]. Teamwork and multidisciplinary approaches are essential to further research in the SBE field, thus enabling more effective clinical trial design, and the hope of quicker benefits to advances in the detection and medical management of disease and public health problems.

References

1. Petricoin EF, et. al. (2000). Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359: 572-577.
2. Petricoin EF, et. al. (2000). Serum proteomic patterns for detection of prostate cancer. *JNCI* 94.
3. Li J, et. al. (2002). Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem* 48: 1296-1304
4. Adam B-L, et. al. (2002). Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men, *Cancer Research* 62: 3609-3614
5. Baggerly KA, et al. (2003). A comprehensive approach to the analysis of matrix-assisted laser desorption/ionization-time of flight proteomics spectra from serum samples. *Proteomics* 9.
6. Johann D, et al. (2004). Novel approaches to visualization and data mining reveals enrichment of diagnostic information in the low amplitude region of serum mass spectra from ovarian cancer patients. *Disease Markers*. In press.
7. Liotta L, et al. (2003). Written in blood. *Nature* 425: 905.
8. Liotta L, et al. (2001). Clinical Proteomics Personalized Molecular Medicine. *JAMA* 286: 2211-2214.
9. Hingorani S, et al. (200=3). Preinvasive and Invasive Ductal Pancreatic Cancer and its Early Detection in the Mouse, *Cancer Cell* 4: 437-450
10. Johann, D, et al. (2004). Clinical Proteomics and Biomarker Discovery. *NY Acad. Sci.* In press.
11. Ntzani E, et al. (2003). "Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment", *Lancet* 362:1439-1444
12. Anderson L, et al. (2002). The human plasma proteome: history, character, and diagnostic prospects. *Molecular and Cellular Proteomics* 1: 845-867

Donald J. Johann, M.D.,[+]

Lance A. Liotta, M.D.,[+]

Emanuel F. Petricoin, Ph.D.[#]

+NCI-FDA Clinical Proteomics Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892

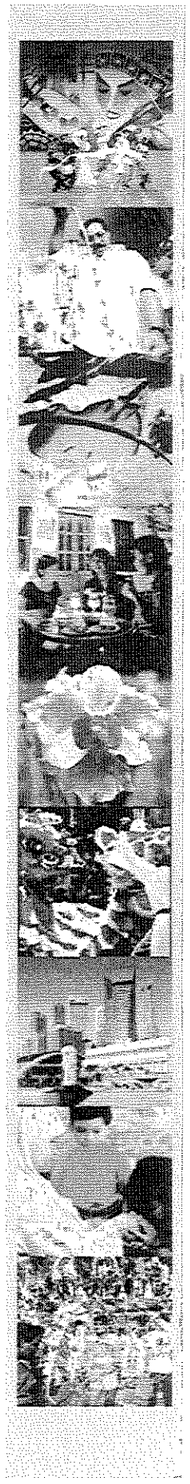
[#] NCI-FDA Clinical Proteomics Program, Office of the Director, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, MD 20892
Email: dj151o@nih.gov

Errata Correction: In the July 2003 issue of the ICSA Bulletin, the last word in the article by D'Agostino, Joe Massaro, and Don Cutlip should be "uninterpretable" instead of "uninterruptible".

Upcoming Controversial Statistical Issue In July 2004

"Genomics Data Analysis"

Interested contributors please contact the Special Topic Editor: Dr. Sue-Jane Wang (wangs@cder.fda.gov).



NUS National University of Singapore

International Chinese Statistical Association 泛華統計協會

The 6th ICSA International Conference, Singapore, July 21-23, 2004

Co-sponsored by the INSTITUTE OF MATHEMATICAL STATISTICS

3RD ANNOUNCEMENT

The Sixth ICSA International Conference will proceed as scheduled at the National University of Singapore (NUS), Singapore, from July 21 to 23, 2004.

Registration forms together with the fees and abstracts should reach the local organizing committee by latest 31 March 2004. The registration fees are as follows:

	On or before 31 March 2004	After 31 March 2004
Regular	U.S. $160	U.S. $190
Student (proof of status required)	U.S. $80	U.S. $95

The registration fee is inclusive of all conference materials, lunches and tea breaks, reception for the first night, dinner for the second night and transportation. It does not include cost of accommodation and local tours.

Participants will have to arrange for their own accommodations. The local organizing committee has negotiated special rates, just for the ICSA 2004 foreign participants, at the Hilton Singapore, Orchard Hotel, Orchard Parade Hotel and RELC International Hotel. Limited on-campus accommodation will be available on a first-come-first served basis.

The Programme Committee has been further expanded and it now includes: Zhiliang Ying (Columbia University and Chair of this committee), Zhidong Bai (NUS), Jia-Hua Chen (University of Waterloo), Louis H. Y. Chen (NUS), Paul Cheung (Chief Statistician of Singapore), Jianqing Fan (Princeton University and Chinese University of Hong Kong), Xuming He (University of Illinois at Urbana-Champaign), Mei-Ling T. Lee (Harvard University) Dennis Lin (Pennysylvania State University), Xiaoli Meng (Harvard University), Kai W. Ng (University of Hong Kong), Qi-man Shao (NUS and University of Oregon), Lingshi Tan (Pfizer Inc), Young Truong (NUS and University of North Carolina at Chapel Hill), Qiwei Yao (London School of Economics), Yi-ching Yao (Institute of Statistical Science, Academic Sinica) and Lincheng Zhao (University of Science and Technology of China)

More information on the conference, hotel accommodation, registration fees and submission of abstracts can be obtained from the NUS-ICSA 2004 website at http://www.stat.nus.edu.sg/ICSA.htm.

Enquiries on the conference can be sent to:

The Secretariat, ICSA 2004
Department of Statistics and Applied Probability
National University of Singapore
6 Science Drive 2
Singapore 117546
Fax: (65)-68723919, Email: icsa@nus.edu.sg

ICSA 2004 APPLIED STATISTICS SYMPOSIUM

June 6-9, 2004 at San Diego, California, U.S.A.

Theme: Statistics in Bio-tech Research and Computing Intensive Methodologies

DATE: June 6 to 9, 2004. Short courses on Sunday, June 6, and technical sessions from Monday, June 7 to Wednesday, June 9.

LOCATION: The San Diego Marriott - La Jolla, 4240 La Jolla Village Dr., La Jolla CA 92037

DEADLINES:

February 28, 2004: Student Award and Travel Fellowships
March 31, 2004: Abstract submission
April 15, 2004: Early short course registration
May 1, 2004: Early Symposium registration (required for all speakers)

CALL FOR PAPERS: The program committee invites talks on all aspects of statistics. Abstracts are due March 31, 2004. Please submit abstracts to: Professor Gang Li, University of California at Los Angeles, email address: vli@ucla.edu. The abstract should include the name, affiliation, mailing address, telephone number, fax number, and e-mail address of the author, and should not exceed 200 words.

