ISSN: 2226-2393

www.icsa.org

July, 2021

Volume 33, Issue 2

International Chinese Statistical Association

泛華統計協會

Bulletin

會刊

ICSA Bulletin

Volume 33/2, July, 2021

ISSN 2226-2393

Editorial Staff

From the Editor

Ming Wang

Dear ICSA Members:

Welcome to the 2021 July issue of the ICSA Bulletin! This issue will cover diversity topics related to the theme of "Opportunities and Challenges in Statistics and Data Science".

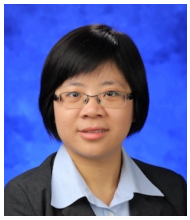
In this issue, we have two featured articles. The first one is written by Dr. Jeffery Morris, Professor and Director of the Division of Biostatistics at Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania. Dr. Morris has tremendous experience in developing quantitative methods to extract knowledge from biomedical big data. In this article, he shared his thoughts on role of biostatisticians in the big data world of modern biomedical research, and discussed the skill set of perspective of biostatisticians in order to play a central role in the emerging big data world of biomedical and population health research. The second one is an interview article with Dr. Naitee Ting, written by Dr. Qiqi Deng from Boehringer Ingelheim Pharmaceuticals and Dr. Joseph C. Cappelleri from Pfizer Inc. Dr. Ting, a lifetime member of ICSA since 1987, worked in Pfizer Inc. between 1987 and 2009, and is currently a director in the biostatistics group of Boehringer Ingelheim Pharmaceuticals. Dr. Ting's research interests include dose response in clinical trials, variance component models, and clinical study designs. Diverse interview questions were covered, including Dr. Ting's grow-up environment, education history, his philosophy for life and community, pharmaceutical working experience, major accomplishment, among others. Thanks to Qiqi and Joseph for providing this opportunity to comprehensively learn Dr. Ting's experience and insights as a pharmaceutical statistician!

As usual, three column articles are included to continue our columnists' stories. In the column "Hints from Hans", Professor Hans Rudolf K̈unsch discussed several issues with regard to teaching, and explained why, what and how to teach and assess teaching efforts. Those members who are particularly working in academia will identify similar teaching experience, thus benefit a great deal from Prof. Hans' thoughts. In the column of "Yi's FDA Stories", Dr. Yi Tsong brought us back to the history when statistics met regulation 1995. In particular, Dr. Tsong mentioned United States Pharmacopeia (USP) procedure and presented a relatively clear description of the compendia used by USP and associated statistical procedures. In the column "XL-Files", Professor Xiao-Li Meng shared his journey working and struggling with "dark" data, and also recommended Prof. David Hand's book, Dark Data, for us. This is a reprint from the author's column "Time Travel and Dark Data" in the IMS Bulletin, February 2020. In addition, we include one memorial article for Dr. Kai Fun Yu, former chief of the Biometry and Mathematical Statistics Branch at the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (NICHD/NIH) and a founding member of ICSA. Drs. Colin Wu, Jiayang Sun and Heping Zhang and Aiyi Liu collected memorizing stories from Dr. Yu's colleagues to express condolences. We are deeply sad to lose such a great member, and Dr. Yu will be always remembered!

Besides these, we also have one article entitled "Data, Data, Everywhere..." from Dr. Nicole Lazar, Professor of Statistics at University of Georgia. Without doubt, we are living in the modern world surrounded by data that are more varied and vaster than ever before. Dr. Lazar introduced a big picture of data science, and also provided tutorial on how to develop into more data literate and uncertainty literate consumers of information. This article is reprinted from the Harvard Data Sciences Review (HDSR) with permission (https://hdsr.mitpress.mit.edu/pub/yrzpcrus/release/2). Further, we have two recaps of panel discussions, one offering leadership advice and the impact of COVID-19 on the workplace in the 2020 ICSA Symposium (https://magazine.amstat.org/blog/2021/04/01/icsa) and the other one providing Insights on real-world evidence in healthcare (https://magazine.amstat.org/blog/2021/06/01/panel-rwe). Both articles are reprinted with permission from Amstat News. We appreciate Dr. Kelly Zou for organizing the panels and coordinating with the panelists for write-ups! Moreover, it is highlighted that our member, Prof. Bin Yu of UC Berkeley, has been awarded an Honorary Doctorate from the University of Lausanne (UNIL), Faculty of Business and Economics, in Switzerland. She was honored as "one of the most influential researchers of her time in Statistics and Data Science, for the excellence and impact of her work, and for her major contributions to the development and advancement of machine learning." Congratulations to Prof. Yu for this achievement!

Turning to ICSA business, we include the letter from the 2021 ICSA President, Dr. Colin Wu, and the letter from the executive director, Dr. Mengling Liu; announcement of the recipients of the 2021 ICSA Awards and 2022 ICSA Pao-Lu Hsu award; the candidates for 2022 ICSA Officers; new fellows of ASA and IMS in our ICSA community; reports from the 2021 program committee; report from Statistics in Biosciences (SIBS) co-editors; and the announcements of the upcoming meetings/conferences at the end of this issue.

In the end, I would like to thank all the contributors, ICSA executives and committee members for their support on our bulletin. Also, a big round of applause is given to my great assistant, Dr. Chixiang Chen, for putting these files together. Hope you enjoy our issue and best wishes to you'll and your family!



Ming Wang, Ph.D.
Editor-in-Chief, ICSA Bulletin
Associate Professor
Division of Biostatistics and Bioinformatics
Penn State College of Medicine

From the 2021 President, ICSA

Colin Wu

Dear ICSA members and friends,

As we enter the second half of 2021, I would like to thank our ICSA members and friends for their persistent effort on being productive and participating in ICSA activities. The first half of 2021 was challenging due to the COVID-19 pandemic. But the disruption did not stop our committee members and volunteers from carrying out their activities and generating novel ideas for ICSA. Their enthusiasm and creativity are the foundation for maintaining ICSA as a vibrant statistical society. As a result, we have continued our time-honored tradition of contributing to statistics and data science with ingenuity and creativity.

Let me start by giving a big applause to our new ASA and IMS fellows. In 2021, sixteen ICSA members became new ASA fellows and eighteen ICSA members became new IMS fellows. Their innovative contributions to statistics and data science greatly strengthened the reputation of ICSA as a leading international statistical society. I would also like to congratulate our new ICSA awardees, including the ICSA Distinguished Achievement Award, the ICSA Distinguished Service Award, the ICSA President Citation, the ICSA Outstanding Young Researcher Awards, and the Pao-Lu Hsu Award, which were given to both young and established outstanding ICSA members who contributed extensively in the new frontiers of statistics and supporting ICSA. Congratulations to all, and thank you for your work and service.

Looking back to the first half of 2021, ICSA continued to engage in its core activities despite the disruption caused by the COVID-19 pandemic. Major activities during this period included organizing ICSA meetings, selecting new editors for ICSA journals and book series, collaborating with our sister statistical societies (i.e., the International Indian Statistical Association (IISA) and the Korean International Statistical Society (KISS)) on diversity and anti-racism, promoting novel statistical methods in new frontiers of data science, and encouraging students and early career statisticians to participate in professional activities.

To further elaborate our achievement and objectives for the rest of 2021, I would like to thank Dr. Guoqing Diao and members of the Program Committee for organizing the upcoming 2021 ICSA Applied Statistics Symposium, which will be held virtually in September 12-15, 2021. I would like to encourage our members to support Dr. Diao and the Program Committee by actively participating in the sessions and panel discussions. Some of you may have also attended, virtually of course, other ICSA cosponsored meetings, such as the Duke-Industry Statistics Symposium (April 21-23, 2021) and the 63th ISI World Statistics Congress 2021 (July 11-16, 2021). A benefit of attending virtual meetings

is that you could attend ISI World Congress without going through the burden of international travel. But, of course, the possible downside is also not having a chance for international travel. Because of the restriction on international travel, it is unfortunate that we have to postpone the 2021 ICSA China Conference to 2022. For future meetings cosponsored by ICSA, you may consider the 8th Workshop on Biostatistics and Bioinformatics (postponed to Spring, 2022), the 77th Annual Deming Conference on Applied Statistics (December 6-8, 2021) and the IMS Asia Pacific Rim Meeting (postponed to January 5-8, 2022).

On the front of ICSA journals, we are fortunate to have seven prominent ICSA members to serve as co-editors and editors-in-chief for Statistica Sinica (co-editors: Rong Chen, Su-Yun Huang and Xiaotong Shen) and Statistics in Biosciences (editors-in-chief: Hongzhe Li and Joan Hu). I would like to thank them for their willingness to lead the effort, and I am sure that, with their stewardship, our journals are in good hands and will continue to thrive. Our journal editors will also organize invited paper sessions at future Joint Statistical Meetings, which will be a great opportunity to showcase the innovative research published by our journals. Given the need for more innovative research in "real-world data/real-world evidence" (RWD/RWE), our journals will give special emphasis on novel statistical machine learning and innovative study designs. In addition to our journals, I would like to thank Dr. Din Chen for continuing his role as editor of the ICSA Springer Book Series. So far, the ICSA book series has published a total of 22 books with 2 new books published in 2019 and 2 more in 2020. I would like to encourage our members to support ICSA by considering our journals and book series as the first in line choice for publishing your work.

In a collaborative effort with our sister Asian American statistical societies, IISA and KISS, executive officers of our three societies held ad hoc meetings throughout the year discussing the pressing issues facing Asian American statisticians. As a result, each of the three societies issued their separate statements for promoting diversity and antiracism, and the ASA Committee on Nomination for 2021 has nominated Dr. Lilly Yue (Deputy Director, Division of Biostatistics, CDRH/FDA) as a candidate for ASA Vice-President. We still have a long way to go in this endeavor. But, I am happy to report that senior leaders of all three societies agree that such collaborations are fruitful and we should continue these efforts in the future.

