A Reminiscence of Dr. Margaret C. Wu’s Statistical Innovation at the National Heart, Lung and Blood Institute

Joseph F. Heyse


Impact and Citations

Give Industry a Chance

A Fundamental Link between Statistics and Humor

2019 ICSA Awards

2019 Student Paper Awards and Travel Grants Recipients
From the Editor

Yi Huang

Dear ICSA Members,

It is my great pleasure to welcome you to the 2019 Jul issue of the *ICSA Bulletin*. This issue focuses on “Statistics: Making an Impact”, same theme as JSM2019.

This issue has two featured articles showing the profound impacts of two distinguished statisticians on government statistics and pharmaceutical statistics respectively. The first one is written by Dr. Colin Wu, our new ICSA 2020 President-Elect, highlighting the tremendous influence and innovation by Dr. Margaret C. Wu on modern statistical methodology and data science research conducted at the National Heart, Lung, and Blood Institute. All her work focused on providing simple and scientifically interpretable results for complex problems deeply rooted in real biomedical studies, which was summarized in four areas, such as informative censoring in longitudinal data, tracking and prediction from serial measurements, and etc.

The second feature article is written by Dr. Jie Chen and Dr. Lisa Lupinacci, on Dr. Joseph F. Heyse, one of the most respected statisticians and influential industry leaders “who helped cultivate the culture of statistical research at Merck and in the biopharmaceutical industry” (Jie Chen). Dr. John Tukey said Joe was one of the most creative statisticians he ever met. He had touched many people’s lives and inspired hundreds of statisticians through work and mentoring and collaborations in his life. Luckily, I was one of them. His profound influence in biopharmaceutical statistics covers health economic statistics (cost-effectiveness), vaccine biostatistics (design studies and analysis for vaccine developments, FDA approval, and post-market safety assessment), early development statistics, and more.

This issue features five column articles. Two are in Dr. Yi Tsong’s column “Yi’s FDA Story: Where Statistics Met Regulation”. With more than 30yrs of FDA experience as a statistician and leader, Yi shares his remarkable personal experience and insider stories about important FDA events with a great sense of humor. In 1988 article, Yi shared his personal story of 1988 AIDS ACT-UP demonstration and the critical thoughts on Food and Drug Administration Act and Spontaneous Reporting System. In 1989 article, he shared his personal experience of the safety assessment of Accutane (Tretinoin) using the adverse event reporting system database. Do you know the connection between the book “The Hot Zone” and Reston city in VA and Ebola virus? The 1989 article has the answer. Dr. Hans Rudolf Küensch shared his critical thoughts and inspiring comments on how to think about impact and citations of research publications in his column “Hints from Hans”. Dr. Terry Speed promoted researchers to join the industry and make more direct impact in real world through employment, or research collaborations, or graduate trainings, under his column “Terence’s Stuff”. Dr. Xiao-Li Meng explains a fundamental link between statistics and humor in “XL-Files”. Want to know why a student asked Prof. Casella about standard deviation always being six? Check out this column article, please.

Turning to ICSA business, this issue includes the letter from 2019 ICSA President, Dr. Heping Zhang and the letter from the executive director, Dr. Gang Li, highlighting selective conferences and professional activities of ICSA under their leadership. After ICSA election this July, the results on 2020 President-Elect and directors of ICSA board (2020-2012) are posted here. Additionally, this issue includes recipients of the 2019 ICSA awards; new fellows of ASA and IMS in ICSA family; *Statistics in Biosciences (SIBS)* co-editors’ report; *Statistica Sinica* co-editors’ report; reports from the 2019 program committee; reports from the 28th ICSA Applied Statistical Symposium; 2019 student paper awards and travel grant recipients, and the announcements of multiple upcoming meetings/confereces at the end of this issue, especially the 2020 ICSA Applied Statistics Symposium.

I would like to thank all the authors and contributors, ICSA executives and committees for their strong support and enthusiasm in the *ICSA Bulletin*, and Xiaoyu Cai for putting it together. Enjoy.

Yi Huang, Ph.D.
Editor-in-Chief, ICSA Bulletin
Associate Professor
Department of Mathematics and Statistics
University of Maryland, Baltimore County
From the 2019 President, ICSA

Heping Zhang

Dear ICSA Members,

It has been a great pleasure and honor to serve as the president of the ICSA, and this opportunity enables me to witness and experience the dynamic and exciting activities that have been happening in our community. Being established in 1987, the ICSA now has over 1,500 members including a few hundreds of students who we offer free membership.

This year began with the successful data science conference that took place on January 11-13, 2019, in Xishuangbanna, Yunnan, China. It was co-organized by ICSA, Yunnan Applied Statistical Association, and Shanghai Jiaotong University. I would like to thank Professor Ming Yuan at Columbia University, the program committee that he chaired, and the local committee for their hard work and contribution.

I had the pleasure to attend and speak briefly in the highly successful 2019 Applied Statistics Symposium and 2019 ICSA China Conference. The 2019 Applied Statistics Symposium was held on June 9-12, 2019 at the Raleigh Convention Center in Raleigh, NC, and the organization committee was led by Professor Wenbin Lu, at the North Carolina State University. This Symposium featured three keynote speakers: Marie Davidian at North Carolina State University, Steve Ruberg at Analytix Thinking, and Lilly Yue at U.S. Food and Drug Administration, representing academia, industry, and government agencies. The 2019 ICSA China Conference took place on July 1-July 4, 2019 on the campus of Nankai University in Tianjin, China. This Conference was organized jointly by the ICSA, Nankai University and Shanghai Jiaotong University. Professor Zhezhen Jin at Columbia University chaired the organization committee. Hongyu Zhao at Yale University and Lixin Zhang at Zhejiang University gave keynote lectures on “Statistical Methods for Genetic Risk Prediction” and “Adaptive-randomization—Models and Theory,” respectively. In both meetings, prizes were given to a number of student participants for their excellent research and presentations. I wish to thank Professors Lu and Jin and their committees for the amazing successes.

To promote next-generation leaders in our community, the ICSA established an Outstanding Young Investigator award this year. The ICSA award committee chaired by Professor Yufeng Liu at the University of North Carolina, Chapel Hill, selected Professor Xi Chen at New York University, Xiaohui Chen at University of Illinois Urbana-Champaign, and Gongjun Xu at University of Michigan as the 2019 recipients. In addition, Professor Tony Cai at University of Pennsylvania is the 2019 ICSA Achievement Award winner. I would like to thank Professor Liu for leading the ICSA award committee, and congratulate these outstanding award winners.

The 11th ICSA International Conference will be held at Zhejiang University, Hangzhou, Zhejiang, China, from December 20, 2019 to December 22, 2019. The organizing committee is chaired by Hongzhe Li at University of Pennsylvania. This international conference will feature three keynote speakers Jianqing Fan at Princeton University, Zhiliang Ying at Columbia University, and Hongyu Zhao at Yale University. Professor Fan will deliver the inaugural Peter Hall lecture, and Professor Zhao will be the recipient of 2020 Pao-Li Hsu Award. I am excited and look forward to the conclusion of 2019 on a high note.

The establishment and growth of ICSA would not be possible without the dedication of many volunteers, especially Dr. Gang Li, the ICSA Executive Director, and his successor, Professor Mengling Liu. It has been my privilege and great pleasure to have been working with Dr. Li, and more recently with Professor Liu. The ICSA has been unfortunate to have them leading the daily operation. Also importantly, the ICSA committees and their chairs have done outstanding jobs. For example, Professor Jianqing Fan led the Special Lectures and his committee selected a list of outstanding ICSA members who will deliver special lectures in future ICSA conferences. Professor Annie Qu at University of Illinois Urbana-Champaign chaired the nomination committee that recommended well-qualified candidates for the leadership of the ICSA in the next few years. Professor Peter Song at University of Michigan chairs the Program Committee that coordinates all of the exciting activities related to the ICSA meetings. I also wish to mention that it has been a rewarding experience to serve in the ICSA Executive Committee and have the opportunity to work closely with Professors Tony Cai (a former ICSA president), Aiyi Liu (the ICSA past-president), and

© www.icsa.org
From the ICSA Executives

Tony Sun (the ICSA president-elect). There are too many individuals whom I feel grateful for their contribution to the ICSA, including all editors and those working at the ICSA office.

Thanks to the outstanding and dedicated contribution of our members, ICSA is now a major association in the statistical society. We are strong and thriving, but we also face challenges like every other society. With your support, we will be stronger!

Thank you all!

Heping Zhang, Ph.D.
2019 President, ICSA
Susan Dwight Bliss Professor of Biostatistics
Professor in the Child Study Center and Professor of Statistics and Data Science
Yale School of Public Health
Yale University

From the Executive Director 2017-2019

Gang Li

Dear ICSA members,

We have a very productive 2019—all our planned projects have progressed smoothly. The three ICSA conference were held with great success. The 2019 ICSA Conference on Data Science was held on January 11-13, 2019, at Xishuangbanna, Yunnan, China, with about 100 attendees. The 2019 ICSA Applied Statistics Symposium was held on June 9-12, 2019 at Raleigh, NC. The theme of this conference was Modernizing Statistics for Emerging Data Science, in recognition of a new era for statisticians with different challenges and opportunities with more than 500 attendees. The 2019 ICSA China Conference took place on July 1-4, 2019, on the campus of Nankai University in Tianjin, China. The Conference was organized jointly by ICSA, Nankai University, and Shanghai Jiaotong University. The conference provided a platform for the exchange of recent research and developments in modern statistical methods, to create collaboration opportunities and to identify new directions for further research. We appreciate our organization committee members and partners in China, including Yunnan University, Nankai University, and Shanghai Jiaotong University, for their tireless devotion to the preparation for conferences and collaborations.

The 11th ICSA International Conference is scheduled on December 19-22, 2019, in Hangzhou, China. The ICSA International Conference is the most prestigious conference of ICSA and held every 3 years. The theme of this conference is Innovation with Statistics and Data Science. We envision that this meeting will attract industrial statisticians and many international statisticians as well as statisticians working in government and academia. Details on the conference can be found at http://cds.zju.edu.cn/ICSA2019.aspx?k1=4&k2=79&k3=80.

To support young statisticians in their career development, ICSA gives 1-2 scholars the Young Researcher Awards each year. The ICSA Young Researcher Awards are given to three scholars in 2019 for the first time. The awards will be presented to them at the ICSA Member Meeting at the Joint Statistical Meeting in Denver, CO. Now it is time for you to consider the nomination for the candidates next year.

Let us work together for the continued success of ICSA. Please let me know if you have any ideas or suggestions.

Gang Li, Ph.D.
ICSA Executive Director (2017-2019)
Director, Statistics and Decision Sciences
Janssen Research & Development, LLC
Results of 2019 ICSA Election

2020 President-Elect
Colin Wu (National Institutes of Health)

Directors of ICSA Board (2020-2022)
(alphabetical order)

• Jason Liao (Merck)
• Bin Nan (University of California)
• Peihua Qiu (University of Florida)
• Jane Zhang (Allergan)
• Yichuan Zhao (Georgia University)

2019 ICSA Awards

Distinguished Achievement Award

In recognition of the distinguished achievement in statistical research and unselfish support of the association.