ICSA STUDENT AWARDS AND TRAVEL FELLOWSHIPS: The deadline is February 28, 2004 (see a separate page in this issue for detailed information). For further questions, please contact Professor Kung Jong Lui, San Diego State University, kjl@rohan.sdsu.edu

PROGRAM AND SHORT COURSES (for updates, see ICSA website at http://www.icsa.org)

EXECUTIVE COMMITTEE
Co-chairs: Nancy Lo (Nancy.Lo@NOAA.Gov) and Gang Li (vli@ucla.edu).
Secretary: Alice Chu (alice.chu@prosanos.com)
Treasurer: Kathy Chi-Burris (kathy.chi-burris@pfizer.com)

Program Committee:
Gang Li (co-chair), Larry Shen (co-chair), Naihua Duan, Keh Shin Lii, Ying Lu, Kung Jong Lui, Edward Pun, Weng-Kee Wong, Eric Yan, Nancy Lo, and Joey C. D. Lin.

Logistic Committee:
Nancy Lo (Chair), Alice Chu, Kathy Chi-Burris, William Yuan, David Shen, George Yu, John Lee, Thomas Lin, Kung Jong Lui, Xun Lin, Christina Show, Eric Yan, Jenny Han, Joey C. D. Lin, Edward Pun, Ming Ji, Yan Wang, Feng He, and Juanjuan Fan.

ICSA 2004 APPLIED STATISTICS SYMPOSIUM
PRELIMINARY PROGRAM

- **Keynote Speakers (June 7-9, 2004):**
  Bradley Efron, Max H. Stein Professor of Humanities and Sciences, Stanford University; President-Elect, ASA.
  http://www-stat.stanford.edu/people/faculty/efron.html
  George Tiao, W. Allen Wallis Professor of Econometrics and Statistics, University of Chicago, http://gsb.uchicago.edu/fac/george.tiao

- **Plenary Sessions (June 7-9, 2004):**
  New Developments of Nonparametric Methods in Financial Economics, by Jianqing Fan

  Current Statistical Issues in Clinical Trials for Drug Development, by Tze Leung Lai

- **Banquet Speaker: Arlene S. Ash, Boston University.**

- **Short Courses (Sunday, June 6, 2004, 9:00AM-5:00PM. See later pages for details)**

	Topic	Instructor
1	Practical Guidance of Generalized Linear Mixed Models	Charles E. McCulloch, University of California, San Francisco
2	Tutorial on Statistical Bioinformatics	Charles (Chip) Lawrence, Wadsworth Center and Rensselaer Polytechnic Institute
3	Cancer Trials for Practitioners – Experimental Design and Efficacy Analysis	Kao-Tai Tsai, Aventis Pharmaceuticals
4	Bootstrap Methods: A Guide for Practitioners	Michael R. Chernick, Novo Nordisk Pharmaceuticals
5	Active Controlled Clinical Trials	Yi Tsong & Sue-Jane Wang, FDA

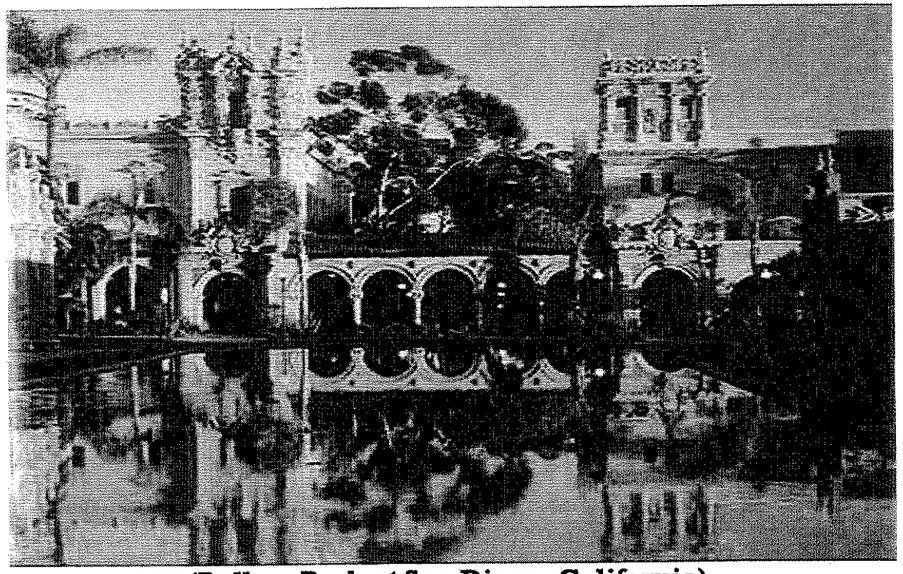
- **Invited Sessions (June 7-9, 2004, subject to change)**

	Session Topic	Organizer	Speakers
1	Recent Advanced in Spatial Statistics	Jun Zhu	Hao Zhang, Washington State University Ji Meng Loh, Columbia University Li Chen, Montserrat Fuentes and Jerry Davis, North Carolina State University
2	Issues of Active Controlled Clinical Trials	Yi Tsong	Steven Snapinn, Merck Ivan Chan, Merck Yong-Cheng Wang*, Gang Chen, and George Y. Chi, FDA
3	Statistics in Financial Econometrics	Jianqing Fan	Yongmiao Hong, Cornell University Per Mykland, University of Chicago Federico M. Bandi, University of Chicago