Looking further into the future, a major task facing us is to promote ICSA as an attractive professional home to students and early career statisticians. Many ICSA members have participated in the discussions on statistical education and training, and ICSA had organized online training sessions for students and young researchers. The ICSA membership and publication committees also took the opportunities of ICSA conferences and training sessions to attract more students and young statisticians to join ICSA. We should continue and expand this effort by offering real benefit and opportunities to our members. I am looking forward to working with you on this important task in the second half of 2021.

Finally, I would like to express my deepest sympathies to Dr. Kai Fun Yu's family. Dr. Yu, former chief of the Biometry and Mathematical Statistics Branch (now known as the Biostatistics and Bioinformatics Branch) at the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (NICHD/NIH), and a founding member of ICSA, passed away on July 8, 2021 at age 70 after a 7-month battle with lymphoma. I first met Dr. Yu in 1994 when my wife and I participated in a weekend study group of Chinese statisticians in the Washington DC area in which Dr. Yu was an active and founding member. Later on Dr. Yu invited me to spend my sabbatical year in 1997 in NICHD collaborating with him and other NIH scientists. That was the first time I came to close encounter with NIH statistics groups, which eventually led to my statistical career at NIH starting in 2002. In that respect, Dr. Yu was my unofficial mentor for my professional career at NIH. I will always cherish the memorable moments I spent with him, and I missed him dearly.

Colin Wu Ph.D.
2021 President, ICSA
Mathematical Statistician
National Heart, Lung and Blood Institute
National Institutes of Health

From the Executive Director 2020-2022

Mengling Liu

Dear ICSA members,

As we have been making progress towards to return to normal work and life environment, the first half of 2021 remained challenging in many fronts. I would like to send my sincere gratitude to all ICSA officers and members to keep us staying on track and going forward. The ICSA Executive Committee met regularly every month to ensure that all administrative functions and tasks were completed on time and had additional meetings with various committees and organizers to provide guidance. The Award Committee (chaired by Dr. Judy Wang) and Nomination Committee (Chaired by Dr. Ming Tan) worked tirelessly to solicit high-quality candidates and identify final winners and nominees. The Special Lecture Committee (chaired by Gang Li) worked efficiently to identify potential keynote speakers for multiple ICSA meetings under a short timeline. The Program Committee (chaired by Dr. Hulin Wu) closed worked with current and future ICSA meeting organizers to ensure the continuity of ICSA meeting activities. In this bulletin, I would like to call your attentions to various committee reports to learn about many ICSA activities by various and join me to thank these committee members for their time and effort dedicated to ICSA service. If you're interested in serving on any of the ICSA committees, please feel free to reach out to me at executive.director@icsa.org.

Although the pandemic made our meeting schedules a little out of sync, the first 2021 ICSA Board Meeting was held on June 6^{th} 2021 virtually. During the meeting, the board members had reviewed all committee activities and reports, and discussed and voted on nominations, awards, and financial plans. The 2021 ICSA Applied Statistics Symposium will be held virtually during September 12-15, 2021, and is currently under intensive planning by the Executive Organizing Committee led by Dr. Guoqing Diao. In addition to all scientific events, we will also hold our second Board meeting, General Member Meeting, and Award Ceremony during the symposium. We look forward to seeing you online and reconnecting. At the end, I would like wish all a safe, happy and restful summer.

Mengling Liu, Ph.D.
ICSA Executive Director (2020-2022)
Professor of Biostatistics
Department of Population Health
Department of Environmental Medicine
NYU Langone Health

2021 ICSA Awards

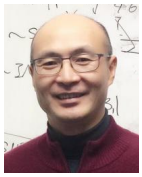
Awards Committee Chair: Judy Wang

2021 ICSA Distinguished Achievement Award

Cun-Hui Zhang, Rutgers University

Citation:
For fundamental and influential contributions in statistical theory and methods, especially in nonparametric statistics; for exemplifying service to and leadership in statistical communities, including the International Chinese Statistical Association; for outstanding mentorship of students and colleagues.

2021 Outstanding Service Award



Heping Zhang, Yale University

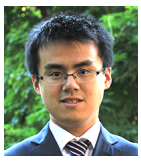
Citation:
In recognition and appreciation of his outstanding service to ICSA for over two decades and his leadership as 2019 ICSA President and a member of the ICSA Board of Directors (2001-2004).

2021 President Citation Award

Yi Tsong, Director, Division of Biometrics VI, CDER, FDA

Citation:
In recognition and appreciation of his dedication and outstanding service to ICSA, especially his long-standing contributions to ICSA activities and his leadership as ICSA President (2006), ICSA Executive Director (2001-2003), and a member of the ICSA Board of Directors (1995-1997).

2021 ICSA Outstanding Young Researcher Award



Anru Zhang, University of Wisconsin-Madison

Citation:
For his significant contributions to high-dimensional statistical inference, non-convex optimization, statistical learning theory, and particularly for his fundamental contributions to tensor learning.

Yifei Sun, Columbia University

Citation:
For her important and outstanding research for lifetime data science, in particular, her novel contribution for recurrent event analysis, stochastic lifetime processes, biased sampling, and predictive modeling.

2022 Pao-Lu Hsu Award

Ji Zhu, University of Michigan

Citation:
For his impactful contributions and pioneering work in the areas of statistical machine learning, statistical network analysis, and their applications in physics and healthcare; for his excellence in teaching, student advising, and service to the profession; and for his engagement in and promotion of statistics education and research in China.

Candidates for 2022 ICSA Officers

Candidates for ICSA 2022 ICSA President-Elect

Gang Li, PhD

Professor,
Departments of Biostatistics & Computational Medicine ,
Director,
UCLA's Jonsson Comprehensive Cancer Center Biostatistics,
University of California, Los Angeles.

Hongzhe Lee, PhD

Professor,
Biostatistics and Statistics, Vice Chair of Research Integration ,
Director,
Center for Statistics in Big Data,
University of Pennsylvania

Candidates for ICSA 2022 Board of Directors

Xun Chen
Sanofi

Wei Zhang
Boehringer Ingelheim

Bo Fu
Sanofi

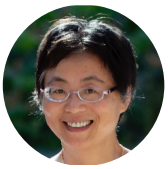
Gang Li
Eisai, Inc.

Yanping Wang
Eli Lilly

Lan Wang
University of Miami

Hao (Helen) Zhang
University of Arizona

Annie Qu
University of California Irvine



Wanli Qiao
George Mason University

Linda Zhao
University of Pennsylvania

Yuehua Cui
Michigan State University

Bin Zhang
Cincinnati Children's Hospital Medical Center

JianQing Shi
Southern University of Science and Technology, Shenzhen, China

Xingqiu Zhao
The Hong Kong Polytechnic University

Eligible ICSA regular members receive further details via e-mail prior to the start of elections, including a summary of the candidates, details on how to access the voting system. Please ensure the Membership Committee has your recent and correct e-mail address on file. For questions, you can reach the Membership Committee in private at membership@icsa.org.

2020/2021 New ASA and IMS Fellows

2020 ASA Fellows

KWUN CHUEN GARY CHAN
University of Washington
For outstanding contributions to the methodology of preferential sampling design, observational data, and complex lifetime data; for substantive biomedical applications; and for dedication to the mentoring of junior statisticians and health researchers.

CHUNG-CHOU H. CHANG
University of Pittsburgh
For demonstrated ability to advance the field of biostatistics; for being an invaluable collaborator and co-investigator for a large number of clinical researchers; and for being a superb teacher and mentor of students throughout the health sciences.

YONG CHEN
University of Pennsylvania Perelman School of Medicine
For major contributions to the methodology of evidence synthesis; for outstanding research in inference under nonstandard conditions, robust inference, and composite likelihood; for bridging statistics and informatics; and for service to the profession.

YUAN JI
The University of Chicago
For pioneering contributions to statistical methodology and application, including interval-based dose finding, big data cancer genomics, and bioinformatics using robust Bayes approaches; for providing exemplary public user interfaces; and for exemplary mentorship and service.

LIANG LI
MD Anderson Cancer Center
For excellent and sustained statistical research and collaboration in the analysis of observational longitudinal cohort studies and chronic disease research and for outstanding service to the profession.

QIZHAI LI
Chinese Academy of Sciences
For outstanding contributions to statistical methodology in cancer genetics and genetic epidemiology, particularly in genomewide association studies and candidate gene analysis, and for key contributions to diagnostic medicine and high-dimensional data analysis in biomedical research.

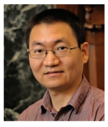
JASON JINZHONG LIAO
Merck & Co., Inc.
For innovative statistical methods and applications in the pharmaceutical industry; for advancing comparability/biosimilars studies, agreement studies, and assay development; for bringing life-saving drugs to market; and for extensive service to the profession.

CHING-TI LIU
Boston University School of Public Health
For outstanding contributions and strong leadership in both the development of innovative statistical methods and in collaborative research to elucidate the genetic basis of complex human traits and for dedicated service to the profession.

SHENG LUO
Duke University
For excellent contributions to statistical methods in longitudinal and survival analysis; for scientific contributions to the field of neurological diseases; for exemplary mentorship of students; and for significant service to the profession.

MICHAEL ROSENBLUM
Johns Hopkins University Bloomberg School of Public Health
For outstanding contributions to statistical methodology and applications, especially with respect to the adaptive design and optimal analysis of randomized trials.

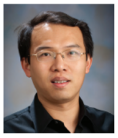
YONGMING QU
Eli Lilly and Company
For substantial and sustained contributions to statistical methodology for clinical trials and drug development; for significant influence and impact on bringing new drugs from research to market; and for excellent service to the profession.