Tony Cai Ph.D., University of Pennsylvania, Philadelphia, PA

For fundamental contributions to high-dimensional statistical inference, adaptive nonparametric function, and minimax optimality, for outstanding mentorship, and for exceptional leadership and service to our profession.

Outstanding Service Award

In recognition and with sincere appreciation for the dedicated effort, unselfish support and outstanding service.

Tony Cai Ph.D., University of Pennsylvania, Philadelphia, PA

For his outstanding services to ICSA including organizing multiple highly successful ICSA conferences including the 10th ICSA International Conference and the 2018 ICSA China Conference, promoting ICSA and its mission in Asia, and dedicated ICSA committee services.

Gang Li Ph.D., University of California at Los Angeles, CA

In recognition of his outstanding service as the ICSA Treasurer 2016-2018, the ICSA Symposium Treasurer 2013-2015, and the local committee Chair of ICSA 2012 Applied Symposium.

Hongliang Shi Blueprint Medicines Inc., Cambridge, MA

In recognition of her dedicated service as the ICSA Treasurer 2016-2018, the ICSA Symposium Treasurer 2013-2015, and the local committee Chair of ICSA 2012 Applied Symposium.

President’s Citation

In grateful appreciation of the generosity, dedication and devoted effort for ICSA.
Chengyong Tang  Ph.D., Temple University, Philadelphia, PA

In recognition and appreciation of his dedicated and outstanding service and leadership as the Chair of the Program Committee for the 2018 ICSA Applied Statistics Symposium.

Min-ge Xie  Ph.D., Rutgers University, Piscataway, NJ

In recognition of his dedicated services to ICSA, including serving as the co-chair of the Organizing Committee for the 2018 ICSA Applied Statistics Symposium.

Xi Chen  Ph.D., New York University, New York, NY

For his impact on bridging between machine learning, statistics, and optimization, in particular, his novel contributions to statistical inference based on stochastic optimization, distributed data inference, sequential learning, and their applications to revenue management, and crowdsourcing.

Outstanding Young Researcher Award

In recognition of the outstanding research in statistical theory, methodology, and/or applications.

Xiaohui Chen  Ph.D., University of Illinois at Urbana-Champaign, IL

For fundamental contributions to high-dimensional statistics, time series analysis and statistical machine learning, in particular to Gaussian approximation to U-statistics and covariance matrix estimation.

Gongjun Xu  Ph.D., University of Michigan, Ann Arbor, MI

For his fundamental contributions to latent-class models and statistical modeling and inference for educational measurements.
New Fellows of ASA

- Huiman X. Barnhart, Professor of Biostatistics, Duke University
- Jinbo Chen, Professor, University of Pennsylvania
- Haoda Fu, Senior Research Adviser, Eli Lilly and Company
- Jianhua Hu, Professor, University of Columbia
- Hongkai Ji, Professor, Johns Hopkins Bloomberg School of Public Health
- Jiashun Jin, Professor, Carnegie Mellon University
- Katerina Kechris, Professor, Colorado School of Public Health
- Jia Li, Professor, Penn State University
- Yehua Li, Professor, University of California at Riverside
- Pei Wang, Professor, Icahn School of Medicine at Mount Sinai
- Min Yang, Professor, University of Illinois at Chicago
- Xiangrong Yin, Professor, University of Kentucky
- Menggang Yu, Professor, University of Wisconsin-Madison
- Lanju Zhang, Director and Research Fellow, AbbVie
- Hui Zou, Professor, University of Minnesota

New Fellows of IMS

Jeng-Min Chiou, Academia Sinica
For contributions to methodology for clustering, classification, and prediction with functional data.

Cynthia Rudin, Duke University
For contributions to interpretable machine learning algorithms, prediction in large scale medical databases, and theoretical properties of ranking algorithms.

Xiaofeng Shao, University of Illinois, Urbana-Champaign
For contributions to non-parametric statistical inference for multivariate time series, in particular to the asymptotic theory for time series analysis via moments and cumulants.

Yuedong Wang, University of California, Santa Barbara
For contributions to non-parametric regression and computational statistics, in particular smoothing spline methodology for dependent observations and applications to bioinformatics and biomedical modeling.

Hongquan Xu, University of California, Los Angeles
For contributions to experimental design, computer experiments, and functional data analysis, in particular to nonregular fractional factorial designs and spacefilling designs.
Report from Statistics in Biosciences (SIBS)

Hongzhe Li and Mei-Cheng Wang

Statistics in Biosciences (SIBS) is one of the two official journals established by ICSA. The journal was established about 10 years ago and has focus on development and application of statistical methods and their interface with other quantitative methods, such as computational and mathematical methods, in biological and life science, health science, and biopharmaceutical and biotechnological science. ICSA members can find more information of the journal from the following website: [http://www.springer.com/statistics/journal/12561?detailsPage=editorialBoard](http://www.springer.com/statistics/journal/12561?detailsPage=editorialBoard)

SIBS publishes both regular articles and topic-oriented papers in Special Issues. The last 5 Special Issues that SIBS published includes ‘SAMSI-Beyond Bioinformatics,’ ‘Statistical Methods in Organ Failure and Transplantation,’ ‘Statistics and Genomics: Emerging Issues and Solutions,’ ‘Statistical methods for clinical trials and precision medicine’ and ‘Challenges in Computational Neuroscience.’ In July 2019 the journal will publish a special issue on ‘Medical Device Data,’ which is guest edited by Jaroslaw Harezlak and Chongzhi Di. As the ICSA has become the 4th largest statistical association in the world, we expect SIBS to grow into a leading journal over time.

Prof. Hongzhe Li/Hongzhe Lee
Professor of Biostatistics and Statistics
Director, Center for Statistics in Big Data
Vice Chair for Integrative Research
Department of Biostatistics, Epidemiology and Informatics
University of Pennsylvania

Mei-Cheng Wang, Ph.D.
Professor
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
Submissions and Acceptance Statistics

In the past 12 months (August 1, 2018 to July 31, 2019), 450 manuscripts were submitted to Statistica Sinica, which include one submission to a special issue and nine comments to a discussion paper (see Table 1). The numbers of manuscripts submitted and accepted for the past six years are shown in Table 2. The submission and acceptance rates during 2015 to 2017 is higher because of the special issues, and the rate in the past 11 months is almost the same as the rate before 2015. The review status for the past three years is displayed in Table 3, and the top 10 countries with the highest submissions for the past three years are shown in Table 4.

Table 1. Manuscript received from August 2018 to July 2019

<table>
<thead>
<tr>
<th>Manuscript Type</th>
<th>Number of Manuscripts</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invited Comment</td>
<td>9</td>
<td>2.0%</td>
</tr>
<tr>
<td>Original Article</td>
<td>440</td>
<td>97.8%</td>
</tr>
<tr>
<td>Special Issue Paper - Big Data in Environmental Studies</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>450</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Submissions and acceptances for the past 6 years.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>95</td>
<td>76</td>
<td>104</td>
<td>150</td>
<td>146</td>
<td>88</td>
</tr>
<tr>
<td>Submission</td>
<td>439</td>
<td>409</td>
<td>567</td>
<td>531</td>
<td>486</td>
<td>450</td>
</tr>
</tbody>
</table>

Table 3. Review status for the past 3 years.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejected w/o external review</td>
<td>204</td>
<td>226</td>
<td>212</td>
</tr>
<tr>
<td>Rejected with external review</td>
<td>103</td>
<td>100</td>
<td>77</td>
</tr>
<tr>
<td>Rejected with revision allowed</td>
<td>81</td>
<td>57</td>
<td>25</td>
</tr>
<tr>
<td>Major/Minor revision</td>
<td>3</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>First submission under review</td>
<td>0</td>
<td>1</td>
<td>73</td>
</tr>
<tr>
<td>Revision under review</td>
<td>0</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Withdrawn</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Accepted</td>
<td>135</td>
<td>92</td>
<td>26</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>531</strong></td>
<td><strong>486</strong></td>
<td><strong>450</strong></td>
</tr>
</tbody>
</table>
Manuscript Processing Time

Table 5 shows the turnaround statistics of initial decision for the past three years, with the decision times censored on July 2, 2019. About 75% of the editorial decisions during 2018-2019 take less than 81 days, but 5% take over 138 days.

<table>
<thead>
<tr>
<th>Period</th>
<th>5th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>95th</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug 2016 – Jul 2017</td>
<td>3</td>
<td>9</td>
<td>30</td>
<td>102</td>
<td>184</td>
<td>526</td>
</tr>
<tr>
<td>Aug 2017 – Jul 2018</td>
<td>3</td>
<td>9</td>
<td>25</td>
<td>94</td>
<td>159</td>
<td>478</td>
</tr>
<tr>
<td>Aug 2018 – Jul 2019</td>
<td>2</td>
<td>9</td>
<td>20</td>
<td>85</td>
<td>154</td>
<td>364*</td>
</tr>
</tbody>
</table>

*Additional 73 manuscripts awaiting initial decision

Backlog for Publication

In the past year, we have published five issues containing 133 articles. There remain 156 accepted manuscripts waiting to be published. Among them, 32 will appear in two special issues in Oct 2017. The backlog is about 16 months from acceptance to publication.

Rankings and Impact Factors

Table 6 shows the ranks of Statistica Sinica based on the 2-Year Impact Factor and the 5-Year Impact Factor provided by the Journal Citation Reports (JCR) in the area of Statistics and Probability from 2010 to 2018. Table 7 shows the ranks of Statistica Sinica in Scimago Journal Rankings among all journals of Statistics and
Probability in the Scopus database from 2010-2018. The ranking is performed using the algorithm Google PageRank.