	Session Topic	Organizer	Speakers
4	Current Methodologies in Pharmaceutical Statistics	Kerry B. Hafner	Rafe Donahue, Glaxo Smith Kline Jimmy Wang, PRA International Charlie (Guoliang) Cao, Takeda Pharmaceuticals North America
5	Assessment of Measurement Agreement	Richard Runze Li and Annie Qu (Chair)	Lawrence I. K. Lin, Baxter Healthcare Corporation Runze Li and Mosuk Chow, Pennsylvania State University John Lu, National Institute of Standards and Technology
6	Recent Advances in Survival Analysis	Gang Li	Dorota Dabrowska, UCLA Wei Wang, Harvard University Ronghui Xu, Harvard University
7	Data Mining in Chemistry and Chinese Medicine	Kai-Tai Fang	Yizhen Liang, Central South China University, China Kai-Tai Fang, Hong Kong Baptist University, Hong Kong, China Aijun Zhang*, Wai-Yan Ha, Yu-Hui Hu, Ricky N.S. Wong and Kai-Tai Fang, Hong Kong Baptist University, Hong Kong, China
8	New Development in Medical Diagnostic and Screening Tests	Xiaohua Andrew Zhou and Kung-Jong Lui (Chair)	Margaret Pepe, University of Washington Vanja Dukic, University of Chicago Xiao-Hua Andrew Zhou, University of Washignton and VA Puget Sound Health Care System
9	Statistical Methods for AIDS Clinical Research	Hulin Wu	Zhezhen Jin, Columbia University Yangxin Huang* and Hulin Wu, Frontier Science & Technology Research Foundation Lang Wu, University of British Columbia
10	Design of Experiments	Ching-Shui Cheng	Weng Kee Wong, UCLA Hongquan Xu, UCLA Kenny Ye, SUNY at Stony Brook
11	Design and Analysis of Dose Response Studies	Naitee Ting and Tao Wang (Chair)	Douglas M Potter, University of Pittsburgh Cancer Institute Dr. Jason Hsu, Ohio State University Naitee Ting, Pfizer Global Research & Development
12	Statistical Applications in Accounting, Economics and Finance	Ruey S. Tsay	Michael C. Davis* and James Hamilton, Univ. of Missouri and UC San Diego Jimmy Ye, Baruch College Ruey S. Tsay, University of Chicago
13	Empirical Likelihood and Its Applications	Songxi Chen	Songxi Chen, National University of Singapore Gang Li, UCLA Ian McKeague, Florida State University Jin Qin, Memorial Sloan-Kettering Cancer Center

	Session Topic	Organizer	Speakers
14	Computing Intensive Methodologies in Bayesian Statistics	Minghui Chen	Jun Liu, Harvard University Steve MacEachern, Ohio State University Ming-Hui Chen, University of Connecticut
15	Bayesian Inference and Graphic Methods for Complex Data Analysis	Xiaoli Meng	Iain Pardoe, University of Oregon Cavan Reilly, University of Minnesota Xiao-Li Meng, Harvard University
16	New Development in Quality Improvements	Smiley W. Cheng	Gemai Chen and Lingyun Zhang, University of Calgary Maria Tong, Sanford Papermate Corporation Youn-Min Chou, U. of Texas at San Antonio
17	Statistical Methods for the Analysis of DNA and Tissue Microarray Data	Steve Horvath	Xinping Cui, UC Riverside Steve Horvath, UCLA Jim Veitch, Corimbia, Inc.
18	Environmental and Ecological Study	Dongchu Sun	Ling Chen, Florida International University Chong Z. He, U. of Missouri-Columbia Hoon Kim, California State Polytechnic U.
19	Aspects of Clinical Trials	Grace Yang	Jian-Lun Xu and Richard Fagerstrom, Philip Prorok and Barnett Kramer George YH Chi, FDA Kai-Fun Yu and Aiui Liu, NIH
20	Bioengineering and Statistics	Nancy Lo	Yihua Zhao, UCSD John Shyy, UC Riverside Amy Sung, UCSD
21	New Developments in Longitudinal Data Analysis	Annie Qu	Ying Qing Chen, UC Berkeley Wei Pan, University of Minnesota Annie Qu, Oregon State University Peter Song, University of Michigan
22	Adaptive Design for Clinical Trials	JianWen Cai	Yu Shen, MD Anderson Cancer Center Qing Liu, Johnson & Johnson Peter Thall, MD Anderson Cancer Center
23	Sequential Clinical Trials	Peng Huang	Aiyi Liu, NIH H.M. James Hung, FDA Peng Huang, Medical U. of South Carolina Sue Jane Wang, FDA
24	Statistics for Natural Resources	Nancy Lo	Steven K. Thompson, Penn State University Mark Maunder, IATTC Din Chen, Int'l Pacific Halibut Commission
25	Design and Analysis of Cardiovascular Clinical Trials	H.M. James Hung	H.M. James Hung, CDER, FDA Gordon Lan, Aventis Pharmaceuticals Steve Snapinn, Merck Research Lab
26	Biostatistical Research in Mainland China	Ji-Qian Fang and Ying Lu (Chair)	Feng Chen, Nan Jing Medical University Hua Jin, South China Normal University Cai Xia Li, Zhongshang University

	Session Topic	Organizer	Speakers
27	Intensive Computing in Genetic Applications	Frank Shen	Nanxiang Ge and Liu Hong, Aventis Pharmaceutical Co. Lue-Ping Zhao, Fred-Hutchinson Cancer Research Center Heping Zhang, Yale University
28	Better Writing Skills for Success	Weng Kee Wong	Xiao-Li Meng, Havard University Don Ylvisaker, UCLA
29	Applied and Interdisciplinary Research: Case Studies (*Jointly sponsored by the San Diego Chapter of ASA*)	Duane Steffey, Jacinte Jean (Chair)	Joey Lin, SDSU Reena Deutsch, UCSD Mikhail Golovnya, Salford Systems
30	Recent Advances in Drug Discovery and Development	C. F. Jeff Wu and Eric Yan (Chair)	Greg Dyson, U of Michigan Xiaoli Shirley Hou, Merck Research Labs Sergei Leonov, GlaxoSmithkline
31	Statistical Methods with its application tó public health	Chinsan Lee	Chinsan Lee, National Sun Yat-Sen University, Taiwan Huey-miin Hsueh, National Chengchi University, Taiwan Yunchan Chi and Mei-Chi Huang, National Cheng-Kung University, Taiwan



(Balboa Park at San Diego, California)

INTERNATIONAL CHINESE STATISTICAL ASSOCIATION
2004 APPLIED STATISTICS SYMPOSIUM

Registration Form

JUNE 6-9, 2004

San Diego Marriott – La Jolla, 4240 La Jolla Village Dr. La Jolla CA, 92037

Name: (English) _____ (Chinese, if any)_____

Affiliation: _____ Affiliation category: Industry ☐ Gov. ☐ Academic ☐

Mailing Address: _____

Phone: (___)_____ Fax: (___)_____ E-Mail: _____

Speaker: Yes ☐ No ☐ ☐Overhead ☐ Slide ☐ Powerpoint ☐ Other: _____

1. Symposium Registration

Please check the appropriate box:

Membership Type	By May 1, 2004	Check (√)	After May 1, 2004	Check (√)
Regular Member	$140	☐	$160	☐
Regular Nonmember	$180	☐	$200	☐
Student Member	$50	☐	$70	☐
Student Nonmember	$70	☐	$90	☐

2. ICSA 2004 Membership

Nonmember attendees will receive one-year membership: Please print Membership Application Form from http://www.icsa.org and mail it with this Registration Form.