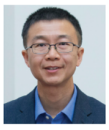
XIAOFENG SHAO
University of Illinois at Urbana-Champaign
For pioneering nonparametric inference for time series and high-dimensional inference using nonlinear dependence metrics, especially in developing self-normalization for dependent data and bootstrapping time series, and for excellence in student mentoring and service to the profession.

YIYUAN SHE
Florida State University
For novel and sustained contributions to high-dimensional and robust statistics; for promotion of sound application of statistical theory and optimization in signal processing and machine learning; and for excellence in mentoring, instruction, and service to the profession.

DAMLA SENTURK
University of California, Los Angeles
For methodological contributions in semiparametric modeling and functional data analysis; for innovative applications in neuroscience and other allied disciplines; for outstanding teaching and mentoring; and for exemplary service to the profession.



HAONAN WANG
Colorado State University
For pioneering work in object-oriented data analysis; for fundamental contributions to statistical learning, spatial statistics, and model selection; and for servic to the profession.



ZHENGYUAN ZHU
Iowa State University
For excellence in survey practice; for superb contributions to statistical theory and methods in spatial statistics, spatial sampling design, survey statistics, and functional data analysis; and for service to the profession.

2021 ASA Fellows

- Thomas M. Braun, Ph.D, Professor of Biostatics, School of Public Health, University of Michigan
- Jie Chen, Ph.D, Sr. VP, Head of Biometric, Overland Pharmaceuticals
- Zhen Chen, Ph.D, Sr. Investigator, NICHD' s Biostatistics and Bioinformatics Branch, NIH
- Li-Shan Huang, Ph.D, Professor, Institute of Statistics, National Tsing Hua University
- Yan Li, Ph.D. Professor, Epidemiology and Biostatistics, University of Maryland
- Fang Liu, Ph.D, Professor, Department of Applied and Computational Mathematics and Statistics, University of Notre Dame

- Bo Lu, Ph.D, Professor, Division of Biostatistics, The Ohio State University
- Weiwen Miao, Ph.D, Professor, Department of Mathematics and Statistics, Haverford College
- David Ian Ohlssen, Ph.D, Head of Advanced Exploratory Analytics at Novartis
- Kimberly F. Sellers, Ph.D, Professor, Department of Mathematics and Statistics, Georgetown University
- Rui Song, Ph.D, Professor, Department of Statistics, North Carolina State University
- Lilly Wang, Ph.D, Professor, Department of Statistics, Iowa State University
- Xiao Wang, Ph.D, Professor, Department of Statistics, Purdue University
- Xiaofei Wang, Ph.D, Professor, Department of Biostatics and Bioinformatics, Duke University
- Hongquan Xu, Ph.D, Professor and Chair, Department of Statistics, University of California, Los Angeles
- Yichuan Zhao, Ph.D, Prof., Department of Mathematics and Statistics, Georgia State University

2020 IMS Fellows

- Guang Cheng, Purdue University: *For outstanding work in methodology and theory of statistics, especially in high dimensional data, semiparametric estimation and inference, big data, and machine learning.*
- Chenlei Leng, University of Warwick: *For fundamental contributions to the theory and practice of high-dimensional statistics, statistical machine learning, model selection, and network data analysis.*
- Wei Pan, University of Minnesota: *For his important contributions to survival analysis, correlated data analysis, statistical learning, bioinformatics, and applications to biology and medicine, and for his dedicated services to the profession.*
- Lily Wang, Iowa State University: *For contributions to spatial, survey, image and functional analysis using nonparametric and semiparametric methods, especially to partially linear models, confidence envelopes and bivariate smoothing.*
- Grace Y. Yi, University of Western Ontario: *For research excellence in developing theory and methods for the analysis of survival data and longitudinal data in statistical and biostatistical applications, and for world-leading contributions to the analysis of missing and mismeasured data.*
- Li-Xin Zhang, Zhejiang University: *For important contributions to difficult problems in probability and statistical inference; and for excellence in mentoring and services.*

2021 IMS Fellows

- Kwun Chuen Gary Chan, University of Washington: *For outstanding contributions to the methodology of preferential sampling design, observational data, and complex lifetime data; for substantive leadership in the application of statistics in public health and biomedical research.*
- Scott H. Holan, University of Missouri: *For significant research contributions to time series, spatial statistics, spatio-temporal modeling, for applications to official statistics, and for his dedicated service to the statistics profession.*
- Jian Huang, University of Iowa:*For fundamental contributions to high-dimensional statistics, survival analysis, and statistical genetics and genomics.*

- Jing Lei, Carnegie Mellon University:*For fundamental and innovative contributions to sparse Principal Component Analysis, conformal prediction, network analysis, and spectral methods, optimal transport, cross-validation, and differential privacy, as well as applied work on the analysis of tissue and single-cell RNA sequencing data and autism.*
- Lexin Li, University of California, Berkeley: *For influential contributions to sufficient dimension reduction, tensor regression, and neuroimaging applications, and for outstanding service to the profession.*
- Huazhen Lin, Southwestern University of Finance and Economics: *For outstanding contributions to theory and methods for survival analysis and semiparametric/nonparametric modeling; for dedicated service to the profession; and for strong leadership in statistics education and development in China.*
- Steven MacEachern, The Ohio State University: *For fundamental and influential contributions to Bayesian statistics and inference, especially in the area of Bayesian nonparametrics, for dedicated mentoring of students and young researchers, and service to the profession.*
- Marianthi Markatou, University at Buffalo: *For outstanding accomplishments on fundamental research in robust statistics and for a strong commitment to the advancement of statistical science through interdisciplinary research and professional service.*
- Yiyuan She, Florida State University: *For novel and sustained contributions to high dimensional and robust statistics, for excellence in mentoring, instruction, and service to the profession.*
- Peter Xuekun Song, University of Michigan: *For pathbreaking developments in likelihood inference, for outstanding contribution in medical research, and for exceptional service and mentorship.*
- Rui Song, North Carolina State University: *For significant contributions to machine learning methods, dynamic treatment regime, and efficient and non-standard statistical inference.*
- Zhiqiang Tan, Rutgers, The State University of New Jersey: *For seminal contributions to statistical theory and practice, including Monte Carlo sampling and causal inference, to significant contributions in collaborative research, and for contributions to the profession, including via editorial and organizational efforts.*
- Nian-Sheng Tang, Yunnan University: *For excellent research accomplishments on statistical inference with missing data and distinguished contributions to the promotion of statistics in developing regions.*
- Xiao Wang, Purdue University: *For significant contributions to nonparametric statistics, shape-restricted inference, and functional data analysis, and for dedicated professional service and students' mentoring.*
- Changbao Wu, University of Waterloo: *For important and original research contributions to survey sampling theory and official statistics, especially for the development of model-calibration theory and techniques, empirical likelihood methods for complex surveys, and robust inferential procedures for analyzing non-probability samples.*
- Min-ge Xie, Rutgers, The State University of New Jersey: *For outstanding contributions to statistical research especially confidence distributions, fusion learning, meta-analysis, estimating equations, and statistical applications in biomedical sciences, industry, engineering, and environmental science. For exemplary service to the profession and university.*
- Zhengjun Zhang, University of Wisconsin–Madison: *For significant contributions to extreme value statistics and risk management, including tail dependence measure and nonlinear dependence measure construction and inference, max-linear competing factor models, nonlinear time series models for high-frequency financial data; and for conscientious editorial and other services to the profession.*
- Yijun Zuo, Michigan State University: *For pioneering and path-breaking contributions to the area of data depth, theory and methodology, and its applications.*

ICSA Financial Report: Janurary 1 Through June 30, 2021

International Chinese Statistical Association
Profit and Loss
January 1, 2021 through June 30, 2021

Beginning Cash Balance (Bank/Paypal accounts)	**1/1/2021**	**$**	**507,367.77**
Income:		**$**	**93,523.79**
Membership from Paypal Account		$	9,280.00
Membership from Institute of Mathematical Statistics		$	600.00
Membership by check		$	1,000.00
From JP Hsu account		$	2,000.00
Job Posting		$	150.00
Interest		$	105.45
Income from 2020 Houston Applied symposium		$	80,388.34
Total Income		**$**	**93,523.79**
Expense:		**$**	**(13,588.58)**
IT Cost		$	(11,319.60)
Support for NISS workshop		$	(2,000.00)
Fax Fee		$	(8.25)
1099 forms Filing Fee		$	(27.92)
Paypal Fee		$	(232.81)
Total Expense		**$**	**(13,588.58)**
Net Total Income		**$**	**79,935.21**
Transfer			
To 2021 Applied Symposium Account		$	(5,000.00)
Ending Cash Balance (Bank/Paypal accounts)	**6/30/2021**		**$582,302.98**
ASSETS			
Main Checking/Savings/PayPal		$	582,302.98
Vanguard Investment Balance		$	577,057.21
TOTAL ASSETS		**$**	**1,159,360.19**
LIABILITIES & EQUITY			
Equity			
Main Accounts Opening Balance January 1, 2021		$	507,367.77
July 1 to December 31, 2020 Net Income(+)/Expense(-)		*$*	*79,935.21*
Transfer to 2021 Symposium Account		*$*	*(5,000.00)*
Symposium Bank Accounts (2018, 2020, 2021) Opening Balance Jan 1, 2021		$	111,292.87
January 1 to June 30, 2021 Net Income(+)/Expense(-)		*$*	*(101,592.62)*
Vanguard investment account opening balance on Jan 1, 2021		$	524,890.52
January 1 to June 30, 2021 Investment Profit(+)/Loss(-)			*$52,166.69*
Total Equity		**$**	**1,169,060.44**
TOTAL LIABILITIES & EQUITY		**$**	**1,169,060.44**

Rongwei (Rochelle) Fu, PhD,
ICSA Treasurer (2021-2023)
Professor
OHSU-PSU School of Public Health