**Table 6.** JCR rankings for recent 9 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Journals</th>
<th>Ranking (2-Year Impact Factor)</th>
<th>Ranking (5-Year Impact Factor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>123</td>
<td>71 (0.947)</td>
<td>66 (1.256)</td>
</tr>
<tr>
<td>2017</td>
<td>123</td>
<td>71 (0.886)</td>
<td>51 (1.399)</td>
</tr>
<tr>
<td>2016</td>
<td>124</td>
<td>70 (0.899)</td>
<td>46 (1.632)</td>
</tr>
<tr>
<td>2015</td>
<td>123</td>
<td>66 (0.838)</td>
<td>42 (1.611)</td>
</tr>
<tr>
<td>2014</td>
<td>122</td>
<td>44 (1.158)</td>
<td>36 (1.591)</td>
</tr>
<tr>
<td>2013</td>
<td>119</td>
<td>37 (1.226)</td>
<td>44 (1.365)</td>
</tr>
<tr>
<td>2012</td>
<td>117</td>
<td>25 (1.440)</td>
<td>41 (1.418)</td>
</tr>
<tr>
<td>2011</td>
<td>116</td>
<td>49 (1.017)</td>
<td>51 (1.167)</td>
</tr>
<tr>
<td>2010</td>
<td>110</td>
<td>54 (0.956)</td>
<td>59 (1.020)</td>
</tr>
</tbody>
</table>

**Table 7.** SCIImago journal rankings for recent 9 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Number of Journal</th>
<th>Journal Rank</th>
<th>Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>219</td>
<td>41</td>
<td>Q1</td>
</tr>
<tr>
<td>2017</td>
<td>196</td>
<td>23</td>
<td>Q1</td>
</tr>
<tr>
<td>2016</td>
<td>183</td>
<td>26</td>
<td>Q1</td>
</tr>
<tr>
<td>2015</td>
<td>179</td>
<td>20</td>
<td>Q1</td>
</tr>
<tr>
<td>2014</td>
<td>179</td>
<td>14</td>
<td>Q1</td>
</tr>
<tr>
<td>2013</td>
<td>179</td>
<td>12</td>
<td>Q1</td>
</tr>
<tr>
<td>2012</td>
<td>176</td>
<td>19</td>
<td>Q1</td>
</tr>
<tr>
<td>2011</td>
<td>164</td>
<td>23</td>
<td>Q1</td>
</tr>
<tr>
<td>2010</td>
<td>163</td>
<td>36</td>
<td>Q1</td>
</tr>
</tbody>
</table>
Report from the Program Committee

Peter X.K. Song

During the period of January to July, 2019, ICSA has successfully organized two major conferences. One is the ICSA Applied Statistical Symposium held at the Raleigh Convention Center in Raleigh, NC during June 9–12, 2019 and the other is the 2019 ICSA China Conference held in Tianjin, China during July 1–4, 2019.

Chaired by Dr. Wenbin Lu, Professor at the Department of Statistics of North Carolina State University, the Applied Statistical Symposium was themed as “Modernizing Statistics for Emerging Data Science” in recognition of a new era for statisticians with different challenges and opportunities. It attracted more than 400 participants and offered 9 short courses and 103 scientific sessions, including three keynote speeches, one featured session and two student paper award sessions, as well as exciting social events. The second was hosted jointly by ICSA, Nankai University and Shanghai Jiaotong University. It took place on the campus of Nankai University on July 1-4, 2019. Chaired by Dr. Zhezhen Jin (Chair), Professor at the Department of Biostatistics of Columbia University, the scientific program of the conference focused on creating collaboration opportunities and identifying new directions for further research. It also attracted over 400 participants, offered over 80 invited sessions and one large invited poster session, including two plenary speakers, and gave out five Junior Researcher Paper Awards and one Junior Researcher Poster Award. Both conferences were big successes! In addition to timely and high-quality scientific programs assembled by the respective program committees, the success was also built critically upon the great contributions and support of the large numbers of volunteers and co-sponsors.

In addition to the two major conferences above, ICSA has also organized various other activities in the first half of year 2019, including some co-sponsored conferences. Among them, the ICSA Conference on Data Science was held during January 11-13, 2019 in Xishuangbanna, Yunnan, China, the Duke Industrial Statistics Symposium in the theme of Innovative Design and Analysis for Complex Clinical Trials for Drug and Device Developments took place in Durham, NC in April 10-12, 2019, and the 9th IMS-FIPS Workshop was held on June 15-16, 2019 in Shanghai, China.

By the time this comes out, ICSA has already geared up with additional exciting events in the second half of year 2019. The 11th ICSA International Conference will be held Hangzhou, China on December 20-22, 2019. The theme of this conference concerns Innovation with Statistics and Data Science, and the scientific program committee is chaired by Dr. Hongzhe Li. The ICSA-Canada Chapter will hold a conference “Advances and Innovations in Statistics and Data Science” on August 9-11, 2019 in Kingston, ON, Canada. In addition, several ICSA-sponsored international conferences include a Conference on Current Trends in Survey Statistics on August 13-16, 2019 in Singapore, the 6th International Symposium on Biopharmaceutical Statistics (ISBS2019) on August 26-30, 2019 in Kyoto, Japan, and the 11th International Conference on Multiple Comparison Procedures (MCP) on December 12-15, 2019 in Taipei.

While the ICSA program committee is working with several hosts to finalize both scientific programs and logistic matters regarding ICSA’s major symposia and conferences to be held in year 2020-2021, we would like to invite organizations who are interested in holding “2022 ICSA-China conference” and “the 12th ICSA International Conference” to contact the committee chair Peter Song (pxsong@umich.edu). Your support to ICSA is highly appreciated!
Report from the 28th ICSA Applied Statistics Symposium

Wenbin Lu and Donglin Zeng

The 28th ICSA Applied Statistics Symposium was successfully held at Raleigh Convention Center, North Carolina, during June 9-12, 2019. The theme of the conference is Modernizing Statistics for Emerging Data Science, in recognition of a new era for statisticians with different challenges and opportunities. Lilly Yue of U.S. Food and Drug Administration, Stephen Ruberg of Analytix Thinking, and Marie Davidian of North Carolina State University delivered keynote presentations. There were 99 invited sessions, 1 student paper award session and 3 contributed sessions. In addition, there were 9 short courses including 6 half day and 3 full day courses, covering topics from clinical trials, causal inference, precision medicine to network data analysis and deep learning for big data. Over 500 participants working in academia, government, and industry from all over the world attended the conference. Over 150 participants attended the conference banquet. David Banks of Duke University delivered an inspiring banquet speech. The topic of David’s speech was about how Chinese researchers have changed the world of statistics.

The conference received 35 applications for the student paper award competition. Zhou Lan of North Carolina State University, Jitong Lou of University of North Carolina at Chapel Hill, Tao Sun of University of Pittsburgh, Xiangyu Liu of University of Texas Health Science Center at Houston and Jennifer Starling of University of Texas at Austin were selected for the student paper award and Xizhou Guo of University of Michigan and Ann Arbor was selected for the Jiann-Ping Hsu Pharmaceutical and Regulatory Sciences Student Paper Award. The six awardees received the awards during the conference banquet and presented their papers in a special student paper award session. In addition, to encourage more students to participate, the program committee offered six student paper contribution awards so that they can present their works in a contributed session of the conference.

The program committee of the 28th ICSA Applied Statistics Symposium would like to thank our sponsors: AbbVie, Boehringer Ingelheim, FMD K&L, Merck, Pfizer, Sanofi Pasteur, SAS and TechData for the generous support. Special thanks also go to the student volunteers of Duke University, North Carolina State University, University of North Carolina at Chapel Hill for their great efforts and hard work for providing the excellent services to the conference.

Wenbin Lu, PhD
Chair of Organizing Committee
Professor of Statistics
North Carolina State University

Donglin Zeng, PhD
Chair of Scientific Program Committee
Professor of Biostatistics
University of North Carolina at Chapel Hill
2019 Student Paper Awards and Travel Grants Recipients

A student paper award competition was held for the 2019 ICSA Applied Statistics Symposium. The award committee consists of

- Dr. Luo Xiao, Chair, North Carolina State University
- Dr. Gen Li, Columbia University
- Dr. Meng Li, Rice University
- Dr. Suyu Liu, University of Texas MD Anderson Cancer Center
- Dr. Min Qian, Columbia University
- Dr. Rui Song, North Carolina State University
- Dr. Yang Song, Vertex Pharmaceuticals Inc.
- Dr. Yanxun Xu, Johns Hopkins University
- Dr. Lingzhou Xue, Pennsylvania State University
- Dr. Shu Yang, North Carolina State University
- Dr. Jiajia Zhang, University of South Carolina
- Dr. Yichuan Zhao, Georgia State University
- Dr. Yingqi Zhao, Fred Hutchinson Cancer Research Center
- Dr. Ruoqing Zhu, University of Illinois at Urbana-Champaign
- Dr. Yunzhang Zhu, Ohio State University

Congratulations to the following 2019 ICSA Awards Recipients!

ICS A Applied Statistical Symposium Student Paper Awards

- Tao Sun, University of Pittsburgh

Jiann-Ping Hsu Pharmaceutical and Regulatory Sciences Student Paper Award

- Xinzhou Guo, University of Michigan

We heartily thank all of the participants for their contributions. We congratulate the winners for their great achievements. The abstracts of the winning papers are as follows.

Authors: Zhou Lan, Brian J Reich, Joseph Guinness and Dipankar Bandyopadhyay

Title: Geostatistical Modeling of Positive Definite Matrices Using the Spatial Wishart Process

Abstract: Geostatistical modeling for continuous point-referenced data has been extensively applied to neuroimaging because it produces efficient and valid statistical inference. However, diffusion tensor imaging (DTI), a neuroimaging characterizing the brain structure produces a positive definite (p.d.) matrix for each voxel. Current geostatistical modeling has not been extended to p.d. matrices because introducing spatial dependence among positive definite matrices properly is challenging. In this paper, we use the spatial Wishart process, a spatial stochastic process (random field) where each p.d. matrix-variate marginally follows a Wishart distribution, and spatial dependence between random matrices is induced by latent Gaussian processes. This process is valid on an uncountable collection of spatial locations and is almost surely continuous, leading to a reasonable means of modeling spatial dependence. Motivated by a DTI dataset of cocaine users, we propose a spatial matrix-variate regression model based on the spatial Wishart process. A problematic issue is that the spatial Wishart process has no closed-form density function. Hence, we propose approximation methods to obtain a feasible working model. A local likelihood approximation method is also applied to achieve fast computation. The simulation studies and real data analysis demonstrate that the working model produces reliable inference and improved performance compared to other methods.

Authors: Xiangyu Liu, Jing Ning, Xuming He and Ruosha Li

Title: Semiparametric Modelling and Estimation of
the Global Percentile Outcome

Abstract: Biomedical studies are often challenged by the lack of a single primary outcome that can comprehensively capture the multidimensional impairments and symptoms of a disease. To address this, global composite outcomes are commonly formulated to integrate multiple individual outcomes and to achieve comprehensive assessments of the overall disease severity and burden. The global rank-sum outcome and the corresponding K-sample test procedures have been successfully applied in many clinical studies, but existing methods do not lend themselves to regression modeling. In this work, we consider sensible regression strategies to evaluate covariate effects on the global percentile outcome (GPO), under the transformed linear model and the monotonic index model respectively. Possing minimal and realistic assumptions, we develop estimation and inference procedures that account for the special features of the GPO. Asymptotic results were established rigorously using U-statistic and U-process techniques. Simulation studies suggest that the proposed methods perform satisfactorily under realistic sample sizes. An application to a Parkinson’s disease dataset illustrates the practical utilities of the proposed methods.