For ICSA member attendees, please renew your membership: Check the appropriate box below:

Annual ICSA regular membership:	$40	☐	Annual ICSA student membership:	$20	☐
Lifetime ICSA permanent membership:	$400	☐			

3. Short Course Registration* (Sunday, June 6 from 9:00 AM to 5:00 PM)

	By April 15, 2004	Check (√)	After April 15, 2004	Check (√)
Non-student	$300	☐	$350	☐
Student	$50	☐	$60	☐

*The short course registration fee includes breakfast, lunch, and coffee breaks.

Please select the short course you would like to attend:

Topic	Instructor	Check (√)
1. Practical Guidance of Generalized Linear Mixed Models	Professor Charles E. McCulloch, *University of California, San Francisco*	☐
2. Tutorial on Statistical Bioinformatics	Professor Jun Liu, *Harvard University*	☐
3. Bootstrap Methods; A guide for Practitioners	Dr. Michael R. Chernick, *Novo Nordisk Pharmaceuticals*	☐
4. Cancer Trials for Practitioners – Experimental Design and Efficacy Analysis	Dr. Kao-Tai Tsai, *Aventis Pharmaceuticals*	☐
5. Active Controlled Clinical Trials	Drs. Yi Tsong and Sue-Jane Wang, *FDA*	☐

Note: A short course may be canceled due to lack of participation. This fee will be fully refunded in such a case.

4 Tour To Tijuana, Rosarito or Ensenada Mexico At Your Leisure

Great opportunities for fun, shopping and sight-seeing in the cities and/or along the beautiful coast of Baja California: Tijuana, Rosarita, and/or Ensenada!! Please make reservation directly with ***CONTACTOURS*** to visit any of the three cities of Mexico during your stay in San Diego. The tour bus will pick you up at the hotel, http://www.contactours.com/tourssightseeing.html or call 619-477-8687. Tourists are required to carry proper IDs: Photo ID's required for U.S. and Canadian citizens. Passport, I-94 Form, Multiple Entry Visa OR Resident Alien Card required for non-U.S. Citizens. Tours to attractions in San Diego and Los Angles are also available.

5. Total Conference Payment

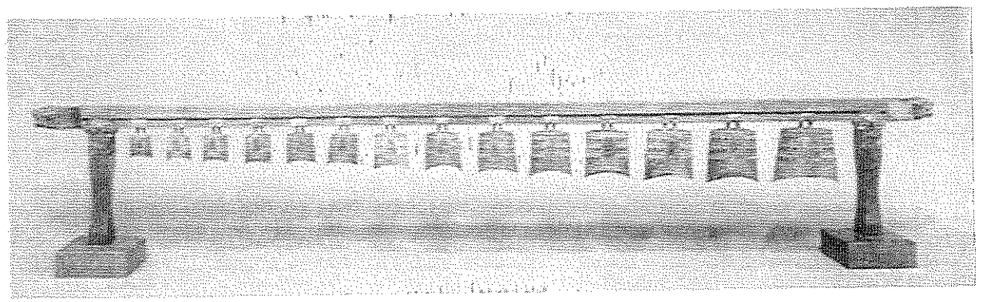
Symposium Registration: = $_____
ICSA Membership: = $_____
Short Course: = $_____
Banquet on Monday night, June 10:
$20/person x ____ people = $_____
$10/child x ____ children (under 10 years old) = $_____
Total for Banquet = $_____
Donation to ICSA: _____ = $_____

Total Payment: $_____
(Please make check payable to 2004 ICSA Symposium)

PLEASE SEND COMPLETED REGISTRATION FORM WITH CHECK TO:

Kathy Chi-Burris
Treasurer, ICSA 2004 Applied Statistics Symposium
c/o Pfizer Global Research & Development La Jolla Laboratories
11085 Torreyana Road
San Diego, CA 92121
Phone: (858) 622-7375
Fax: (858) 678-8248
E-Mail: kathy.chi-burris@pfizer.com

Cancellation Policy: Unless approved by the Committee, all symposium participants must register. Full refund for cancellation will be made if requested on or before May 1; 80% refund will be made if requested after May 1 but on or before May 15, 2004.



ICSA Student Awards & Travel Fellowships

The 13th Annual ICSA Applied Statistics Symposium will be held on June 6-9, 2004 at the San Diego Marriott- La Jolla, La Jolla, CA. The Program Committee will again sponsor the Student Awards and Travel Fellowships. The main purpose of the award is to encourage student members of ICSA to participate and present their research work at this annual meeting.

Qualifications: The student must be an ICSA member (or join at the time of manuscript submission), a degree candidate in any term during 2004 at an accredited institute and be able to register and present the work at the 2004 symposium.

Manuscripts should be prepared double spaced using Biometrics or JASA guidelines for authors. They must be no more than 20 pages in length exclusive of tables and figures. Use one-inch margins and no smaller than 12 point type. The work must be that of the student and be relevant to applications in a variety of fields including biomedicine, business, etc. The manuscript may be co-authored with a faculty adviser and/or a small number of collaborators, although the student must be the first author.

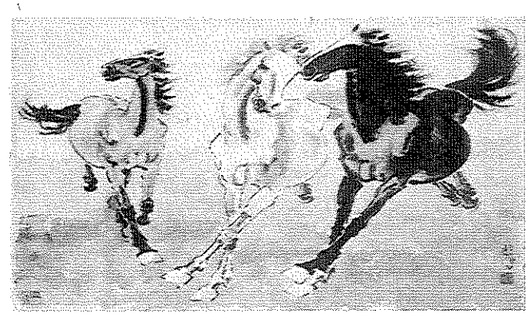
Review and Selection Process: Three review members of the Award Committee, appointed by the Chair of the Committee, will receive blinded copies of the submitted manuscripts from the Committee Chair and review them based on the following criteria: The manuscript should be well motivated by an application relevant to the specific field(s). The methodology developed should be applicable to the motivating problem. Inclusion of an application of the proposed methodology to a particular study will be favorably considered. Clarity of presentation in writing will be considered as well. Up to 3 travel award winners will be selected by the Awards Committee chaired by Prof. Kung-Jong Lui. All winners will each receive a certificate, $400, and tuition for one short course of their choice. The winners will be notified by April 15, 2004.