Report from the ICSA Program Committee

ICSA Program Committee

- Hulin Wu, Chair, ICSA 2021 Program Committee (Hulin.Wu@uth.tmc.edu)
- Xuming He (2019-2021, JSM Representative 2020, xmhe@umich.edu)
- Aiyi Liu (2020-2022, JSM Representative 2021, liua@mail.nih.gov)
- Pei Wang (2021-2023, JSM Representative 2022, pei.wang@mssm.edu)
- Guoqing Diao (2020-2022, ICSA Symposium 2021, gdiao@email.gwu.edu)
- Samuel Wu (2021-2023, ICSA Symposium 2022, samwu@biostat.ufl.edu)
- Hongzhe Lee (2021-2022, ICSA International Conference 2019, hongzhe@pennmedicine.upenn.edu)
- Xin-Yuan Song (2021-2023, ICSA International Conference 2022, xysong@sta.cuhk.edu.hk)
- Alan Y Chiang (2019-2021, achiang@celgene.com)
- Bin Nan (2019-2021, nanb@uci.edu)
- Ji Zhu (2019-2021, jizhu@umich.edu)
- Qingning Zhou (2020-2022, qzhou8@uncc.edu)
- Liang Zhu (2020-2022, Liang_Zhu@eisai.com)
- Jie Chen (2020-2022, jiechen0713@gmail.com)

Conferences and Meetings in 2021

- ICSA co-sponsored the 7^{th} Duke-Industry Statistics Symposium: Moved to virtual, April 21-23, 2021, Durham, NC. More than 330 participants across the USA participated in the 7th Duke-Industry Statistics Symposium (DISS2021). The symposium theme is "Emerging Clinical Initiatives in Pharmaceutical Development: Methodology and Regulatory Perspectives." The first day of the Symposium was devoted to five short courses. The second day started with a keynote speech by Dr ShaAvhrée Buckman-Garner, Director of the Office of Translational Sciences (OTS), Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration. It was followed with 25 parallel sessions on the second and the third days. The symposium also included a virtual poster session and a job fair. The Journal of Biopharmaceutical Statistics (JBS) will publish a Special Issue based on the invited talks and posters. More information about the symposium can be found at https://sites.duke.edu/diss/.
- ICSA co-sponsored the 4^{th} Regulatory Science and Biostatistics Summit, Guangzhou, China, November 12-14, 2021. Theme: Practice and Consensus-A Joint Journey for Development and Regulation. The Summit is sponsored by the China Tripartite Coordination Committee (TCC) and organized by the Guangdong Biostatistical Society and Southern Medical University. It is co-sponsored by the International Chinese Statistical Association (ICSA). The target audience includes regulators, quantitative scientists, medical physicians, regulatory affairs staff, and drug development professionals from government agencies, biopharmaceutical industry and academic community, with approximately 700 anticipated participants. The Summit consists of one-day preconference short sources and two-day scientific sessions including three keynote talks, Professor Shi Yuankai of Chinese Academy of Medicine will share his insights on innovation in drug development from PI perspectives, Professor CHEN Pingyan of Southern Medical University will discuss patient-focused drug development and patient-reported outcomes, and Dr. CHEN Jie of Overland Pharmaceuticals will present statistical innovation in changing landscape of drug development and regulation. The Summit Website: https://xueqiu.com/4376152234/180638632.
- 2021 ICSA Applied Statistics Symposium: Moved to virtual due to COVID-19, September 12-15, 2021. Organizing Committee Chair: Dr. Guoqing Diao at George Washington University. The 2021 ICSA Applied Statistics Symposium, themed "Leading with Statistics and Innovation", features three keynote speakers: Dr. Scott Evans (George Washington University), Dr. Nicholas P. Jewell (Univer-

sity of London), and Dr. Ram Tiwari (Bristol Myers Squibb). Sponsored by more than thirteen industry partners, the symposium will offer six short courses, 109 invited sessions, several poster sessions, and the student paper awards session. The symposium will also include a special panel on Leadership and Communication led by Dr. Jiayang Sun (George Mason University) and Dr. Hulin Wu (The University of Texas Health Science Center at Houston) and social activities organized by Dr. Kelly Zou (Viatris). The symposium website: https://symposium2021.icsa.org. Two conferences in 2021 postponed due to COVID-19:

- 2021 ICSA China Conference: Postponed to July 1-4, 2022 at Xi'an University of Finance and Economics, Xi'an, China.
- ICSA co-sponsored the 8th workshop on Biostatistics and Bioinformatics, Spring 2021, Atlanta, GA: Postponed to 2022 (date to be determined).

Future Conferences and Meetings

The COVID-19 continues to impact every aspect of our planned and planning activities. The program committee has reviewed and approved one proposal to host the 2023 ICSA Applied Statistics Symposium from the University of Michigan (Co-Chairs: Drs. Jian Kang and Gongjun Xu). Two universities in USA expressed interest in hosting 2024-2025 ICSA Applied Statistics Symposium and three universities in China expressed interest in hosting the ICSA China Conference in future.

Below is a list of planned upcoming events sponsored or co-sponsored by ICSA. We will keep you updated as the COVID-19 situation evolves.

- 2022 ICSA Applied Statistics Symposium, June 19-22, 2022: Chaired by Professor Samuel Wu at University of Florida, Gainesville, FL. Theme: Statistical Innovation in the Era of Artificial Intelligence and Data Science. Keynote speakers: Drs. Susan Murphy (Harvard), David Siegmund (Stanford), and Jianqing Fan (Princeton).
- 2022 ICSA China Conference: July 1-4, 2022 at Xi'an University of Finance and Economics, Xi'an, China. Scientific Program Committee Co-Chairs: Professor Yingying Fan at fanyingy@marshall.usc.edu and Professor Chunjie Wang at wangchunjie@ccut.edu.cn.
- 2022 12^{th} ICSA International Conference: December 18-20, 2022 at the Chinese University of Hong Kong. Program Committee Chair: Dr. Tony Sun and Local Committee Chair: Dr. Xin-Yuan Song.
- 2023 ICSA China Conference: To be hosted by the Southwest Jiaotong University, Chengdu, China (Date to be determined).
- ICSA co-sponsored the 2022 IMS Asia Pacific Rim Meeting (January 4-7, 2022, Melbourne, Australia): Postpone to January 2023.
- 2023 ICSA Applied Statistics Symposium: To be hosted by the University of Michigan in Ann Arbor, Michigan (Date to be determined).

If you have any comments or suggestions on ICSA programs, please contact Professor Hulin Wu at the University of Texas Health Science Center at Houston (Hulin.Wu@uth.tmc.edu).

Hulin Wu, PhD,
ICSA Program Committee Chair,
Professor,
University of Texas Health Science Center at Houston.

Report from Statistics in Biosciences (SIBS)

Hongzhe Li and Joan Hu

Statistics in Biosciences (SIBS) is one of the two official journals established by ICSA. The journal currently has three issues a year in print and electronic form. It publishes articles 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 from the following website: https://www.springer.com/journal/12561.

SIBS publishes both regular articles and topic-oriented papers in Special Issues. In 2021 the journal has published one regular issue and one special issue on "Statistics in Microbiome and Metagenomics" guest edited by Dr. Huilin Li at New York University Langone Health. The 8 papers in the special issue cover a variety of topics on analysis of microbiome and metagenomics data. The regular issue includes papers on modeling early phase of COVID-19 outbreak, detection of short segments for copy number variation analysis, sensitivity analysis for Mendelian randomization and several papers in survival analysis.

There are two special issues that are now in preparation for publications in 2021-22. One special issue is on "Machine Learning Algorithms in Genomics and Genetics" guest edited by Dr. Yingying Wei at the Chinese University of Hong Kong. Another issue is guest-edited by Dr. Lanju Zhang at Vertex and Dr. Naitee Ting at Boehringer Ingelheim on "Leveraging External Data to Improve Trial Efficiency". Both issues include excellent papers on a wide range of topics in genomics and clinical trial design and analysis. If you have interests in proposing new special issues for SIBS, please contact the editors-in-chief.

Professor Mei-Cheng Wang at Johns Hopkins Bloomberg School of Public Health finished her term at the Editors-in-Chief in the end of 2020. We would like to thank Mei-Cheng for her selfless dedication to the journal and her outstanding editorship!

Hongzhe Li , PhD,
Professor, Department of Biostatistics and Epidemiology
University of Pennsylvania
Philadelphia, PA 19104, USA.

Joan Hu, PhD,
Professor of Statistics,
Department of Statistics and Actuarial Science,
Simon Fraser University,
Burnaby, BC V5A1S6, Canada.

Introduction for Bin Yu's Honorary Doctorate Award

Dean Jean-Philippe Bonardi

Editorial: Prof. Bin Yu of UC Berkeley has been awarded an Honorary Doctorate from the University of Lausanne (UNIL), Faculty of Business and Economics, in Switzerland. She was honored as "one of the most influential researchers of her time in Statistics and Data Science, for the excellence and impact of her work, and for her major contributions to the development and advancement of machine learning." Bin was interviewed by journalist Nathalie Randin, with an introduction by Dean Jean-Philippe Bonardi of UNIL in French, June 4, 2021. (English translation).

It is a great honor to stand before you today to award an Honorary Doctorate to Professor Bin Yu. This is obviously a difficult task given the breadth and importance of her contributions, but I will nevertheless try.

After obtaining a BS degree in Mathematics in 1984 from Peking University, Professor Bin Yu went on to study Statistics at the University of California, Berkeley, where she obtained a PhD in 1990. Since 2006, she has held the Chancellor's Professorship in the Department of Statistics and the Department of Electrical Engineering and Computer Sciences at Berkeley.