Authors: Jitong Lou, Yuanjia Wang, Lang Li and Donglin Zeng

Title: Integrative Analysis of Irregularly Measured and Mixed-Type Biomarkers in Electronic Health Records Method with Statistical Guarantees

Abstract: Electronic health records (EHRs) has increasingly become an important data source for personalized medicine. In EHRs, disease biomarkers from the same patient are recorded longitudinally at clinical encounters. In order to comprehensively assess patient’s disease comorbidity and susceptibility, it is necessary to characterize these biomarkers over time in an integrative way. However, there exist some challenges such as, in EHRs, the biomarkers are measured sparsely at irregular and informative clinical encounters. In this paper, we propose multivariate generalized linear models to analyze mixed-type biomarkers, where latent multivariate Gaussian temporal processes are introduced to capture between-marker dependence over time. We allow the covariate effects and covariance matrix of the latent processes to be time-dependent. For inference, we apply kernel-weighted estimating equations based on the method of moments where kernel weights account for the heterogeneous intensity of measurement times. We investigate the asymptotic properties of the derived estimators.

Under the multivariate models, we integrate the irregularly measured biomarkers of mixed modes into composite scores that reflect patients’ underlying health status. We illustrate the finite-sample performance of our method through extensive simulation studies. Lastly, we apply our method to analyze a large sample of EHRs of Type 2 Diabetes (T2D) patients and show its utility to characterize patients’ comorbidity and disease progression while accounting for challenges of EHRs. Keywords: Electronic health records, latent process, kernel smoothing, generalized linear models, method of moments, type 2 diabetes

Authors: Jennifer Starling

Title: BART with Targeted Smoothing: an Analysis of Patient-Specific Stillbirth Risk

Abstract: We introduce BART with Targeted Smoothing, or tsBART, a new Bayesian tree-based model for nonparametric regression. The goal of tsBART is to introduce smoothness over a single target covariate t, while not necessarily requiring smoothness over other covariates x. TsBART is based on the Bayesian Additive Regression Trees (BART) model, an ensemble of regression trees. TsBART extends BART by parameterizing each tree’s terminal nodes with smooth functions of t, rather than independent scalars. Like BART, tsBART captures complex nonlinear relationships and interactions among the predictors. But unlike BART, tsBART guarantees that the response surface will be smooth in the target covariate. This improves interpretability and helps regularize the estimate. After introducing and benchmarking the tsBART model, we apply it to our motivating example: pregnancy outcomes data from the National Center for Health Statistics. Our aim is to provide patient-specific estimates of stillbirth risk across gestational age (t), based on maternal and fetal risk factors (x). Obstetricians expect stillbirth risk to vary smoothly over gestational age, but not necessarily over other covariates, and tsBART has been designed precisely to reflect this structural knowledge. The results of our analysis show the clear superiority of the tsBART model for quantifying stillbirth risk, thereby providing patients and doctors with better information for managing the risk of fetal mortality. All methods described here are implemented in the R package tsbart.

Authors: Tao Sun, Wei Chen and Ying Ding

Title: GWAS-based AMD Progression Using a Copula Semiparametric Model
Abstract: The genome-wide association studies (GWAS) of Age-related Macular Degeneration (AMD), a progressive bilateral eye disease, is the first and most successful GWAS research, where the massive GWAS data provide unprecedented opportunities to study disease risk and progression. This research is motivated by discovering genetic causes and making accurate prediction for AMD progression. For genetic variant identification, we develop a copula-based semiparametric approach for modeling and testing bivariate censored data. Specifically, the joint likelihood is modeled through a two-parameter Archimedean copula, which can flexibly characterize the dependence between two margins. The marginal distributions are modeled through a semiparametric transformation model using sieves, with the proportional hazards or odds model being a special case. We propose a sieve maximum likelihood estimation procedure and develop a generalized score test for testing the regression parameter(s). We apply our method to a genome-wide analysis of AMD progression to identify susceptible risk variants for the disease progression. Lastly, we build a novel GWAS-based survival neural network prediction model for AMD progression. Our results demonstrate how the synergy of wealthy GWAS data and deep learning can effectively predict survival probabilities.

Authors: Xinzhou Guo, Ruosha Li and Xuming He

Title: A Quantitative Assessment of Risk for Subgroup Pursuit in Clinical Trials

Abstract: In clinical studies, when to recommend or decide further pursuit of the most promising subgroup that we have observed from an existing trial is a very important question. It is well recognized that the observed treatment effect size of the best identified subgroup tends to be too optimistic and, therefore, any subgroup pursuit needs to be made with a careful statistical consideration. In this paper, we address the issue of bias in subgroup pursuit and provide a quantitative screening measure that can be used in the decision-making of subgroup pursuit. In addition, we propose a confidence bound on the best subgroup treatment effect from clinical data. The proposed quantitative analysis is model-free, transparent and easy to compute, and can help make a better-informed decision of subgroup pursuit in clinical trials.
A Reminiscence of Dr. Margaret C. Wu’s Statistical Innovation at the National Heart, Lung and Blood Institute

Colin O. Wu

In today’s era of “Big Data” and “Real World Data/Real World Evidence” (RWD/RWE), innovative methods for exploratory statistical analysis have been well recognized as one of the key components of data science—a multidisciplinary field that is becoming increasingly popular across the scientific community. For example, Goal 3 of the NIH Strategic Plan for Data Science is “Support the Development and Dissemination of Advanced Data Management, Analytics, and Visualization Tools,” which includes “Extracting understanding from large-scale or complex biomedical research data requires algorithms, software, models, statistics, visualization tools, and other advanced approaches such as machine learning, deep learning, and artificial intelligence.” (https://datascience.nih.gov/strategicplan) The National Heart, Lung and Blood Institute (NHLBI) has also identified “leveraging emerging opportunities in data science to open new frontiers” as one of the objectives of its Strategic Vision (https://www.nhlbi.nih.gov/about/strategic-vision). As a statistician, it is exciting for me to see that these important objectives have finally garnered the tremendous amount of attention they truly deserve in the biomedical research community. But, I am also constantly reminded that the field of statistics as a critical component of data science for “large-scale or complex biomedical research data” is not new. Many early visionary statisticians have made significant contributions to the development of critical analytical tools that have become the integral parts of today’s data science.

Dr. Margaret C. Wu (not related to me), a biostatistician who spent her career at the NHLBI Office of Biostatistics Research between 1973 and 2001, was one of these early visionary statisticians whose seminal work from the late 1970’s to the early 2000s had tremendous influence on the modern statistical methodology in data science. It is most notable that Dr. Margaret Wu’s methodological contributions were almost entirely motivated by real scientific problems encountered by the project investigators and biomedical researchers. Her statistical results were focused on providing simple and scientifically interpretable solutions with sufficient mathematical rigor. This was a period when many of the complex data structures, practical issues and shortfalls of classical statistical techniques in large biomedical studies were not widely known to the general statistical community. Much of the statistical terminology with which we are so familiar today, such as “error-in-variable regression,” “unbalanced repeated measurements,” “informative censoring,” “nonrandom missing data,” “nonignorable dropout,” and “longitudinal tracking” were non-existent prior to Dr. Margaret Wu’s work. In this respect, Dr. Margaret Wu is a pioneer who spearheaded many of the statistical analysis tools that we take for granted today.

Her work was primarily motivated by real scientific problems encountered by the project investigators and biomedical researchers. Her statistical results were focused on providing simple and scientifically interpretable solutions with sufficient mathematical rigor. This was a period when many of the complex data structures, practical issues and shortfalls of classical statistical techniques in large biomedical studies were not widely known to the general statistical community. Much of the statistical terminology with which we are so familiar today, such as “error-in-variable regression,” “unbalanced repeated measurements,” “informative censoring,” “nonrandom missing data,” “nonignorable dropout,” and “longitudinal tracking” were non-existent prior to Dr. Margaret Wu’s work. In this respect, Dr. Margaret Wu is a pioneer who spearheaded many of the statistical analysis tools that we take for granted today.

I joined NHLBI in early 2002, shortly after Dr. Margaret Wu’s retirement in September 2001. My
Invited Articles

appreciation of Dr. Margaret Wu and her NHLBI colleagues’ innovative work began when I started participating in several high impact NHLBI long-term epidemiological studies. The complexity of practical data and potential mechanisms of biological processes made me understand that seemingly elegant theoretical results could be worthless and soon forgotten if they could not provide meaningful solutions to real problems. The beauty of the trailblazing work by Dr. Margaret Wu and her collaborators rests on its applicability to real practical settings (something close to today’s “Big Data” and “RWD”). The corresponding theoretical developments are all based on practically meaningful assumptions and meant to provide useful insight and interpretations. To make a connection between Dr. Margaret Wu’s publications and potential research questions which we may face today, I briefly comment on the following four areas:

**Informative Censoring in Longitudinal Data.** In a recent special issue paper in Statistics in Medicine, Albert (2019) [1] provided a systematic review of one of Dr. Margaret Wu’s most influential contributions on statistical methods: correcting nonignorable dropout (or nonrandom missing data) bias in longitudinal data analysis with the shared random parameter models. These methods were motivated by the realization that most high impact and long-term biomedical studies have complex repeated measurements, and missing data in these studies often are not “missing at random.” Although most of Dr. Margaret Wu’s publications in this area appeared in the 1980’s and 1990’s [3, 8, 9, 10, 11, 12, 13, 14], the methods developed in these papers are still the standard approaches today for large studies with complex longitudinal data structures. One reason that Dr. Margaret Wu and her colleagues’ methodological contributions have stood the test of time is that these methods are interpretable and tractable. Indeed, these methods can provide useful insights into the underlying relationships between longitudinal variables with complex data structures. Moreover, they rely on tractable mathematical structures which are computationally feasible. Although some of the “Big Data” and “RWD” problems we are facing today may be more complicated than the data structures discussed in Dr. Margaret Wu’s publications, such as the potential selection bias caused by incorporating datasets from multiple different longitudinal studies, I believe that the basic framework developed in her work lays the foundation for future research in developing innovative “RWE” analytical methods.

**Tracking and Prediction from Serial Measurements.** A main advantage of a long-term longitudinal study over a simpler cross-sectional study is that, with repeated measurements over time, it is possible to estimate and predict a subject’s specific outcomes in the future without solely resorting to the population averages. This type of “subject-specific prediction” (or simply termed “tracking”) is particularly useful, for example, in pediatric studies, where it is unlikely to observe hard time-to-event health outcomes, such as death or hospitalization. In Dr. Margaret Wu’s work with Professor James H. Ware in 1981 [5], a practical solution for individual outcome tracking and prediction is provided using the now well-known linear mixed effects model. Unfortunately, it took several years for the original work of Ware and Wu (1981) to attract widespread attention. We now recognize its potential for mining big longitudinal data or electronic health record data for health status and disease risk predictions.

**Clinical Trial Designs with Repeated Measurements.** Since almost every major clinical trial conducted at NHLBI has some repeated measurements, designing clinical trials with repeatedly measured outcomes is also a major theme in Dr. Margaret Wu’s research in the 1980’s and 1990’s [6, 7, 15]. These are all clever trial designs because they take into account the challenges of subject withdrawal, staggered entry, informative censoring, and sequential monitoring. In my opinion, these ideas for trial designs were far ahead of their time, because the concepts of informative censoring and other potential complications were not well understood by many clinical trial practitioners in the 1980s and 1990s. However, the push for “RWD/RWE” and pragmatic clinical trials demands pragmatic trial designs which take into account these challenges. I believe these early ideas of Dr. Margaret Wu and her colleagues will continue to be quite useful for future clinical trial designs.