Submission of Manuscripts: Manuscripts should be received and postmarked no later than February 28, 2004. The submission should include: (1) A cover letter, (2) One complete title page with author(s), institutional affiliation, mailing address, phone/fax numbers and e-mail address, (3) Five copies of the manuscripts with only a title, but no information on authors or affiliation, on the first page, (4) Two copies of abstract, (5) Two copies of the ICSA membership application for non-members.

Membership forms can be downloaded from http://www.icsa.org

Note: All student winners of 2003 and 2004 will be acknowledged at 2004 symposium.

All materials should be mailed to:
Professor Kung-Jong Lui
(kjl@rohan.sdsu.edu)
Department of Mathematics and Statistics
San Diego State University
San Diego, CA 92182-7720



International Chinese Statistical Association

泛華統計協會

Membership Application & Renewal Form

Name	(Last)	(Middle)	(First)
(English)			
(Chinese)			
Address			
Office	Address:		
	City:		
	State:	Zip Code:	Country:
	Email:	Telephone:	FAX:
Home	Address:		
	City:		
	State:	Zip Code:	Country:
	Email:	Telephone:	FAX:
Education			
	Degree:	Year Graduated:	
	University:		
Professional Occupation & Title			
	Occupation:	Title:	
Membership Fees			
	Regular (US$40)		
	Student (US$20)		
	Permanent (US$400)		
	Spouse (50%)		
	Donations		
	Total Amount Paid: US$		
Statistical Area of Interest (circle all applicable):			
	A: Agriculture	B: Business / Economics	
	C: Computing / Graphics	D: Education	
	E: Engineering	F: Health Sciences	
	G: Probability	H: Social Sciences	
	I: Biostatistics	N: Theory & Methodology	
Please Make Check Payable to: I.C.S.A. Mail This Form & Fees to: ICSA c/o Yi Tsong, Ph.D. 13215 Lazy Glen Lane, Herndon, VA 20171			

Calendar of Meetings

July 21-23, 2004 - The Sixth ICSA International Conference
Location: Singapore.
More information available at: Zhiliang Ying zying@stat.columbia.edu.

July 26-30, 2004, IMS Annual Meeting/ 6th Bernoulli World Congress
Location: Barcelona, Spain
http://www.imub.ub.es/events/wc2004/

The program covers a wide range of topics in statistics and probability, presenting recent developments and the state of the art in a variety of modern research topics and in applications such as mathematical finance and statistical bioinformatics. The program include up to twelve Special Invited Lectures given by leading specialists, thirty-five Invited Paper Sessions and a large number of contributed talks. David Nualart is the Chairman of the Organizing Committee and Wilfrid Kendall is the Chairman of the Scientific Committee.

August 4-6, 2004 The Seventh North American New Researchers Conference
Location: York University, Toronto, Canada

The New Researchers' Conference (NRC) is an annual conference sponsored by the Institute of Mathematical Statistics dedicated to fostering interaction between new researchers in the fields of Probability and Statistics. The Seventh NRC will be held at York University in Toronto, August 4-7, 2004, immediately prior to the Joint Statistical Meetings, also to be held in Toronto

The purpose of the conference is to provide a comfortable setting for new researchers to share their research and make connections with their peers in an informal setting. The conference is kept relatively small (50-60 participants) so as to maximize the intellectual and social interaction. All participants are expected to give a short, expository talk or contribute a poster on their research. Anyone who has received a Ph.D. since 1999 is eligible to attend, although priority will be given to first time participants. Abstracts of these talks and posters will appear in the IMS Bulletin.

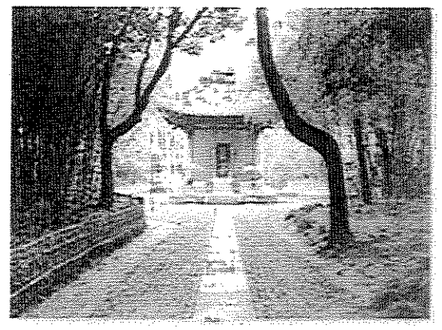
Deadline for receipt of applications is February 15, 2004. Please apply promptly since the number of participants is limited. In addition, we expect to receive funding to defray travel and housing costs for participants. To apply, please submit letter of interest curriculum vitae, and title and abstract of presentation preference for a talk or poster to

Peter Song
Dept of Mathematics and Statistics
N520 Ross Building, 4700 Keele St
York University, Toronto, Ontario, Canada M3J 1P3
Email: song@mathstat.yorku.ca
http://www.math.yorku.ca/StatsSection/NRC

August 6-7, 2004, New Directions in Probability Theory
Location: Fields Institute, Toronto, Canada
http://www.imstat.org/meetings/NDPT/default.htm

The meeting is co-sponsored by the Institute of Mathematical Statistics (IMS) and the Fields Institute for Research in Mathematical Sciences. It will be held at the Fields Institute.

It is intended for a general probability audience interested in recent developments in probability theory. The topics of the session are Random Walks with Self-Repulsion, Random Matrices, Random Media, Super-processes, and Markov Chains with Algorithms. There will be no registration fee for the meeting. However, space at the Fields Institute is limited and early registration is recommended.



Regional Activities

Hailiang Yang, Ph.D.

Hong-Kong Area

Here I report some of the activities of 2003 from the Hong Kong statistical community. I apologize for any missing information.

Bernoulli Society East Asian and Pacific Regional (EAPR) Conference 2003

The Bernoulli Society East Asian and Pacific Regional (EAPR) Conference 2003 will be held at The Hong Kong University of Science and Technology (HKUST) on 18-20 December 2003. The conference is organized by HKUST under the auspices of the East Asian and Pacific Regional Committee of the Bernoulli Society. Keynote speakers are Friedrich Goetze, Zhi-ming Ma, Tze Leung Lai, Kei Takeuchi, Wing Hung Wong and C. F. Jeff Wu.
For more information about the conference, please visit http://www.bm.ust.hk/~eapr2003.