There are several reasons why the Faculty of Business and Economics suggested her name to the University of Lausanne's administration. The first is related to her international reputation, which is nearly unrivalled in her field. Professor Bin Yu is considered one of the most influential statisticians of her time. She has received numerous prizes and awards, including the Guggenheim Fellowship in 2006 and the Elizabeth Scott Award in 2018, a distinction that recognizes a statistician for their efforts in fostering the careers of women in statistics. Bin Yu is also an elected member of several prestigious scientific institutes and societies such as the Institute of Mathematical Statistics, for which she served as President from 2012 to 2015. She was also one of the first action editors of the Journal of Machine Learning Research.

The second reason why we are extremely pleased to award this Honorary Doctorate to Professor Bin Yu is related to the impact of her work in the discipline of Data Science, that will undoubtedly be central to tomorrow's economy. In her research, Professor Bin Yu is actively involved in the development of theoretical foundations for random forests, deep learning, artificial intelligence, interpretable machine learning, and veridical data science, all with the aim of better understanding these models from a mathematical, statistical, and computational perspective. This work contributes to solving scientific problems in many fields, such as neuroscience, genomics and remote sensing.

The final reason why we are so proud to recognize Professor Bin Yu today is because of her character and her openness to others and to the world. Her contributions are far from being confined to the scientific community, and are part of collective efforts to build a better world. In this sense, Professor Bin Yu has recently received significant media coverage, particularly with respect to her work on the prediction of COVID-19 severity in the United States.

So there you have it, reputation, impact, and openness: three things that guide us at the Faculty of Business and Economics. Thank you, Professor Bin Yu, for leading the way. We look forward to working with you for years to come!

I would now like to invite you to watch an interview of Professor Bin Yu by journalist Nathalie Randin.

In Memory of Dr. Kai Fun Yu (1950-2021)

Colin O. Wu, Jiayang Sun, Heping Zhang, Aiyi Liu

Dr. Kai Fun Yu, former chief of the Biometry and Mathematical Statistics Branch (now known as the Biostatistics and Bioinformatics Branch) at the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (NICHD/NIH), and a longtime member of the International Chinese Statistical Association, passed away on July 8, 2021, at age 70 after a 7-month battle with lymphoma. Dr. Yu's death was a big loss to all the friends and colleagues who knew or worked with him, as well as the statistical community. Over a remarkable teaching and research career in statistics, spanning across four decades, Dr. Yu had nurtured a highly reputable biostatistics research group within NIH, mentored several generations of young statisticians and taught countless students. His contributions to research, teaching, and leadership have established a great role model for young statisticians at NIH and academic institutions.

Kai was born in Southern China in 1950 and spent his childhood and teenage years in Hong Kong. He moved to the United States and received his bachelor's degree (Summa Cum Laude) in mathematics from Dartmouth College in 1973 and then his Ph.D. in mathematical statistics under the supervision of Prof. Y.S. Chow from Columbia University in 1978. Kai started his career in academia from 1978 to 1990, working as an assistant professor at Yale University (1978-1984), visiting associate professor at Stanford (1986) and Columbia (1985 and 1986), and associate professor at University of South Carolina (1984-1990), where he supervised the Ph.D. dissertation of Dr. Miin-Shen Yang who in turn had 17 Ph.D. students in Taiwan.[5] Kai had visited Stanford in 1980 and learned Efron's bootstrap before working with Professors Tze Leung Lai and Herbert Robbins. Their joint work, Lai, Robbins and Yu (1983), on estimating the center of a symmetric distribution was published in the Proceedings of the National Academic of Sciences.[3] After he moved to South Carolina, Kai spent significant amount of time in 1985 and 1986 visiting Columbia University and participated in conferences and discussions with other visiting prominent statisticians, including Butch Tsiatis (then at Harvard), Gordon Lan (then at NHLBI), Dan Anbar (then at Abbott Labs) and Nancy Geller (then at Sloan Kettering). Kai also served as a statistical consultant for local and state agencies in Connecticut and South Carolina during the period of 1981 to 1989. In 1990, Kai moved to the federal government as a senior investigator in NICHD/NIH and served as chief of the Biometry and Mathematical Statistics Branch (BMSB) at NICHD between 2000 and 2008. Under his leadership, the BMSB became established as one of the leading groups for statistical methodology in pediatric studies and research in child health and human development. Kai's research reputation and leadership skills have played a key role in establishing and maintaining BMSB as one of the NIH's top biostatistics research groups. After retiring from the NIH in 2008, Kai returned to his academic career and taught statistics and stochastic processes courses in universities in China.

Kai was a longtime member of the American Statistical Association (ASA) and a founding member of the International Chinese Statistical Association (ICSA). He served on multiple committees, institutional review boards, and editorial boards, and contributed immensely to the statistical society. He was a fellow of the ASA and an elected member of the International Statistical Institute (ISI). One of the most notable activities in Kai's committee work was his involvement in NICHD's tenure and promotion committee, in which he contributed to the collegial and productive working environment at NICHD. Many talented statisticians had directly or indirectly benefited from Kai's effort and support at their professional careers.

Kai made fundamental contributions to the theory and applications of probability and statistics. His earlier work in the 1980s focused on the theory and applications of stochastic processes, such as sequential procedures, renewal theorem, distribution functions, and branching processes.[1,2,3,4] Inspired by the biomedical studies at NICHD, Kai gradually shifted his attention in the 1990s and early 2000s to statistical methodology and applications in clinical and epidemiological studies.[6,7,8,9,10] Functioning as a senior member and a leader of the biostatistics group in NICHD, Kai played a critical role in many

collaborative research projects at NICHD. These efforts resulted in important discoveries in child and maternal health.[11,12,13] Kai's efficiency, creativity, and leadership were truly inspiring.

Throughout his long career, Kai made many friends beyond his workplaces. With a warm and humble personality, his generosity, kindness and wisdom touched many of us. He loved to learn new things and share his insights. Talking with him was fun and interesting. He could turn a causal conversation about everyday things in life into a fun and insightful discussion. Those of us who were his colleagues, collaborators and friends will always remember the experience. He will be sorely missed.

Reference

1. Chow, Y.S. and Yu, K.F. (1981). The performance of a sequential procedure for the estimation of the mean. Annals of Statistics, 9, 194-189.
2. Chow, Y.S., Hsiung, C.A. and Yu, K.F. (1983). A renewal theorem and its applications to some sequential procedures. Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete, 64, 241-250.
3. Lai, T.L., Robbins, H. and Yu, K.F. (1983). Adaptive choice of mean and median in estimating the center of a symmetric distribution. Proc. Nat. Acad. Sci. USA, 90, 5803-5806.
4. Chow, Y.S. and Yu, K.F. (1984). Some limit theorems for a subcritical branching process with immigration. Journal of Applied Probability, 21, 50-57.
5. Yang, M.S. (1989). On limit theorems of the fuzzy C-means clustering procedure. Ph.D. thesis, Department of Statistics, University of South Carolina.
6. Troendle, J.F. and Yu, K.F. (1994). A note on testing the Hardy-Weinberg law across strata. Annals of Human Genetics, 58, 397-402.
7. Simons-Morton, B., Crump, A.D., Haynie, D.L., Saylor, K.E., Eitel, P. and Yu, K.F. (1999). Psychosocial, school, and parent factors associated with recent smoking among early adolescent boys and girls. Preventive Medicine, 28, 138-148.
8. Wu, C.O. and Yu, K.F. (2002). Nonparametric varying-coefficient models for the analysis of longitudinal data. International Statistical Review, 70, 373-393.
9. Liu, A., Wu, C., Yu, K.F. and Gehan, E. (2005). Supplementary analysis of probabilities at the termination of a group sequential phase II trial. Statistics in Medicine, 24, 1009-1027.
10. Troendle, J.F. and Yu, K.F. (2006). Likelihood approaches to the nonparametric two-sample problem for right-censored data. Statistics in Medicine, 25, 2284-2298.
11. Levine, R.J., Lam, C., Qian, C., Yu, K.F., Maynard, S.E., Sachs, B.P., Sibai, B.M., Epstein, F.H., Romero, R., Thadhani, R. and Karumanchi, S.A. for the CPEP Study Group (2006). Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. New England Journal of Medicine, 355, 992-1005.
12. Brotman, R.M., Klebanoff, M.A.,Nansel, T.R.,Andrews, W.W., Schwebke, J.R., Zhang, J., Yu, K.F., Zenilman, J.M. and Scharfstein, D.O. (2008). A longitudinal study of vaginal douching and bacterial vaginosis – A marginal structural modeling analysis. American Journal of Epidemiology, 168, 188-196.
13. Levine, R.J., Vatten, L.J., Horowitz, G.L., Qian, C., Romundstad, P.R., Yu, K.F., Hollenberg, A.N., Hellevik, A.I., Asvold, B.O. and Karumanchi, S.A. (2009). Preeclampsia, soluble Fms-like Tyrosine Kinase 1, and the risk of reduced thyroid function. British Medical Journal, 339, b4336.

Colin O. Wu, Ph.D.
Mathematical Statistician,
Office of Biostatistics Research, Division of Intramural Research, NHLBI
6705 Rockledge Drive, Bethesda, MD 20892
wuc@nhlbi.nih.gov

Jiayang Sun, Ph.D.
Professor, Chair, and Bernard Dunn Eminent Scholar,
Department of Statistics,
George Mason University,
Fairfax, VA 22030

Heping Zhang, Ph.D.
Susan Dwight Bliss Professor,
Yale School of Public Health,
Director, Collaborative Center for Statistics in Science,
300 George Street, suite 523,
New Haven, CT 06511

In Memory of Dr. Kai Fun Yu

Aiyi Liu, Ph.D.
Senior Investigator,
Biostatistics and Bioinformatics Branch,
Division of Intramural Population Health Research, NICHD,
6710B Rockledge Drive, Bethesda, MD 20817

Memories from friends and colleagues

Kai Fun Fu was my unofficial mentor at the National Institutes of Health (NIH). I started, fresh out of graduate school, at the Biometry and Mathematical Statistics Branch (BMSB) of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, in January 1992. Kai had joined the BMSB in 1990. Kai and I soon became friends and collaborators. Our first paper together was published in 1994. Kai was easy to approach for statistical advice. I learned quickly that if I needed someone to sound out an idea on, Kai was always available. His mathematical abilities outpaced my own, but I was able to understand Kai in a way that often eluded me when communicating with others. Naturally, I relied on Kai to bring clarity to me on many topics. I can't think of anyone other than my thesis advisor who had a bigger impact on my career.