**Influential Statistical Practice.** In addition to her many methodological contributions, Dr. Margaret Wu’s work on statistical practice has also motivated many influential statistical methods developed by other well-known statisticians. A notable example is Halperin, Wu and Gordon (1979) [4] in which the authors discovered through an application to the famous Framingham Heart Study that the differences of baseline risk variables between cases and non-cases in a cohort study should be evaluated through an appropriate mathematical model. This finding was noted in Carroll et al (1984) [2] and served as a precursor for developing the well-known “error-in-variable regression,” which is now a major branch
of statistics that is widely used in real applications.

Although it is impossible to comment on the full impact of Dr. Margaret Wu and her colleagues’ statistical research, I hope that this essay serves as an appreciation of their trailblazing work. Many of their innovative ideas were ahead of their time and provide creative insight into current statistical research and data science. We are indebted to their efforts.

References


Acknowledgement: The views expressed in this essay are those the author’s and do not necessarily represent the views of the NHLBI, the National Institutes of Health, or the US Department of Health and Human Services. I would like to thank Drs. Nancy L. Geller, Eric Leifer, Xin Tian, James Troendle and Myron Waclawiw for comments and suggestions.
Invited Articles

Colin O. Wu, Ph.D.
Office of Biostatistics Research,
National Heart, Lung and Blood Institute,
National Institutes of Health

Joseph F. Heyse
Jie Chen and Lisa Lupinacci

A veteran statistician who helped cultivate the culture of statistical research at Merck and in the biopharmaceutical industry

Joseph F. Heyse, one of the most respected and influential leaders in biopharmaceutical statistics, passed away Friday, May 31, 2019. He was 67.

Joe joined Merck & Co., Inc., in 1976 as a statistician supporting preclinical research and immediately began to have a profound influence on the statistics profession by developing statistical methods that subsequently became industry standards.

In 1987, Joe began supporting clinical pharmaceutical research, providing technical oversight for health economics studies to support the marketing of Merck products. In 1990, he was named the director of a newly established department of Health Economic Statistics, which furnished cost-effectiveness evaluations of all Merck products, and proceeded to become a well-recognized leader in the field of health economic statistics.

In 1993, Joe assumed responsibility for the Clinical Biostatistics group at Merck’s Pennsylvania site, including Vaccine Biostatistics which was undergoing unprecedented growth. His impact on human health in this role is best summarized by Senior Vice President of Medical Affairs, Eliav Barr, who said, “Joe led the [clinical] statistics group in an era of extraordinary challenges and productivity at Merck. At that time, Merck developed vaccines against rotavirus, a leading cause of infant morbidity and mortality; human papillomavirus, which causes most cases of cervical, genital, and head/neck cancer; and varicella zoster virus, which causes chickenpox in children and often debilitating herpes zoster in adults. Each of these programs required novel, large, complicated studies to address key efficacy and safety questions. Joe’s innovative spirit, keen understanding of clinical research and the underlying disease, deep font of statistics knowledge, and can-do attitude was instrumental in the success of these studies and the subsequent availability of these vaccines. In short, his work was instrumental to the availability of vaccines that have saved hundreds of thousands of lives and prevented untold misery.”

In 2009, Joe transitioned to the head of the Early Development Statistics group at Merck, and most recently, he was in his fifth year as a Scientific Associate Vice President and Head of the Biostatistics Methodology Research group which provides statistical support to research design, analysis and reporting relating to product development through the application of existing and innovative statistical methods.

Throughout his career, Joe published books and peer-reviewed journal papers extensively. John Tukey, one of the most distinguished individuals in the field of statistics, and one of Joe’s earliest mentors, once noted that Joe was one of the most creative statisticians he ever met.

Joe received his M.S. in statistics from Villanova University in 1975 and an MBA in Economics from Temple University in 1979. He was selected...
to participate in the Merck Research Laboratories Doctoral Program in 1984 and received his Ph.D. in statistics from Temple University the following year. He was an elected Fellow of the American Statistical Association (ASA) in 1997 and of the American Association for the Advancement of Science (AAAS) in 2006. Joe was the founding editor of Statistics in Biopharmaceutical Research and editor of Statistical Methods in Medical Research. He served as Vice President and President of Philadelphia Chapter in 1988-1989 and Biopharmaceutical Section Program Chair of ASA in 1994-1995. Joe also served on doctoral dissertation advisory committees at Temple University and the University of Maryland Baltimore County (UMBC). “I still remember his encouraging words and strong support to junior researchers ... he always encourages me to pursue my dream, don’t wait. He is such an inspiring leader to work with and a role model for us to follow.” written by Yi Huang, a professor of statistics at UMBC.

In his 42 years at Merck, Joe had an extraordinary impact on human health. “Joe not only directly contributed to the successful development of dozens of pharmaceutical products, but he personally and proudly developed, mentored, coached and collaborated with hundreds of statisticians who have, in turn, become successful drug and vaccine developers and leaders in their profession both at Merck and in the external statistics community,” said Lisa Lupinacci, Associate Vice President of Late Development Statistics at Merck. “Joe was indeed the catalyst for my career at Merck [and] his encouragement and counsel for so many years have directly contributed to my professional and personal achievements. I will certainly miss Joe’s generosity, coaching, inspirational stories, candor, integrity, wisdom and humor.” said Amy Gillespie, Associate Vice President for Statistical Programming at Merck.

Keaven Anderson, a distinguished scientist at Merck, recalled his interactions with Joe by noting, “His thinking always drew us into a future vision that we could not imagine on our own! I have so much to thank him for.”

Joe will be remembered as a humble and kind person, who was quicker to give credit than accept it and was most willing to share his time and talents to help mentor others along the way.

Joe will be dearly missed by his wife Lil, their daughters Angelina and Gabby, and his many friends and colleagues at Merck and in the biopharmaceutical statistics community around the world.

Jie Chen, Ph.D.
Distinguished Scientist
Biostatistics and Research Decision Sciences
Merck Research Laboratories
Merck & Co., Inc.

Lisa Lupinacci, Ph.D.
AVP, Biostatistics
Late Development Statistics
Merck Research Laboratories
Merck & Co., Inc.
1988

Yi Tsong

*This article reflects the review of the author and should not be construed to represent FDA's views or policies.

Settling in Northern Virginia:
As I mentioned in my 1987 report, when I joined FDA, I moved to Rockville, Maryland, alone in March. By the summer of 1987, my family moved to join me. The four months between March and July was the time window for me to get a basic idea what the housing market looked like. I checked the housing in Rockville, Maryland, and Reston/Herndon in Virginia. They are about the same distance from Interstate 495 beltway. Meanwhile, my wife Patricia was approved to transfer as an employee to Boeing office in Reston, Virginia. Boeing rented a hotel room at Wolftrap hotel in Vienna, Virginia. After selling the house in Houston, my parents-in-law took Stephanie, then four years old, to visit Pat’s sister living in Detroit. Pat brought daughter Jennifer, then eight years old, to Virginia to join me. When they arrived in Dulles Airport, my housemate, Dr. Kao-Tai Tsai was kind enough to join me to pick up and escort them to the hotel. It was the first time we drove on this part of the Virginia in dark night. We prepared a map and decided to take toll road exit from Wiehle Avenue south to Sunrise Valley road. Driving east it would be Hunter Mill Road leading to south. It connected to Lawyers Road and Maple Avenue which was the main street of Vienna, Virginia. But little did we know that part of the Hunter Mill-Lawyers Road route was so hilly like a narrow roller coaster ride without lighting. All four of us were driving fearfully for the 10 minutes distance before reaching the hotel. To ease myself and my family, I sang the Bobby McFerrin’s 1988 popular song “don’t worry, be happy” song with words like,

“Don’t worry, be happy
In every life we have some trouble
But when you worry you make it double
Don’t worry, be happy, …”

https://www.youtube.com/watch?v=d-diB65scQU

In later years, I often drove on the same roads with much more comfort. But I did blow a tire once by the sharp edge of the road.

A few days after checking in at the WolfTrap Hotel in Vienna, we moved to an apartment in Reston for six months. It was time for Stephanie to come home to join us. I still remember the scene of Stephanie coming out of the tunnel bridge from the airplane and seeing her sister Jennifer at the entrance. She was so happy, opened her arms, giggling loud while running to meet Jennifer. Jennifer was equally happy to run forward to catch her. Without a camera or photo, I had the scene carved in my mind of the love between the two sisters.

By March 1988, the home in Herndon Virginia was built. We finally moved into the home. It is hard to believe that Pat and I raised the daughters and lived there for more than 32 years until both of us retired.

History of Reston (taken from Wikipedia):
Reston is one of the leading “New Town” planned communities in the United States. Founded in 1964, Reston was influenced by the Garden City movement that emphasized planned, self-contained communities that intermingled green space, residential neighborhoods, and commercial development. The intent of Reston’s founder, Robert E. Simon, was to build a town that would revolutionize post—World War II concepts of land use and residential/corporate development in suburban America. Reston was frequently (2012, 2014, 2016, 2017 and 2018) ranked in the list of “Best Places to Live in America” by Money magazine for its expanses of parks, lakes, golf courses, and bridle paths as well as the numerous shopping and dining opportunities in Reston Town Center. Beginning in 2017, however, high-density commercial and residential developments along the Dulles Toll Road began to spark concerns among residents about local government’s ability to ensure that key infrastructure, including roads, schools, and parks, would remain in sync with the accelerating pace of new construction. In the early days of Colonial America, the land on which Reston sits was part of the Northern Neck Proprietary, a vast grant by King Charles II to Lord Thomas Fairfax that extended from the Potomac River to the Rappahannock. The property remained in the Fairfax family until they sold it in 1852. Carl A. Wiehle and William Dunn bought 6,449 acres in northern Fairfax County along the Washington & Old Dominion (W&OD) Railroad line in 1886, later dividing the land between them, with Wiehle retaining the acreage north of the railroad line. Wiehle envisioned founding a town on the property, including a hotel, parks, and community center, but completed only a handful of homes before his death in 1901.

The apartment we lived in Reston was located close to Lake Ann, one of the two lakes received wa-
ter from Potomac River. Though the lake is small, it has a very peaceful scenery. I drove on the road over the lake to Dulles toll road to work every morning. When I drove over the lake I watched the water vaporizing on the surface of Lake Ann, it gave me a sense of calm and happiness feeling.