Workshop on Control Theory and Applications

This workshop is to be hold at the University of Hong Kong from December 13-15, 2003. The workshop mainly organized by professor Tze Leung Lai. More than 10 experts from France, USA, Singapore, Mainland China and Hong Kong are invited to given a talk. Professor Tze-Leung Lai will visit the University of Hong Kong as C.V. Starr Professor. For further information about the workshop, please see: http://www.hku.hk/math/

Other Activities and News

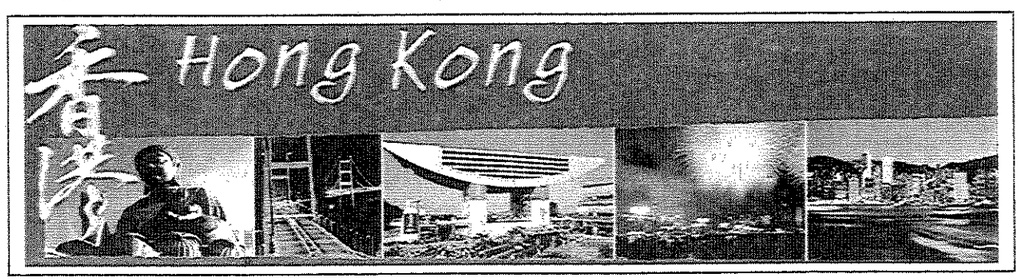
Professor John Aitchison visited The University of Hong Kong in November and December of 2003. Professor Aitchison presented a seminar at the Department of Statistics and Actuarial Science during his visit. He also gave a seminar to the Hong Kong Statistical Society on December 1, 2003.

Professor W.K. Fung was elected as the president of the Hong Kong Statistical Society. Dr. Fung is a professor of the Department of Statistics and Actuarial Science, The University of Hong Kong. Mr. K.C. Leung of Census and Statistics Department was elected as Vice-president.

Professor W.K. Li and Professor W.K. Fung of the Department of Statistics and Actuarial Science, University of Hong Kong were elected Fellow of the American Statistical Association.

Dr. L.X. Zhu of the Department of Statistics and Actuarial Science, University of Hong Kong was elected Fellow of Institute of Mathematical Statistics.

Hailiang Yang, Ph.D. is Associate Professor of the Statistics and Actuarial Science Department, The University of Hong Kong.
Email: hlyang@hkusua.hku.hk



Investigator and Postdoctoral Fellow Positions in Biostatistics and Bioinformatics

National Health Research Institutes (NHRI)
Taiwan, R.O.C.

The National Health Research Institutes (NHRI) is a rapidly growing non-profit organization supported by the Government of R.O.C. **The Division of Biostatistics and Bioinformatics** in the NHRI is seeking outstanding researchers for the positions at the levels of Assistant Investigator, Associate Investigator, and Investigator (equivalent to Assistant Professor, Associate Professor, and Professor in Universities). These are fully funded positions with sufficient opportunities for independent and collaborative research. Postdoctoral Fellow positions are also available. Candidates should have a Ph.D. or equivalent degree in **statistical science**, **epidemiology**, **computational biology**, **computer science** or related fields. Good skills in problem solving, interpersonal communication and coordination will be a plus.

The Division actively engages in NHRI intramural and extramural research design, data management and analysis, and commits to the advancement of biostatistics and bioinformatics research. Current research includes clinical trials, genetic studies and epidemiological studies etc.. Bioinformatics Core Laboratory is established by the Division to facilitate genomic studies. Methodological research focus on the topics which have broad and valid biomedical applications We are establishing a multi-disciplinary group to work together analyzing many interesting data in a variety of studies to elucidate health related scientific problems.

Application should include: A letter of intent, curriculum vitae with publication list, a brief research proposal, reprints of selected publications, and three reference letters sent directly to :

Dr. Chao A. Hsiung, Director
Division of Biostatistics and Bioinformatics
National Health Research Institutes
128 Yen-Chiu-Yuan Road, Sec II, Taipei 115, Taiwan, R.O.C.
Tel: 886-2-2653-4401 ext. 7110
Fax: 886-2-2789-0253
E-mail: hsiung1@nhri.org.tw

NHRI web site - http://www.nhri.org.tw

Guanghua School of Management
Peking University

Faculty position openings
in
Department of Business Statistics and Econometrics

Guanghua School of Management at Peking University, a leading business school in China, has established a new **Department of Business Statistics and Econometrics**. We are conducting a worldwide recruiting for full time faculty at all ranks, junior/senior econometricians and statisticians. Applicants should have a Ph.D. degree in statistics or economics and an established academic record or strong research potential.

The School offers undergraduate, MBA, EMBA, M.S. and Ph.D. programs, with an enrollment of 2,000 students, of whom 1,200 are in the MBA program. Please visit *www.gsm.pku.edu.cn* and *www1.gsm.pku.edu.cn/~stat* for more information.

The starting salary for assistant, associate and full professors are US$40,000, $50,000 and $60,000, respectively (all payable in Chinese currency). In addition to the standard salary and fringe benefits, successful candidates for full-time faculty positions will also receive the following special benefits.

1. Housing benefits. The university provides a two- or three-bedroom apartment; those who wish to buy their own apartment will receive a lump-sum subsidy of 200,000 RMB.
2. Settle-down allowance of 150,000RMB.
3. A start-up research grant of 50,000RMB per year for the first two years.
4. Passage allowance. The School will reimburse one-way economy class airfares for the appointee and his/her immediate family members to fly to Beijing for duty.
5. Flexible sabbatical leave. Sabbatical leaves can be arranged for faculty with well-established records and previous appointments in overseas academic institutions.
6. Publication reward: the current reward for each English publication varies from 10,000RMB to 40,000RMB (depending on journal ranks).

Please contact Professor Rong Chen at (US) 312-996-2323 ( ) for more information. Applicants please send a CV, three reference letters and a research/teaching statement to Prof. Rong Chen, Department of Information and Decision Sciences (MC 294), College of Business Administration, University of Illinois at Chicago, 601 South Morgan Street, Chicago, IL 60607, U. S. A.

New Statistical Programs in China

Zhiliang Ying, Ph.D.

News on the Statistics Program at the Center of Mathematical Sciences (Hangzhou) and the Morningside Center of Mathematics (Beijing)