One time I told Kai it was my birthday. He told me that certain pieces of information, once stored in his brain, would always be retained. Every year he would give me birthday wishes on my birthday. Another time, Kai told me of how he came to work at NIH. He said that while an Assistant Professor at the University of South Carolina, he decided he wanted to do more to improve public health. So, he applied for a job at Merck. While interviewing at Merck, he talked with a vice president. The vice president asked him "Why do you want to work at Merck?". Kai said he was interested in improving public health. The vice president said "If you work at Merck, you work for the shareholders. If you want to improve public health, work at NIH". If anyone but Kai had told me this story, I would have suspected it to be fabricated or embellished. With Kai, I immediately knew it to be true. I will miss my friend and mentor.

James Troendle,
Senior Investigator & Deputy Director,
Office of Biostatistics Research, Division of Intramural Research, NHLBI
6705 Rockledge Drive, Bethesda, MD 20892

It is shocking and with great sadness to learn that Kai left us. I had a brief overlap with Kai at NICHD before his retirement. I was then in the Epidemiology branch but frequently attended group lunches with Kai's Biostatistics branch. The conversations at the lunch tables were of course spontaneous and broad in topics, but many times circled back to British history which Kai would dive into with tremendous depth. At one point, I couldn't help and asked how he could remember all those names, places, and small details. He told me that he read books about British history whenever he was tired of doing statistics!

Kai was a very proud father. Once I mentioned to him the challenge of raising kids. Kai told me that the best strategy is to let kids grow by themselves. He then told the story of his daughter, about how she was so self-motivated that he did not need to worry a thing about her, how she worked so hard in her high school sport team to the captain position, and how she, as a reward, could park in a designated spot of the school's precious parking lot for the whole school year.

May Kai rest in peace!

Zhen Chen,
Senior Investigator,
Biostatistics and Bioinformatics Branch, NICHD,
6710B Rockledge Drive, Bethesda, MD 20817

One of my fondest memories of Kai Yu were the statistician's tennis games, that we had. A few years after Kai started working at NICHD, he mentioned to us that he used to play tennis, so Ed Korn set up weekend doubles matches at outdoor courts near Ed's house – later going indoors during the colder months. The teams usually consisted of Ed and Kai against me and various partners that we found, which was mainly Joe Zytnick. Kai was a fierce competitor, but it was his style of playing that impressed everyone. Essentially, he covered the entire side of the court by himself with Ed floating around to pick off short shots, which often resulted in points for them. It was a frustrating but an effective strategy. This required a tremendous amount of energy from Kai, that seemed unending.

Besides his prowess in tennis, I was also impressed by how highly principled Kai was in his personal and professional life. He always treated people with great respect even when they disagreed with

him. After our tennis games he and I often would linger behind and talk about all sorts of issues relating to personal matters, our jobs at NIH and politics throughout the world. In these discussions I learned a lot about him and some of his early life in Hong Kong under the unpleasant rule of the British.

Barry I. Graubard
Senior Investigator,
Biostatistics Branch, Division of Cancer Epidemiology and Genetics, NCI,
9609 Medical Center Drive, Rockville, MD 20850

With input from
Edward K. Korn,
Mathematical Statistician,
Biostatistics Branch, Division of Cancer Treatment and Diagnosis, NCI,
9609 Medical Center Drive, Rockville, MD 20850

A little known but charming fact about Kai Fun Yu was that he was a member of a "statistical dynasty." His thesis advisor, Y.S. Chow, turned out to be a statistical matchmaker. Chow introduced Kai to his wife-to-be, his daughter Eunice and introduced TL Lai to his wife-to-be, Eunice's sister! Eunice has a master's degree in statistics and worked for Westat for many years. Kai and Eunice's son Jarvis is also a statistician, working for the VA and their daughter is an economist. There are few examples of three generations of statisticians (so far) in one family.

Kai was a beloved mentor and leader among the NIH statisticians and could always be counted on for advice and wise council. He will be sorely missed by all of us.

Nancy L. Geller,
Director, Office of Biostatistics Research,
Division of Intramural Research, NHLBI,
Bethesda, MD 20892

In his over 20 years as a senior investigator at NICHD, Dr. Yu made important contributions to longitudinal modeling methodology and theory. He made important scientific contributions to the understanding of preventive strategies for prevention of pre-eclampsia as the statistical investigator on the Calcium for Pre-eclampsia Prevention (CPEP). Dr. Yu led the Biostatistics Branch at NICHD from 2000-2008. He was a leader who was loved and admired by branch members. Dr. Yu always put junior researchers ahead of his own research agenda. During his 8 years as branch chief, he successfully led two branch investigators through the laborious NIH intramural tenure process. Aside from his academic and leadership accomplishments, Dr. Yu loved the pleasures of life including food. For years, he organized a monthly restaurant club that visited different restaurant across the region.

Paul S. Albert
Senior Investigator & Cheif,
Biostatistics Branch, Division of Cancer Epidemiology and Genetics, NCI,
9609 Medical Center Drive, Rockville, MD 20850

An Interview with Naitee Ting

Qiqi Deng, Boehringer Ingelheim Pharmaceuticals*
Joseph C. Cappelleri, Pfizer Inc.
**Email:* qiqi.deng@boehringer-ingelheim.com

Summary

Dr. Naitee Ting is currently a director in the biostatistics group of Boehringer Ingelheim. Naitee earned his Ph.D. in statistics in 1987 from Colorado State University. He received an M.S. degree from Mississippi State University (1979, statistics) and a B.S. degree from College of Chinese Culture (1976, forestry). He was with Pfizer Inc. between 1987 and 2009. His research interests include dose response in clinical trials, variance component models, and clinical study designs. Naitee has published over 40 articles either in peer reviewed journals or book chapters. In addition, he has edited the books "Dose Finding in Drug Development", "Applied Statistics in Biomedicine and Clinical Trials Design", and "Design and Analysis of Subgroups with Biopharmaceutical Applications." Naitee has also co-authored two books "Fundamental Concepts for New Clinical Trialists" and "Phase II Clinical Development of New Drugs." He served as an associate editor for the Journal of Biopharmaceutical Statistics and on the editori board for Therapeutic Innovation & Regulatory Science (formerly Drug Information Journal). He is an adjunct professor of Statistics at the University of Connecticut and Columbia University. Naitee has been a Fellow of American Statistical Association since 2014.

In 1987, Naitee joined the International Chinese Statistical Association (ICSA) as a lifetime member when the organization first allowed it. He prepared the section proposal when formulating the Biometric Section within ICSA back in 1990s and was the first secretary of the ICSA Biometrics Section. He is also one of the key contributor for establishing the ICSA's applied journal Statistics in Biosciences. Naitee served as executive director of ICSA between Jan 1998 and Dec 2000. Naitee has actively participated in the organizing committee of multiple ICSA symposia over the years. He co-chaired the Applied Statistics Symposium in 1998 and chaired the 2015 Joint ICSA and Graybill Conference.

Partial supports are gratefully acknowledged from ONR grant N00014-17-1-2176, NSF grants DMS-1613002, and IIS 1741340, and the Center for Science of Information (CSoI), a US NSF Science and Technology Center, under grant agreement CCF-0939370.

PART I: Upbringing and Education

1)Where did you grow up? Tell us about your upbringing.

I was born in 1953 and raised in Taiwan, Republic of China. I have never been a good student. In those years at Taiwan, primary schools ranked students in each class. Class sizes were above 40 or 50, and my rank in the class was never below 40. I became accustomed to being overweight since very young and my reactions were slow. Consequently, I was never good at any sports. I was slow in many aspects,which discouraged me from reading and writing. In my primary school years, homework assignments were mostly about reading and writing. Being on the slow side, I did not do well. Because my grades were not good, I disliked homework assignments. In my primary school, like elsewhere, if a student had good grades, he or she was a good student. For those who did not earn good grades, they try to prove themselves by playing sports. I was good at neither.

I grew up in a big family- grandparents from my father's side, my parents, and six in my generation (four sisters, one brother, and myself). I grew up in a family of 10 members, and I have been a troublemaker at home. In April of 1964, while I was in fifth grade, my family decided that I was too much to handle, and they sent me to live with my grand uncle "三爷", who is the younger brother of my paternal grandfather. This move was one of the most important turning points in my life. My grand uncle completely changed me for the better. He attended a "normal" school for education (a school to train teachers) and became a teacher before World War II. During WWII, he participated in the war fighting against Japanese soldiers. He escaped to Taiwan by himself during the Chinese Civil War between communists and Kuomintang (KMT, the Chinese Nationalist Party).

Because 三爷 used to be a teacher, he knows how to educate children, especially recalcitrant teenagers like I was. Between 1964 and 1971, he brought me up. In those years, 三爷 taught me right from wrong; he educated me and provided me with the value system that I carried for the rest of my life. I graduated from high school in 1971 and left 三爷.

2)Please share your experiences at the universities you attended?