1988 Summer Olympic Games:
By the summer, the 1988 Summer Olympics were held in Seoul, South Korea. This was the last Olympic Games for Soviet Union and East Germany before they ceased to exist. This is the year Florence Griffith Joyner set the Olympic record in the 100-metre and 200-metre dashes and received both gold medals. Furthermore, she added a gold in the 4x100 relay and a silver in the 4x400 relay. After losing consecutively the gold medals in years, 1988 Olympic is the last year that the US Olympic basketball team was represented by only non-professional players. This is the Olympic Game that Seoul Olympic Committee produced and distributed an official song to publicize the Games. The song “Hand in Hand” was written by Italian composer Giorgio Moroder and American songwriter Tom Whitlock. It was with lyrics like “Hand in hand we stand
All across the land
We can make this world
A better place in which to live,...”

However, USA Team also has its own Olympic song sang by Whitney Houston, called “One moment in time” with lyrics like “Give me one moment in time
When I’m more than I thought I could be
When all of my dreams are a heartbeat away
And the answers are all up to me
Give me one moment in time
When I’m racing with destiny
Then in that one moment of time
I will feel
I will feel eternity.”

“One moment in time” became Houston’s third song reached UK number-1 hit and second number-1 song in German hits. It reached number 5 of the US Billboard chart. It further became her all-time best performance at the 1989 Grammy ceremony.

Food and Drug Administration Act and Spontaneous Reporting System:
A sponsor seeking to market a new drug product is required by law to have an application (New Drug Application) reviewed and approved by FDA. The principal requirement of an NDA is that the safety and efficacy of the drug is supported by data collected through phase I to III trials. Often the clinical trial comprises most patients with age lower than sixty and no complication of other disease. The trial is typically designed with a short study duration for adverse event since many adverse reactions has a incidence rate of lower than 0.1% with long exposure. Once the drug is marketed, the same drug product is available to a patient population that is broader and not well studied. The drug may be used for indications of different disease states; for longer duration; patients, such as elderly, children and women not studied in pre-marketing studies; in patients exposed also to other treatments; used off-label (use for treatment not stated in the label), etc. Postmarket adverse drug event (ADE) surveillance is designed for exactly the purpose of identifying the low frequency reactions that are not caught in pre-marketing; the high-risk group; the drug-drug interaction; the long-term effect of the drug; increased severity and/or frequency of known reactions.

FDA started the Spontaneous Reporting System (SRS) in 1968 for collecting, processing and analyzing adverse drug event (ADE) reports. Most of them were reported on the ADE reporting form designed by FDA (Form 3500 which was later replaced by the MedWatch Form in 1992) and submitted by manufacturers of prescription drugs as a condition for marketing in the United States, while others are submitted voluntarily by concerned health professionals and consumers. Each report contains: patient’s demographic information including age and gender; adverse event information, including date and outcome of event; description of event, evidence, and existing medical history; information of suspect and concomitant medicine(s), including drug name, dose and frequency, diagnosis for use, whether the event abated after drug use was stopped, and whether the event reappeared after the drug was reintroduced; and the reporter’s name and address. Each report was reviewed by a trained professional in the FDA to assure that the form is complete and the information is correct before it was entered into the computer database. Together, these reports in the FDA database make up the Spontaneous Reporting System (SRS) of the United States. In 1997, FDA overhauled the postmarket adverse reaction reporting system in order to accommodate electronic submission and unify the coding with MedDRA (a dictionary of adverse drug reactions) and renamed the system AERS (Adverse Event Reporting System). In comparison to SRS,
the post-1996 AERS is more structured, with more capacity, and uses more detailed coding terminologies. The FDA now receives over 300,000 adverse drug event reports annually, and currently has over 1.6 million reports in its computerized database. It forms a major database for postmarketing adverse drug event monitoring and screening.

However, the information in the AERS database is rich but cannot be used for computation of incidence rates of specific adverse event and drug combinations. The difficulties result from the following four reasons:

1. The method of ascertainment of an adverse event is not uniform among the reporters and non-reporters, and the reporters have their own criteria for identifying an AE. It is generally believed that the number of reports is much smaller than the actual number of events.

2. Although all the reports are reviewed carefully in the FDA, the quality of the information collected for a particular adverse event varies with the reporter.

3. The population at risk is unknown, hence the number of reports is not necessarily closely related to the number of individuals taking the drug.

4. Although there are courses offered in medical schools and other medical institutes for all the students, it is, however, not mandatory to report ADEs by the US medical laws. This is different from ADR reporting systems in countries with social welfare systems which requires reporting by every healthcare handler in the government system.

In spite of these limitations, AERS is a valuable database to be used for signaling the possibility of excessive or outlying ADE and for generating hypotheses.

Epidemiological and statistical applications were often used in the signaling process. The drug-adverse event (ADE) signaling procedures used in United States have evolved in the last three decades.

As part of my work on postmarketing safety assessment, I was responsible for development and modified statistical methods for using the SRS (AERS) and other database such as the Medicaid database and third party reporting systems (including most of the insurance company systems). Details of the many developments will be documented in a journal of the later years.

1988 AIDS ACT-UP demonstration:

1986: One by one, we strolled into the conference room of The University of Texas Medical Branch and took our seats. Finally, Professor Tom Branowski, the professor in Family Medicine and the principal investigator of this research team of Family Health Project walked in with his assistant Ms. Henske. Professor Branowski looked so solemn and stood by the table without sitting down and finally said with a trembling voice “Paul won’t be able to come. He checked in the hospital after diagnosed for infection of HIV virus.” I sucked in my breath and looked around and saw five faces displayed with shock. This was an August morning of 1986. Paul Hook was a professor of sociology and the co-principal investigator of this research project. The objective of Family Health Project is two-fold. The first is to investigate the health status and awareness of the black families in Galveston Island and southern Texas. Typically, those families were in the ghetto areas. Paul was a tall black and attractive professor. His responsibility was to visit and interview the families and collect all the data through interview and questionnaire survey. Those were the years when HIV meant the same as AIDS and unavoidable death.

Six months later, I left Houston and moved to Rockville, Maryland and joined FDA. Two more months later, I received a package of the reprints of the papers published and a letter informed me that Paul was gone.

1988: is the year when the incidence of AIDS reached its high point. In the summer, FDA increased the size of the AIDS drug review team trying to speed up the AIDS drug review process. On October 11, those of us working at the FDA Parklawn building in Rockville were notified that there would be a demonstration of AIDS patients and friends. We were strongly advised to avoid any confrontation with the demonstrators. We should be cautious if yelled at when entering and leaving the office. Since it was likely to have obstacles blocking the entrance of the building, we were allowed to turn back home instead of forcing our way in. There were actually about a few hundred demonstrators who showed up in an orderly demonstration in the morning about 8:00 o’clock. We learned later on that the number increased to more than 2,000. The protesting about the concerns about the length of the drug approval process was brought to the center in the AIDS epidemic. In the mid- and late 1980s, ACT-UP and other HIV activist organizations accused the FDA of unnecessarily delaying the approval of medications to fight HIV and opportunis-
tic infections. “Shame! Shame! Shame!” and “No more deaths!” were yelled by the protesters. This protest was led by John Thomas, head of the AIDS Resource Center in Dallas. He claimed “I’m here because the FDA is holding up drugs that are available, and because people are dying.” John Thomas spoke through a loudspeaker as police motorcycles circled the building and crowd behind him.

“I am here today – we are all here today – because we all have AIDS,” Thomas said. “Some of us have AIDS in our bloodstream. And some of us have AIDS in our minds. We look into the mirror and see a sore that won’t go away, and we are fearful that we are going to be diagnosed. And we all have AIDS in our hearts,” he said. “All of us have lost people we love.”

This protest resulted in 176 arrests.

In August 1990, Dr. Louis Lasagna, then chairman of a presidential advisory panel on drug approval, estimated that thousands of lives were lost each year due to delays in approval and marketing of drugs for cancer and AIDS. Partly in response to these criticisms, the National Institute of Health and FDA issued in 1992, new rules to expedite approval of drugs for life-threatening diseases, and expanded pre-approval access to drugs for patients with limited treatment options. The first of these new rules was the “IND exemption” or “treatment IND” rule, which allowed expanded access to a drug undergoing phase II or III trials (or in extraordinary cases even earlier) if it potentially represented a safer or better alternative to treatments currently available for terminal or serious illness. A second new rule, the “Parallel Track Policy,” allowed a drug company to set up a mechanism for access to a new potentially lifesaving drug by patients who for various reasons would be unable to participate in ongoing clinical trials. The “parallel track” designation could be made at the time of IND submission. The accelerated approval rules were further expanded and codified in 1992.

All of the initial drugs approved for the treatment of HIV/AIDS in the early years were approved through accelerated approval mechanisms. For example, a “treatment IND” was issued for the first HIV drug, AZT, in 1985, and approval was granted just two years later in 1987. Three of the first five drugs targeting HIV were approved in the United States before they were approved in any other country.

John Thomas attended as a consumer representative in every advisory committee meeting of HIV/AIDS new drugs. He made presentations and in tears asked the committee to approve the new drug because of its potential usage. In 1997, I saw him the last time at the advisory committee meeting. He made a presentation in tears and cried, “I pleaded at every AIDS AC meeting for the approval of new treatments for HIV. But all I saw in the last ten years, the approval process made all sponsors much richer but none helped my friends or me.” John dies of AIDS in 1999.
1989

Yi Tsong

*This article reflects the review of the author and should not be construed to represent FDA’s views or policies.

The Accutane Case:

Each federal agency may recruit professional experts to serve as a special government employee (SGE) who may be consulted on special issues, provides service on some special projects or serves on an advisory committee. When FDA has any issues regarding the drug efficacy, safety or manufacturing or quality control, members of the advisory committee may be invited to participate in a public meeting to listen and discuss the issues. The discussion and recommendation of the committee will be used to help the agency to determine what action to take. My first experience of an advisory committee meeting was about the safety of a drug called Accutane (Tretinoin) for which I was involved with post-marketing safety assessment using the AERS (Adverse Event Reporting System) database.

Accutane is a medication used for the treatment of acne and acute promyelocytic leukemia. For acne it is applied to the skin as a cream or ointment. For leukemia it is taken by mouth for up to three months. Its common side effects when used by mouth include shortness of breath, headache, numbness, depression, skin dryness, itchiness, hair loss, vomiting, muscle pains, and vision changes. Other severe side effects include high white blood cell counts and blood clots. When used as a cream side effects include skin redness, peeling, and sun sensitivity. Use during pregnancy is known to harm the baby. Accutane was patented in 1957 and approved for medical use in 1962. The compound was first studied in the 1960s at Roche Laboratories in Switzerland by Werner Bollag as a treatment for skin cancer. Experiments completed in 1971 showed that the compound was likely to be ineffective for cancer and, surprisingly, that it could be useful to treat acne. However, they also showed that the compound was likely to cause birth defects.

So in light of the events around thalidomide (of which I will report in in a later issue of this journal), Roche abandoned the product. In 1975, Gary Peck and Frank Yoder independently rediscovered the drug’s use as a treatment of cystic acne (the most severe form of acne) while studying it as a treatment for lamellar ichthyosis, and published that work. Roche resumed work on the drug. In clinical trials, subjects were carefully screened to avoid including women who were or might become pregnant. Roche’s New Drug Application for isotretinoin for the treatment of acne included data showing that the drug caused birth defects in rabbits. The FDA approved the application in 1982.