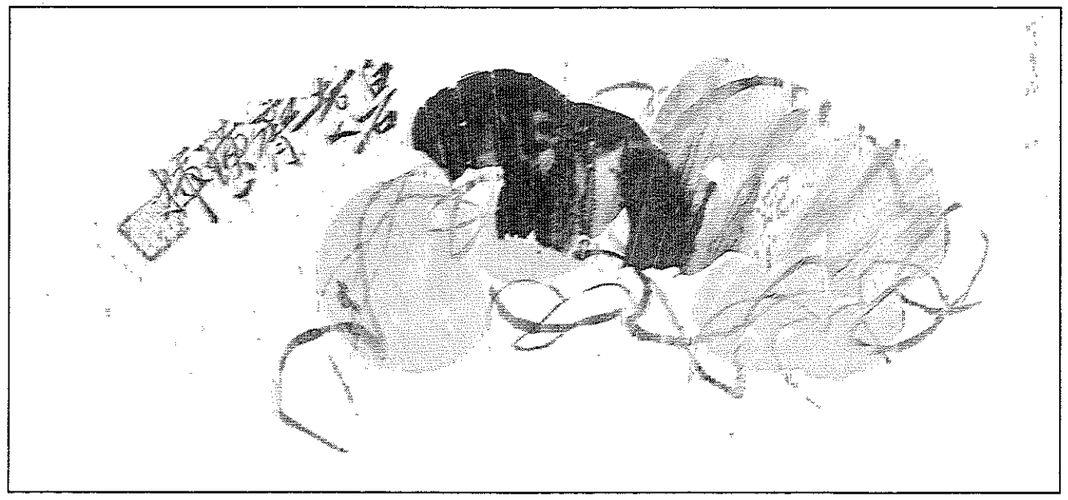
At the suggestion of Professor Shing-Tung Yau, the Morningside Institute of Mathematics of the Chinese Academy of Science and the Center of Mathematical Sciences of Zhejiang University jointly sponsor a regular statistics program. This program is aimed at promoting cutting-edge researches in statistics and its interactions to other areas of science within the statistical/mathematical community. It plans to invite international scholars and experts to give series of lectures on the state-of-the-art researches to young scientists and educators in China.

The first part of Workshop on Financial Statistics was held successfully on December 11-27, 2003 at the Center of Mathematical Sciences in Hangzhou. The principal lecturers were Professor Tze Leung Lai of Stanford University, Professor Ragnar Norberg of London School of Economics, Professor Sam Wong of Hong Kong University of Science and Technology and Dr. Frank Zhang of Morgan-Stanley. Their lectures cover a wide range of mathematical and statistical tools in financial and insurance industries. About 100 mostly young statisticians from all over China attended the Workshop.

The second part of the workshop is scheduled to take place at the Morningside Center of Mathematics in Beijing during the later part of June 2004. It will cover additional topics such as financial engineering, computational methods in finance, statistical inferences for financial models etc.

Also under planning are additional workshops. They include: (1). Nonparametric methods and limit theorems, Hangzhou, winter of 2004-5; (2). Information technology and statistics, Beijing, summer of 2005; (3). Biostatistics, Beijing, summer of 2005.

For more information regarding the program and participation, please contact either Shen, Xiaotong xshen@stat.umn.edu or Ying, Zhiliang zying@stat.columbia.edu.



New Graduate Study Programs in Germany

Prof. Dr. M. Denker & Prof. Dr. A. Munk

The Georg-August-Universität Göttingen has opened a new PhD-program in Applied Statistics and Empirical Methods starting October 1, 2002. The program is offered by the Center for Statistics and is open to students in Agricultural Science, Biology, Economics, Forest Science, Mathematics, Medicine or Social Science with a strong background in statistics. All courses are taught in English. Tuition fees are waived and registration fees are less than 200 Euro annually. Scholarships are available for outstanding candidates only. Applicants should hold a M.Sc. (or equivalent). Good knowledge of the English language and statistical methods are required.

OUR TRAINING PROGRAM

Current topics in statistical methods and data analysis techniques are taught by internationally recognized experts. This includes lectures, tutorials, reading courses, seminars and computer courses. Various Labs and field experiments offer the possibility to train the practical skills of data collection and evaluation.

Statistics courses offered by the ZfS include

- Financial time series analysis
- Smoothing techniques
- Genetic epidemiology
- Pattern recognition
- Resampling methods and algorithms
- Survey sampling
- Spatial statistics
- Survival analysis

The Center for Statistics is constituted by seven faculties which offer additionally a variety of accompanying lectures and courses. This includes the M.Sc. and PhD programs in the International School of Mathematics and Computer Science.

CUTTING EDGE PHD PROJECTS

The three-year PhD period emphasizes independent and interdisciplinary research of our students. It is intended that your PhD project result in a major scientific contribution to your field of specialization. Large interest in development and application of current statistical methods and software is expected. Each PhD student will be advised by three selected faculty members (one serving as the principal advisor) who monitor progress and advice the student in a doctoral project. A PhD degree can be received in any of the accompanying faculties.

CENTER FOR STATISTICS

Courses and research projects are offered by the following researchers and institutions:

- Centre for Social Science Methodology: Steffen Kühnel: linear structural equation models, multivariate models for social science data, survey techniques; Volker Müller-Benedict: modelling and simulation, quantitative empirical methods, analysis of socialstructures
- Department of Genetic Epidemiology: Heike Bickeböller: statistical genetics, epidemiology, and clinical projects.
- Department of Medical Statistics: Edgar Brunner: nonparametric methods, factorial designs, repeated measures; Reinhard Hilgers: design, quality management and analysis of clinical trials, multicenter trials.
- Georg-Elias-Müller Institute of Psychology: Gerd Lüer: human memory and space orientation, research methods,cognitive ergonomics; Thomas Rammsayer: cognitive neuroscience, individual differences, temporal information processing; Michael Waldmann: causality, experimental methods, cognitive psychology.
- Institute for Forest Biometry and Applied Computer Science: Joachim Saborowski: sampling techniques, spatial statistics, regionalization; Branislav Sloboda: computer aided plant modelling, biometric growth models, integrated forest information systems.

- Institute for Statistics and Econometrics: Walter Zucchini: time series, model selection, environmetrics; Jörg Breitung: financial econometrics, analysis of panel data, cointegration
- Institute of Forest Management: Klaus von Gadow: tree growth analysis, forest spatial structures.
- Department of Agricultural Economics: Michael Leserer: econometrics, applied multivariate analysis, applied decision analysis.
- Institute of Mathematical Stochastics: Manfred Denker: nonparametric statistics, nonlinear time series, limit theorems; Heinrich Hering: stochastic analysis, stochastic processes; Ulrich Krengel: optimal stopping, game theory, ergodic theorems; Axel Munk: medical statistics, nonparametric modelling, pattern recognition.

APPLICATION

Applicants are requested to hold an excellent M.Sc. (or equivalent) in biosciences, computer sciences, economics, forest sciences mathematics, physics, psychology, social sciences or statistics. English proficiency must be proved (by TOEFL or IELTS or equivalent).

Applications must be submitted for the winter term (October-February) until June 1, for the summer term (April-July) until January 1.

THIS YEAR'S APPLICATION DEADLINE IS PROLONGED UNTIL JULY 1.