In those years at Taiwan, higher education resources were limited and only selected students could enter college. In 1971 I participated in a "college entrance exam" and failed. From the summer of 1971 to the next summer, I was not allowed to attend college. I went to cram school (补习班) to prepare for the next exam in 1972. During that year, I felt inferior and had low self-esteem for a long time –saying to myself, Naitee, you are no good at anything. This realization awaken me to look internally within myself. Starting from 1971, I have been working on myself with an attempt to overcome this psychological deficiency. Today at age 68, people ask me, "What motivates you to continually be productive?" My answer is that I want to prove that I am not totally useless.

The "college entrance exam" is setup in the following way: 1. Before taking the exam, each student needs to specify a list of departments at various universities that they would like to join (rank from top choice to the last choice); 2. All students take the exam in the same two days; 3. Students' grades are ranked from high to low; 4. Put the student with the highest score into the top choice that he/she selected, then move to the student with the second highest score, assign that student into his/her top choice. Repeat this process until each department is filled. In 1972, I completed my wish list and then took the exam. Among the lowest choices on my list was the Department of Forestry, College of Chinese Culture.

I went to that department, became a student with a major in forestry, while clueless about it, and not knowing where my future would be. Throughout my four years in college, I was wondering what kind of job opportunities existed for me. In my junior year, my father asked me to prepare for graduate study in the U.S. That was not my call-I was hoping to stay in Taiwan and participate in the international trade business. However, in a traditional Chinese family, the father's order carried (and carries) a lot of weight. Hence I applied to U.S. universities for graduate study in forestry. Again, because my grades were not good, I did not get into most of the schools I applied to. Luckily, however, I was accepted by Mississippi State University (MSU).

I came to MSU in 1976 to study forestry. In fall 1976, the first semester I arrived at the U.S., I took two statistical courses. Both courses were taught by Professor Charles J. Monlezun, who opened up my eyes to the field of statistics. In 1977, I switched my major into statistics and graduated in 1979 with a master's degree in statistics from MSU.

My first job was to work as a demographer in the Department of Sociology at MSU. I was at that job for about one and a half years until my green card application was turned down by the U.S. immigration services. So I changed job and went to work as a computer programmer at Jackson Park Hospital in Chicago. I was in Chicago between 1981 and 1984. By 1983, I have been a programmer for over two years, and I had to learn new programming languages every time there is an "upgrade." I began to understand that, in the field of computer programming, I needed to keep learning new computer languages but would not be able to accumulate domain knowledge for professional growth. At this time, I thought that I would need a Ph.D. degree. I applied to Ph.D. programs and was accepted by Colorado State University (CSU). I started my Ph.D. program in January of 1984.

In September of 1987, I completed my Ph.D. in statistics from CSU and then joined Pfizer Inc. In 2009, I left Pfizer and joined Boehringer Ingelheim Pharmaceuticals Inc (B.I.). I have been with B.I. up to the present.

Given the solid training at Colorado State University, I thought that I had a good understanding of mathematical statistics. Now, at Pfizer, I was applying my knowledge to solve real-world problems. However, during the first few years at Pfizer, I was still struggling because the thought process for me to view a problem was grounded in more mathematical- based thinking from my recent Ph.D. training. In dealing with real-world problems, however, the viewpoint is more practical and conceptual - and less mathematical. Many years later, I gradually realized the struggles I went through during my Ph.D. years (when my mathematical training was initially subpar) and my earlier years at Pfizer (where the emphasis was on the applied and conceptual, less on the mathematical).

From the real world going to mathematics is a process of abstract thinking. Math is the queen of sciences because it is deductive. All other sciences like biology, chemistry or physics depend on observations. Math is pure logical thinking, without the need of any observations. Hence the transition from the real world to math is an abstract process. On the other hand, from math to the real world is an application process. One example is that physics is applied math, dynamics is applied physics, material dynamics is applied dynamics, and finally engineering is applied material dynamics. Engineering solves real-world problems and builds constructions and products. Similarly, probability is applied math, statistics is applied probability, biostatistics is applied statistics, and finally clinical trials is ap-

plied biostatistics.

My personal experience started as a forestry major when I was looking at the larger picture, namely, the real world without much knowledge of mathematics that lies at the core of science. For any problem on hand, I tended to think from outside in, an abstract thinking process. Then I went to CSU to start my Ph.D. program. In the years I was there, the training was looking at the world from inside out-without solid scientific foundation, the logic will not follow. In other words, if the math is not rigorous, the scientific foundation is not solid, then (figuratively speaking) the building will collapse, the car will not run, and the boat will sink. From my increasingly more advanced Ph.D. training in statistics, I only knew how to look at the real-world applications from the inside out, from the sphere of math to the larger world of reality. But designing and executing clinical trials are practical applications, and statisticians need to communicate with non-statisticians on a regular basis. This inside out way of thinking did not initially work for me in a pharmaceutical context.

Many years later, after obtaining much experience and lessons learned, I finally realized that my huge challenge existedbecause the Ph.D. training forced me to change my habit of thinking from outside in (from real-world applications to mathematical formulations) to inside out (from math to real world). After working on clinical trials, I realized that the environment forced me to change from thinking inside out to thinking outside in again. Both types of training are very valuable to me and have their respective places. Today I have matured and can look at a practical problem from both sides. Things are much clearer to me. I am very grateful for those hard-learned lessons to realize that both perspectives can co-exist and prove valuable.

3)What people were or have been most influential to you?

Countless people helped me in the path of my life –families, teachers, friends, and many strangers. To them, I am forever grateful. As to people who are most influential to me, I can think of at least three – my grand uncle, Professor Charles J. Monlezun, and Professor Franklin A. Graybill (from CSU).

My grand uncle (三爷) made a profound and positive impact in my life. He taught me the way of living, and the multiple meanings of life. Most importantly, he shaped my values. 三爷 was married in China during WWII, and, when the communists took over China in 1948, he escaped to Taiwan by himself, leaving his wife and daughter in China. The years that I was with him was the turning point of my life. In those years, 三爷 lived a very simple life. He taught me the importance of simplicity and how to simplify things that look complicated. 三爷 respects natural resources, turns off lights when he left a room, he buys only necessities and never wasted any penny on things just "nice to have." This lifestyle taught me to live a simple and modest life and conserve natural resources whenever possible.

One huge benefit I learned from 三爷 is "Health is Number One." He exercised on a regular basis and paid close attention to his health status. When I was young, I ate a lot and was overweight. During the years I was with 三爷, he asked me to exercise. He taught me how to live a healthy lifestyle. Since then I have maintained a healthy lifestyle. I hugely benefitted from 三爷's teachings.

In the spring of 1989, I learned that Pfizer employees were invited to swim at a local pool. Soon after trying to swim, I became addicted to this form of exercise. Up to now, I have been swimming every morning for over 30 years. It has become the most important part of my daily routine, a prerequisite to start the day. Every time I travel, I look for hotels with a swimming pool. Swimming gives me a sense of calm. In the morning, while I am in the pool, my mind is very clear. I take this opportunity to reflect on what I did yesterday and to plan my activities for today. I would say that most of my important ideas have come from the swimming pool.

Professor Charles J. Monlezun positively changed my life. In the fall of 1976, after I first arrived at MSU, my professor in the forestry department advised me to take one course in forestry and two statistical courses. It turned out that both statistical courses were taught by Dr. Monlezun. The way he presented the statistical topics stimulated a lot of interests in me. I found this field of science fascinating! I then signed up for another statistical course from Dr. Monlezun in the next spring semester.

In the summer of 1977, I went to Dr. Monlezun and expressed my interest in switching to major in statistics. He enthusiastically helped me with a research assistantship, and I became a student in statistics beginning in the fall of 1977. I received my master's degree in statistics in 1979, under the guidance and mentorship of Professor Monlezun. I am forever grateful for his teaching, his guidance, and his care of students.

The third person who had profound influence on my career is Professor Franklin A. Graybill. I joined the Department of Statistics at Colorado State Uni-

versity as a graduate student in January of 1984. Given my background in forestry, and the fact that I have been out of school for almost five years, I am aware of the fact that my math skills were weak for a Ph.D. program in statistics. But I did not imagine that my psychological preparation would be far from sufficient. The reality was that not only did I lack good mathematical training, but I also worked in non-math field for too long. I did not practice any mathematical thinking. Starting with the first homework, I realized that I needed calculus to do the assignment, and I had almost no background in calculus. While other students were working relatively smoothly on their homework, I dusted off my calculus textbook and started from the definition of derivatives.

In the years I took Ph.D. courses, I signed up for three courses each semester, and each course had weekly homework. On average, I spent 30 hours on each homework assignment and thus spent a total of 90 hours each week just on doing homework. Luckily, I passed the qualifying exam. Then, in fall 1985, I approached Professor Graybill to see if he would take me as his Ph.D. student. Dr. Graybill apparently understood my poor background in math. He told me that "Yes, I take you, and I have a problem for you. You can do it." I never imagined that sentence would change my life.

Given my weak math foundation, when I saw professors in the statistics department doing research work with heavy mathematics, I was convinced that I would not be fully active in statistical research. That is why I spent my entire career in industry; I simply have not had the guts of trying academia. The first research problem that Professor Graybill assigned to me turned out to be a relatively easy problem. I was able to solve that problem and eventually got it published. In 1987, I finished my dissertation and received my Ph.D. degree. Before graduation, Professor Graybill told me to send my dissertation to Professor Rick Burdick of Arizona State University. Later Rick helped convert my dissertation into publishable manuscripts and get them published. Rick and I became very close research collaborators for an extended time. In the late 1980's up to early 1990's, we co-authored at least a half dozen publications by "Ting, Burdick, Graybill."

Dr. Graybill gave me the confidence to do research. His words "I have a problem for you, you can do it" removed my mental block that I was not prepared for a career in research. Many years later, after I published my books and papers about dose finding, my close friends told me that "your contributions in dose finding has nothing to do with Professor Graybill." My response is that "Everything I published should be credited to Dr. Graybill –he removed my mental block."