Scientists involved in the clinical trials published articles warning of birth defects at the same time the drug was launched in the US, but nonetheless Accutane was taken up quickly and widely, both among dermatologists and general practitioners. Cases of birth defects showed up in the first year, leading the FDA to begin publishing case reports and to Roche sending warning letters to doctors and placing warning stickers on drug bottles, and including stronger warnings on the label. Lawsuits against Roche started to be filed. In 1983 the FDA’s advisory committee was convened and recommended stronger measures, which the FDA took and were at that time unprecedented: warning blood banks not to accept blood from people taking the drug and adding a warning to the label advising women to start taking contraceptives a month before starting the drug. However, use of the drug continued to grow, as did the number of babies born with birth defects. In 1985 the label was updated to include a boxed warning (see below).

In early 1988 the FDA called for a second advi-
sory committee. An internal FDA memo was leaked to the New York Times a few days before the meeting stating that around 1,000 babies had been born with birth defects due to Accutane, that up to around 1,000 miscarriages had been caused, and that between 5,000 and 7,000 women had had abortions due to Accutane. This led to a storm of media attention. In the committee meeting, dermatologists and Roche each argued to keep the drug on the market but to increase education efforts; pediatricians, FDA safety reviewers and the Centers for Disease Control (CDC) argued to withdraw the drug from the market. The committee recommended to restrict physicians who could prescribe the drug and to require a second opinion before it could be prescribed. The FDA, believing it did not have authority under the law to restrict who had the right to prescribe the drug, kept the drug on the market but took further unprecedented measures: it required Roche to make warnings yet more visible and graphic, provide doctors with informed consent forms to be used when prescribing the drug, and to conduct follow up studies to test whether the measures were reducing exposure of pregnant women to the drug. Roche implemented those measures and offered to pay for contraception counseling and pregnancy testing for women prescribed the drug - the program was called the “Pregnancy Prevention Program”.

A CDC report published in 2000 showed problems with the Pregnancy Prevention Program and showed that the increase in prescriptions was from off-label use, and prompted Roche to revamp its program, renaming it the “Targeted Pregnancy Prevention Program” and adding label changes like requirements for two pregnancy tests, two kinds of contraception, and for doctors to provide pharmacists with prescriptions directly; providing additional educational materials, and providing free pregnancy tests. The FDA had another advisory meeting in late 2000 that again debated how to prevent pregnant women from being exposed to the drug; dermatologists testified about the remarkable efficacy of the drug, the psychological impact of acne, and demanded autonomy to prescribe the drug; others argued that the drug be withdrawn or much stricter measures be taken.

In 2001 the FDA announced a new regulatory scheme called SMART (System to Manage Accutane Related Teratogenicity) that required Roche to provide defined training materials to doctors, and for doctors to sign and return a letter to Roche acknowledging that they had reviewed the training materials, for Roche to then send stickers to doctors which they would have to place on prescriptions they give people after they have confirmed a negative pregnancy test; prescriptions could only be written for 30 days and could not be renewed, thus requiring a new pregnancy test for each prescription.

As reported in Wikipedia that in February 2002, Roche’s patents for Accutane expired, and there are now many other companies selling cheaper generic versions of the drug. On June 29, 2009, Roche Pharmaceuticals, the original creator and distributor of Accutane, officially discontinued both the manufacture and distribution of their Accutane brand in the United States due to what the company described as business reasons related to low market share (below 5%), coupled with the high cost of defending personal-injury lawsuits brought by some people who took the drug. Generic versions of Accutane remained available in the United States through various manufacturers however. Roche USA continues to defend Accutane and claims to have treated over 13 million people since its introduction in 1982.

It was also reported that several trials over inflammatory bowel disease claims have been held in the United States thus far, with many of them resulting in multimillion-dollar judgments against the makers of Accutane or its generics. F. Hoffmann-La Roche Ltd. continues to manufacture and distribute Accutane under the brand name RoAccutane outside of the United States.

In 1976, World Health Organization reported more than 600 cases of Ebola found in Congo and Sudan. It raised major concerns in health organizations worldwide. In 1989, there was an outbreak of Ebola-related cases found in Northern Virginia near where I live. Even though it is not FDA work related, I found it is interesting to also report as follows.

**HOT ZONE:**

1943 Isaac Newton Square, Reston, VA was about 3 miles from the apartment we lived in 1987 and less than 7 miles from my current home in Herndon VA. A company called Hazelton Research operated a quarantine center there for monkeys that were destined for laboratories. In October 1989, when an unusually high number of their monkeys began to die, their veterinarian decided to send some samples to Fort Detrick (USAMRIID) for study. At the time, it was believed that the virus was Simian hemorrhagic fever virus, a viral hemorrhagic fever harmless to humans but almost always fatal to other primates. Early during the testing process in biosafety level 3, when one of the flasks appeared to be contaminated with harm-
less pseudomonas bacterium, two USAMRIID scientists exposed themselves to the virus by wafting the flask. When they eventually tested the samples with known Level 4 agents, only EBOV reacted with the unknown samples. They decided not to tell anyone about their exposure, but they did secretly tested their blood every day. After one of the monkey house staff members became ill with nausea and violent vomiting, USAMRIID was given permission to send in a team to euthanize all the monkeys at the facility and collect tissue samples. They later determined that, while the virus is terrifyingly lethal to monkeys, humans can be infected with it without any health effects at all. This virus is now known as Reston virus (RESTV), a relative of Ebola. The virus found was a mutated form of the original Ebola virus, and was initially mistaken for Simian Hemorrhagic Fever (SHV). The original Reston facility located at 1946 Isaac Newton Square was torn down sometime between 1995 and 1998.

Richard Preston investigated the event and wrote a book called “The Hot Zone”. Due to the detailed and graphic descriptions of the effects of exotic tropical diseases, as well as the revelation that an ebola virus was found a few miles away from Washington D.C., The Hot Zone was hailed by many as a chilling and accurate story of lethal viruses and their encounters with humans. Because Preston’s writing style is that of a “science fact” thriller, some critics accused Preston of dramatizing and exaggerating the effects of an Ebola virus infection and embellishing facts with his own imagination. Since its publication over a decade ago, however, The Hot Zone is generally regarded as a nonfiction work and acknowledged for its masterful dramatization. In his blurb, horror writer Stephen King called the book, “one of the most horrifying things I’ve ever read.”
Impact and Citations

Hans Rudolf Künsch

Research can be a frustrating and lonely enterprise when all your attempts to prove a conjecture lead nowhere, or when you receive reports by referees who didn’t understand the new idea of your paper, or when nobody among your colleagues shows any interest in what you are trying to do. So why do you not give up? One reason is extrinsic: Successful research is an essential condition to obtain a position, a promotion or a salary rise. But there is also the pure intrinsic motivation, “the pleasure of finding things out” as the title of a book by Richard Feynman says (Perseus, 1999). A third reason is presumably that we would like to make an impact on science and – more ambitiously – on society. This third motivation interacts with the two others: Success of research is usually judged by its impact, and the pleasure of finding things out is much higher if what you have found is also relevant for others.

The number of citations of a researcher is an indicator for her or his impact that is better than the number of publications. The main reason for using citations in a scientific work is to distinguish between your own work and that of others, thereby respecting intellectual property. If you cite someone else’s work, you acknowledge that this work has had an impact on your research, because it has been useful or even essential for you to arrive at your own result or conclusion. However, the number of citations as an impact measure has its shortcomings too. For instance, the cited work is often not the origin of the idea or result that was used, but it is more easily accessible or understandable than the original source. This means that someone with highly original ideas that are difficult to understand may not get as many citations as he or she deserves. This is similar to “Stigler’s Law of eponomy” (due Robert K. Merton and maybe others) which states that no scientific discovery is named after the person who discovered it. Often a citation also serves to convince the reader of the importance of the topic considered. In that case, the cited author is typically someone well known with a high reputation, but not necessarily the one who has made the most essential contribution. In other cases the cited work contains a somewhat different approach to the same problem and the reason for citing it is that it allows the author to show the superiority of the new method. Finally one should not forget that some authors find ways to manipulate citation statistics to their advantage, e.g. by citing earlier work of their own or of their network even though it has little relevance for the paper on hand, or by requiring citations of their own work when refereeing someone else’s manuscript.

In order to really understand what a citation means one has to look at the paper where this citation has been made. The quality of research of a scientist or a whole department should therefore not be evaluated on the basis of citation statistics alone. Reducing citation statistics to a single index is even worse. A serious evaluation of someone’s research requires that one looks at some of his or her papers. I think it is therefore a good idea to list your five best or most important papers (which need not be your most cited ones) on your cv.

For the same reason, you should not base your intrinsic motivation for research too much on your own citation statistics. Progress in research is not only achieved by singular contributions of a few geniuses, but also by a complex web of interactions between many different contributions. This has been noticed already in the 18th century by Georg Christoph Lichtenberg (1742-1799), a German physicist and satirist. Today he is best known for his aphorisms, concise and witty philosophical statements. He wrote once “If I hadn’t written this book, then in 1000 years from today between 6 and 7 pm people in many German towns would speak about entirely different things than those they will actually be speaking about.” Surely, he didn’t expect that people in a 1000 years in Germany would still be reading and discussing this one book, but rather that everything that is published continues to have some unpredictable indirect effect in the future, even if it is completely forgotten in a short time – a bit like the famous butterfly in Brazil of Ed Lorenz that sets off a tornado in the US by flapping his wings.

Moreover, you can also make an impact by other things than your publications. The guidance and feedback you give to your students, the questions you ask after a seminar talk, the suggestions you give in a referee report all can have a positive effect on the scientific progress. Obviously, by the same means you can also have a negative impact by missing to see the value of an important, but badly described idea. To give an example of a positive impact I received in my own career, my most cited paper that introduced the block bootstrap was triggered by a question from Colin Mallows at Bell
Labs. I had told him about my earlier work on defining an influence function for time series data. He then remarked that the influence function has a close relation to the jackknife and that I therefore should look at the jackknife for time series.

I hope that you experience in your career not only the pleasure of finding things out, but also the pleasure of sharing your ideas with others. It will have a positive impact even though it is not always measurable.

A more detailed discussion about the use and misuse of citation statistics can be found in the paper by R. Adler, J. Ewing and P. Taylor, Statistical Science 24 (2009), 1-14.

Give Industry a Chance

Terry Speed


What do people promoting data science and big data want that we statisticians do not seem to have? Why do so few PhD candidates build their theses around a specific application? Why don’t more PhD students in statistics do internships with industry over their summers? Nothing is simple with questions like these, so my initial answers to these questions are bound to be simplistic, probably wrong-headed, and definitely tendentious, but let me give them anyway.