More specifically applicants have to submit the following documents and papers:

1. A signed letter of application to the program.
2. A certified copy of the `Abschlusszeugnis' (MSc diploma). Your final grade must be equivalent to the grade 2.5 at german universities (approximately B- in the american system).
3. A proof of basic knowledge in statistics (e.g. transcripts of records, Übungsscheine,...).
4. A signed statement that no other attempt has been made at a german institution of higher education by the applicant to receive a PhD.
5. A curriculum vita in english or german language, describing the academic career.
6. A statement at which participating faculty the applicant will apply for his PhD degree.
7. Certificate of sufficient knowledge of the English language (TOEFL with 550 points or IELTS with 7 points or equivalent).
8. In addition, the applicants have to provide two letters of recommendation from two professors. These letters should be sent independently.
9. A statement which faculty member is preferred as a principal advisor of the PhD. Alternatively, a statement (e-mail suffices) by a faculty member of the Georg-August Universität Göttingen who agrees to act as a principal advisor for the applicant's PhD dissertation.

Applications have to be submitted to:

Dekanat der Mathematischen Fakultaet
Georg-August-Universitaet Goettingen
Bunsenstr. 3-5
37073 Goettingen Germany.

More detailed information about the procedure can be found on the Web Page of the Center for Statistics, http://www.statistics.uni-goettingen.de.

SELECTION OF CANDIDATES

Applications are reviewed by an admissions committee of scientists participating in the Center for Statistics. Candidates are selected based on their academic qualification, referee's evaluation and personal interviews. Decisions on admission will be made within 2 months after the application deadline.

COSTS, FINANCIAL AIDS & SCHOLARSHIPS

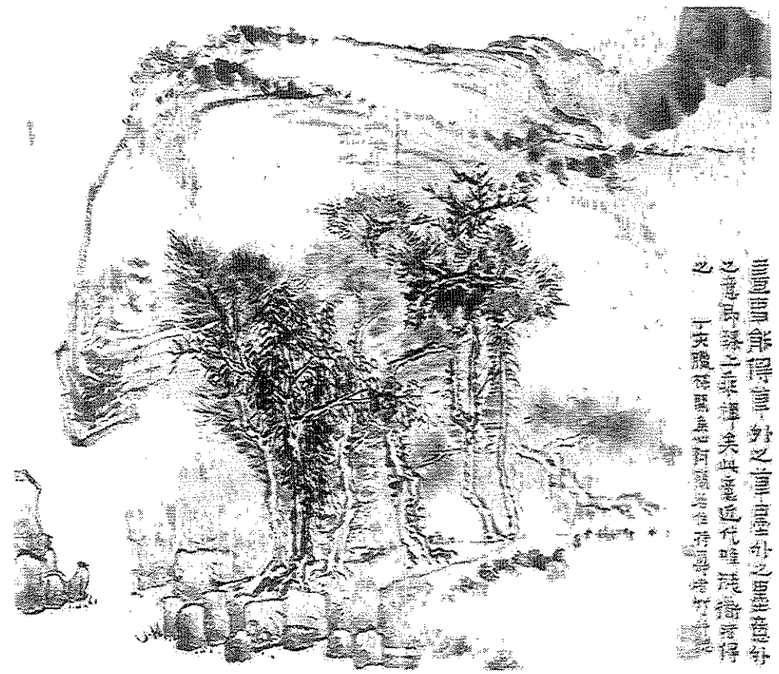
As a general rule, tuition fees at most of the German universities are covered to a large extent by the government. The general administration fee for registration at Göttingen University is less than 200 Euro (180 US$) annually. Living expenses including housing and health insurance ranges from 500-600 Euro per month. All candidates are eligible to apply for a limited number of scholarships. This will be a 3-year funding, about 1,125 Euro per month.

International Chinese Statistical Association
Profit & Loss
January 1, 2003 through December 31, 2003

Ordinary Income/Expense	
Income	
Advertisement	700.00
Banquet at ASA Meeting	4,562.00
Contributions Income	
Membership Dues	7,777.00
Total Income	13,149.00
Expense	
Bank Service Charges	0.40
Casual Labor	139.00
Contributions to ASA	500.00
Banquet at ASA meeting	5,430.00
Internet Registration	262.55
Miscellaneous	469.39
Postage and Delivery	
Announcement	504.76
Ballot	1,164.45
Book/Journal Donation	1,233.00
Bulletin	3,422.67
Statistica Sinica	20.50
Total Postage and Delivery	6,345.38
Printing and Reproduction	
Jan. Bulletin	4,200.00
July Bulletin	4,000.00
Total Printing and Reproduction	8,200.00
Professional Fees for tax filing	320.00
Program Expense	436.88
Supplies	
Other	732.19
Supplies and others	732.19
Web Page Hosting	1,200.00
Total Expense	24,035.79
Net Ordinary Income	-10,886.79
Other Income/Expense	
Interest Income	608.08
Net Other Income	608.08
Net Income	**-10,278.71**

International Chinese Statistical Association
Balance Sheet
January 1, 2003 through December 31, 2003

ASSETS	
Checking/Savings	
Checking	5,208.36
Savings-Money Market	59,078.05
TOTAL ASSETS	**64,286.41**
LIABILITIES & EQUITY	
Equity	
Opening Balance 1/1/2003	74,565.12
Net Income	-10,278.71
Total Equity	64,286.41
TOTAL LIABILITIES & EQUITY	**64,286.41**



Our Sincere Thanks!

The Editorial Team

As in the last issue, many good friends have taken time from their busy schedules to write for the Bulletin.

We especially appreciate the effort of Professor Stephen Lagakos and Professor Mei-Ling Ting Lee for the article about Professor Zelen. It is absolutely fascinating to learn the wonderful achievement of this great man.

We would also like to express our thanks towards Professor T.L. Lai and Professor Ying for arranging the information about the statistical programs in China. This is the testament of the results of great effort from Professor Lai, Professor Yau, and many others.

We would also like to thank Professor Lu and his committee members for providing the information about the ICSA member meeting at the ASA Joint Conference. The dinner and the party afterwards are vividly shown in the pictures of highlights.

Of course, we would not have the collection of the controversial statistical issues regarding the surrogate endpoint, biomarker, & imaging test without the effort of Dr. Wang. These articles represent the state-of-the-art thinking of important issues in the current clinical practices.

As usual, we would like to thank Dr. Lippman for translating the strategy of Suwen-Tse, the great military strategist in ancient China, and the helping hands of other editorial members and volunteers.

We hope we all can follow the paradigms set by these friends and make ICSA the best that it can be.

孫子兵法

知己知彼，
百戰不殆；
不知彼而知己，
一勝一負；
不知彼，不知己，
每戰必敗。

Know yourself and know others;
A hundred wars are no dangers.

Not know others, but know yourself;
One victory, one defeat.

Not know others and not know yourself;
Every war is a disaster.