One thing I notice about discussions of data science and big data is that they are invariably in the context of specific application areas. Global change, brain signals, earthquake signals, supernovae, social unrest, traffic accidents and smart thermostats are all named on the website of the Berkeley Institute of Data Sciences, part of a three campus $38M initiative launched by the White House in November last year, to be housed in the University Library at Berkeley. One thing I notice about discussions of mathematical statistics is that they are rarely in the context of a specific application area.

Recently I talked to participants in a Mathematics in Industry Study Group (MISG) and to students at an Industry Doctoral Training Centre (IDTC), both in Australia; to postdocs at a US academic institution who are struggling to maintain—not to mention further—their careers there; and to a scientist from a small, successful US company keen to employ statisticians or statistics interns. I found some common themes in our discussions, including the perennial “applied”-vs-applied, and academia-vs-industry divisions. There was also a general feeling that these issues do not get as fully discussed as they should in academia, where most of us reside.

When writing about factors motivating the first Australian Mathematics in Industry Study Group, 30 years ago, one of the organizers remarked that he had attended an applied mathematics conference at which just two or three of the 120 delegates were from industry, and only about 20 out of 56 talks had specific applications in mind, i.e. were really applied. In part, the MISG was started to change this, and it has been very successful. I wonder what proportion of the statistics talks at a typical IMS conference are applied, in the preceding sense, and whether we need an initiative to change this? My feeling is: possibly.

The Industry Doctoral Training Centre seeks to bring together an industrial problem and an industry sponsor with a PhD student and an academic advisor. This seemed a wonderful program: industry gets problems solved, while the students not only get something with direct real-world value to work on for their PhD, they develop communication, teamwork and leadership skills, as well as immersing themselves in the subject matter of their problem. I heard that finding industrial problems and students wanting to work on them was not as hard as finding suitable academic advisors.

Chatting with people, some of whom were on
their second or third postdoc, about the difficulty of getting research grants or tenure-track academic jobs, I couldn’t resist asking them how seriously they had considered other careers. How would they like a job, I asked them, which didn’t require applying for grants, which permitted them to work reasonable hours, paid well, and was enjoyable, challenging and fulfilling? Sounds good, was the response. There are plenty of such jobs outside academia, I told them. They freely conceded their value system had probably been skewed by their long sojourn in academia, and by their respect and admiration for their advisors, and that their belief that academic jobs were the best by far was not necessarily based on complete information. It’s my impression, I told them, that professors are very good at helping students develop into people like themselves, often many more than could reasonably have a career like them, even if that was desirable, but less good at pointing to or promoting alternative careers. That, I said was up to them to explore. In a better world, their advisors would be well informed and be able to discuss a broader range of futures with them.

What should be done? Stat department chairs could establish or strengthen existing links with industry, strongly encourage all grad students to take summer internships, ensure that they all know about careers outside academia, arrange adjunct appointments for interested and suitably qualified people from industry, and encourage their faculty to become involved with industry as well, for example via sabbaticals or summer internships. Then we might hope to see more grad students working on statistical problems from industry for their PhD research, something which would serve the triple purpose of solving an applied problem of real interest and perhaps importance, familiarizing themselves with a field other than statistics, and broadening their career opportunities.

Terry Speed, Ph.D.
Professor and Lab Head
Bioinformatics Division
Walter & Eliza Hall Institute of Medical Research
Parkville, Victoria
Australia

A Fundamental Link between Statistics and Humor

Xiao-Li Meng


My new year’s reading started with a holiday gift: On the Money, a collection of over 400 cartoons in The New Yorker from 1925–2009. No, after months of learning about fundraising, money was least alluring on a day when my alarm clock took a rest. But I could use a few laughs, even at my own expense or with irony. Indeed, the gift was from an alumna, and I wondered if it was meant to be a friendly reminder: relax, don’t take money (and your job) too seriously.

However I did end up taking the book very seriously, reading it word by word. Huh? Reading cartoons? Of course not. But there is an introduction by Malcolm Gladwell, whose name might not be as familiar to statisticians as the title of one of his best-sellers: Outliers (and perhaps also The Tipping Point and Blink). Gladwell’s philosophical introduction also started with irony. As a writer for The New Yorker, he considered the book very strange, “because we are a magazine for people for whom money is a secondary concern. …So what on earth does The New Yorker think about when we think about money?” Regardless of your opinions about his characterization, his short answer was, “we make jokes about it.”

Gladwell’s long answer began with an anecdote of how stunned he was at a corporate retreat, when a CFO used a business-like PowerPoint presentation to tell his life story. It ended with the key point: “People who want the world to conform to the principles of business are Realists. People who think the other way around—this is true whether they
spend their days parsing sonnets or actuarial tables—are Romantics, and the Romantic position...is the comic position.”

Gladwell’s main supporting example is a statistical one. Bernard Madoff’s fund had a 96% “winning” percentage, with “annual gains that fell like clockwork between 10—12 percent.” Why didn’t he mimic the volatility of a hedge fund, with a more enticing long-run gain and a much less suspicious winning pattern? Gladwell’s answer is that it is “because Madoff understood what consistency means in personal terms: it means trustworthiness, mastery, competence, safety.” It was precisely this consistency that convinced Harry Markopolos, Madoff’s bête noire, that Madoff was really a Made-up. As Markopolos argued in his 17-page memo to the SEC, “No major league baseball hitter bats .960, no NFL team has ever gone 96 wins and only four losses over a 100-game span, and you can bet everything you own that no money manager is up 96 percent of the months, either.” In other words, when we take a Realist’s position, Madoff’s made-up consistency points us to exactly the opposite conclusion it had aimed to achieve, that is, he is completely untrustworthy.

By no means is this example meant to glorify the Realists’ position. As Gladwell emphasized: “The victims of Bernie Madoff would have done well to think of Madoff in business terms, not personal terms. Then again, the traders at AIG, who have cost taxpayers many, many multiples of what Madoff cost the world, would have done well to import a healthy dose of personal virtue into their professional practices.”

Despite the painful context, Gladwell’s underlying message inspired me to think about a fundamental link between statistics and humor. The discipline of statistics is essentially about separating commonalities (e.g., patterns, signals) from individualities (e.g., variability, noise). In contrast, the best kind of humor is often the result of judiciously mixing commonalities and individualities to create comic effect.

A story told by Rick Cleary (visiting Harvard from Bentley) at our 2008 holiday party, illustrates this point well. At the end of the last lecture of an introductory course for which he was a teaching assistant, the professor (the late George Casella—whose obituary appears here) encouraged the class to ask any remaining questions on anything that had been covered. A student who had never asked any question before raised her hand. “Professor Casella, I really enjoyed your class, but there is one thing that has puzzled me for the entire semester. Why are standard deviations always six?”

You will be laughing now, or in a few seconds, if you are a real statistician. Otherwise you would be laughing at how nerdy statisticians must be if they can find humor in the number six. What makes this story greatly humorous to statisticians is the mixing of a well-understood commonality (standard deviation is commonly denoted by σ) and an unexpected individuality (the student’s mistaking σ for 6). It would not be humorous at all if six were replaced by one because George, for whatever reason, decided to use the letter l for standard deviation in his course.

What could be a more joyful way to celebrate the International Year of Statistics than by telling the world that Statistics is the most enjoyable profession on earth, because along with every depressing study or erroneous argument there is an enticing recipe for entertaining ourselves?

Xiao-Li Meng, Ph.D.
Whipple V. N. Jones Professor of Statistics
Department of Statistics
Harvard University
Upcoming Events

Please find below a list of upcoming ICSA meetings and co-sponsored meetings. This list also appears on the ICSA website. If you have any questions, please contact Dr. Gang Li, the ICSA Executive Director (gli@its.jnj.com).

ICSA Sponsored Meetings:

The 11th ICSA International Conference
December 20 - December 22, 2019
Hangzhou, Zhejiang, China
Organizing committee Chair: Hongzhe Li (hongzhe@pennmedicine.upenn.edu)

2020 ICSA Applied Statistics Symposium
Huston, TX, USA
May 17 - May 20, 2020
Professor Hulin Wu at University of Texas Health Science Center at Houston (Hulin.Wu@uth.tmc.edu) will be the Chair of the organizing committee.

ICSA Co-sponsored Meetings:

Conference on Current Trends in Survey Statistics
August 13 - August 16, 2019
National University of Singapore, Singapore
For more information about the conference, please visit https://ims.nus.edu.sg/orgsites/2019data/.

The 6th International Symposium on Biopharmaceutical Statistics
August 26 - August 30, 2019
Kyoto, Japan
For more information about the conference, please visit https://www2.aeplan.co.jp/isbs2019/index.html.

The 3rd International Conference on Statistical Distributions and Applications
October 10 - October 12, 2019
Grand Rapids, MI, USA
For more information about the conference, please visit http://people.cst.cmich.edu/lee1c/icosda2019.

The 11th International Multiple Comparisons Procedures (MCP) Conference
December 12 - December 15, 2019
Taipei
For more information about the conference, please visit http://www.mcp-conference.org/hp/2019/.

Chapter Meeting:

The 4rd ICSA —Canada Chapter Symposium, Advances and Innovations in Statistics and Data Science
August 9 - August 11, 2019
Queen’s University in Kingston, Ontario, Canada
The 2020 ICSA Applied Statistics Symposium will be held from Sunday, May 17th to Wednesday, May 20th, 2020, at The Westin Galleria Houston, Houston, Texas, USA. This will be the 29th annual symposium for the International Chinese Statistical Association (ICSA). The theme of this conference is **Advancing Statistics for Data Intelligence**. For more information, please contact symposium2020@icsa.org and visit the conference website https://symposium2020.icsa.org.

**Speakers**

- **Keynote Speaker**
  - Xihong Lin, PhD
  - Professor, Harvard University

- **Keynote Speaker**
  - Michael I. Jordan, PhD
  - Professor, University of California, Berkeley

- **Keynote Speaker**
  - Josh Chen, PhD, Head, Global Biostatistical Sciences, Sanofi Pasteur

- **Banquet Speaker**
  - Hong Ogle, MS
  - Bank of America Houston Market President

**Key Dates**

- Early Bird Registration Deadline: March 15, 2020
- Invited Session Proposal Submission Deadline: November 1, 2019
- Short Course Proposal Submission Deadline: December 15, 2019
- Student Paper Award Application Deadline: February 15, 2020
- Invited Session Abstract Submission Deadline: March 15, 2020
- Contributed Session Abstract Submission Deadline: April 15, 2020
- CV/Resume for Career Services Submission Deadline: May 1, 2020

**Houston**

Houston is the fourth-most populous city in the United States. Notable facilities include the Museum District, the Houston Zoo, and NASA’s Lyndon B. Johnson Space Center. Houston is the seat of the internationally renowned Texas Medical Center (TMC), which contains the world's largest concentration of research and healthcare institutions including University of Texas Health Science Center at Houston, MD Anderson Cancer Center, Baylor College of Medicine, Memorial Hermann Hospital, The Methodist Hospital, and Texas Children's Hospital. Several universities of higher education such as The University of St. Thomas, University of Houston, and Rice University are located within the city.