4)What is your philosophy for life and community?

The two fundamental roots of my values are 1) being honest to myself, and 2) love and care for others. I consider the most difficult thing in life is being honest with myself. In most situations, it is easy to blame other things, especially those out of our own control, instead of looking within our own selves. This distinction becomes very apparent when a project does not go well. Consider a team where something went wrong with its project. The most typical or natural initial answer is, "It is not my fault –I have done my best to help the project succeed." In fact, this would be the best opportunity for me to ask, "What did I do wrong to end up with this result?"

Regarding my first value –being honest to myself - I hope to bring up two experiences. The first is about my smoking habit. I smoked one pack per day between 1971 and 1979. In the 1970's in Taiwan, smoking was very popular. In most social occasions, the conversation starts with sharing a cigarette and lighting it up for each other. In 1971, before entering the college entrance exam, I was under a lot of stress and began to smoke. When I arrived at Mississippi State University in 1976, there were not many smokers in my environment. After switching to the Department of Statistics, I realized that most professors and graduate students are health conscientious. They don't smoke, they exercise regularly. Many professors run and jog. In those years, people asked me when did I plan to quit smoking. My answer was that "I don't see any problem with smoking; many old people are smokers –they still smoke in their 80's and 90's. Smoking does not affect longevity."

In late March of 1979, however, I woke up one day and told myself that with many professors being runners and joggers, I too may try to run or jog. That morning I put on my running shoes and went running. I could not even finish 200 meters. I had difficulty breathing and could not keep up my pace. On my way walking back home, I told myself that I was only 26 years old and thought that I had good health. But in fact my health was pretty bad. I realized that all of these years, I had been simply lying to myself –smoking cigarettes was actually bad to my health. On that day, I decided to quit smoking

forever.

One other example happened in April of 2009. By then, I have been employed by Pfizer Inc. for over 21 years. In the morning of April 21, my boss came to my office and told me that "your position is eliminated." I was 56 years old, and my younger daughter Felicia would start college that fall. I was in a single-income family, the only bread earner of the household. After over 21 years at Pfizer, I now lost my job. That night, I went home, and told my wife that there were three possibilities. The best possibility is that I can find a job, and life continues. The second possibility is that I may find a job but don't know how long it would take, maybe half a year or one year, maybe two to three years. The third possibility raises the question, "Who will hire a jobless 56-year-old guy?" If I were to remain unemployed,, my wife and I can only rely on our savings. I did some quick calculation –if I add up the severance pay, the 401K, the savings and everything, there is a total number. Subtracting my young daughter's college tuition and living expenses, what left for my wife and myself is the dollar amount for the rest of our lives. I told her that, if we were lucky, we may die within 10 to 12 years, which will make us OK financially. If we survive another 30 to 40 years, we will eventually outlive our money.

On April 22, the next morning, I went to swim. By the time I finished swimming, I walked from the pool to the men's locker room. My mind was very clear –Pfizer was facing Lipitor patent expiration in 2011 with a significant loss in annual revenue. Pfizer needed to lay off employees in order for it to survive. When my boss told me the bad news, I asked how many colleagues would be losing jobs this time; he told me that the count was based on a 15% cut. Our team had 40 statisticians, 15% of it is six –a total of six of us would lose our jobs. When I walked out of the pool, I thought that, if the statistical management needs to lay off six people and if my boss decided not to lay me off, he would have to lay off another colleague. To me, all my colleagues are my friends. Instead of seeing my friends losing their job, I rather face this challenge myself. Therefore, on the moment when I was walking to the locker room, I told myself the following: There was no need to complain about Pfizer and my boss, which would be fruitless and counterproductive. I accept this conscious realization and will be moving on in sincerely wishing them the best.

One colleague in upper management at Pfizer told me that there was a sense of bitterness and anger when most people went through the unfortunate experience of job loss. But I did not feel that way. He appreciated that I was not at all bitter or angry. In 2010, I participated in the ICSA Symposium at Indianapolis and met a professional friend. She told me that "Naitee, you've got a very good reputation –nobody heard you say anything bad about Pfizer." Many years later, when I described my experience to another acquaintance, she said, "How come you saw that so quickly? My husband lost his job for over six months now, and he is still asking, "Why me?" After reflecting on this dialogue during my next swim, I realized at that moment when walking out of the swimming pool how important and helpful this assessment was. Because it provided the constructive mindset to close the chapter of my life at Pfizer and begin to concentrate on my next steps. I therefore relished the blessings of the past and embraced the splendors of the present and future.

Throughout my life, I tried to practice "being honest to myself." This practice helped me face one of the harshest challenges in the later part of my life –losing my job at 56 as the sole income earner in a family of four.

In my high school days when I was with my grand uncle (三爷), he often told me that our life mission is to serve –if you can serve one person (yourself), then serve one person. If you can serve 10 people, then serve 10 people. If you can serve more, then serve more. The basic principle is really about love and care for others. That value was ingrained in my mind and this became my purpose of life as my second fundamental value: care and love for others. I think my philosophy of life and community can be simplified into these two values –being honest to myself; love and care for other people.

PART II: Career

5)Tell us about your years working in pharmaceutical industry, including Pfizer and Boehringer Ingelheim.

I hope to discuss this from the following three aspects: 1) my years with Pfizer and Boehringer Ingelheim; 2) my early publications related to dissertation work; and 3) teaching statistical applications in clinical trials and research on dose finding.

1) My years with Pfizer and Boehringer Ingelheim:

I joined Pfizer in September 1987, the first job after my Ph.D. degree. I was with Pfizer until May 2009, for over 21 years. I am most grateful to Pfizer and surely learned very much during those years. I

joined Boehringer Ingelheim (B.I.) in September of 2009. Up to now, I have been with B.I. for over 11 years. If I can summarize my experiences in these two companies, I may say that I learned about clinical development of new drugs during my Pfizer days, and I applied what I learned to mentor young statisticians during my time with B.I.

Of course the two companies have totally different corporate cultures. Pfizer is a publicly owned company, and its chief executive officer has to face stockholders on a quarterly basis. Everyone can find Pfizer's financial status from publicly available documents, which are required by law. On the other hand, B.I. is a privately owned company and the workforce is responsible to the Boehringer family. There are pros and cons for each type of company. Corporate culture is developed based on the fundamental set up of the company structure, along with attitudes of employees who participated in the organization.

When joining Pfizer, I first heard about the term "clinical trials." In all of my previous education, I have never heard that term. Everything related to clinical trial was totally new to me –new drug application (NDA), drug efficacy and drug safety, phases of clinical development, pre-clinical research and development, and so forth. My new world was fascinating! The work about clinical development of new drugs completely captivated me. By the way, I did not know SAS before joining Pfizer, which provided training to advance my SAS skills. In the Pfizer years, my experiences covered most of the major therapeutic areas –cardiovascular, rheumatology, immunology, respiratory, central nervous system, infectious diseases, ophthalmology, and vaccine development. I worked on Phase II and Phase III development, submission processes including advisory committee meetings, label negotiations and drug approval in what was a wonderful set of experiences.

At BI, I first worked in the clinical development for a few years. In those years, I was able to support the development of some cardiovascular studies, and I was mostly involved in respiratory drug programs. In 2015, the vice president of biostatistics and data sciences encouraged me to support pre-clinical research and development. I have been working on the pre-clinical area since then.

In 2014, I was elected as a Fellow of American Statistical Association: "For influential contributions in clinical development of new drugs, especially in dose finding; for prominent contributions in the research of confidence intervals on variance components; for substantial mentoring of students and colleagues; and for prodigious and significant service to statistical communities." For a statistician employed in industry, being an ASA Fellow is a very high honor. This recognition marked the highest point of my career. Thank you, Joe, for nominating and supporting me.

2) My early publications related to dissertation work

I met Dr. David Salsburg (the author of "The Lady Tasting Tea," "How Statistics Revolutionized Science in the Twentieth Century") at Pfizer. He gave me much helpful advice. One very important advice that I received from him was to publish my dissertation work. I followed his advice and worked with Professors Burdick and Graybill to publish my dissertation, which in turn led to extensions of it and more publications. I also sought for opportunities to make presentations of the research work. I not only participated in symposia, conferences and joint statistical meetings, but also in local meetings and university seminars. Dr. Salsburg was a graduate from University of Connecticut (UConn). With his introduction, I went to the statistics department of UConn and presented my dissertation work.

The research topic of my Ph.D. dissertation is about constructing confidence intervals on functions of variance components. While at Pfizer, I continued to work on extensions and applications of this research topic. In my Ph.D. days, I learned about using the "Modified Large Sample (MLS)" approach to construct confidence intervals on the sum and the difference of two variance components. My dissertation was to calculate confidence intervals on a general linear combination of many variance components. After joining Pfizer, in the late 1980's up to the early 1990s, I was able to extend and apply MLS to establish confidence intervals for a ratio of linear combinations of variance components.

In 1997, FDA published a draft guidance on population bioequivalence (PBE) and individual bioequivalence (IBE). This guidance introduced a new set of metics to evaluate bioequivalence. Based on this new metric, my work on ratio of linear combination of variance components provided the best performance as a unique solution to PBE and IBE. It is interesting that 10 years after graduation with my Ph.D. degree, I found a real world application of my dissertation topic. From this draft guidance of FDA, I was able to collaborate and publish a few papers and book chapters about PBE and IBE. I was also able to advise two Ph.D. students working on this topic. In 2003, FDA decided to go back to average bioequivalence (ABE), and the PBE and IBE work became less popular.

During the years I worked on PBE and IBE, I had an opportunity to collaborate with you, Joe. One of Dr. Cappelleri's expertise is psychometrics. Joe told me that my dissertation work is applicable to health measurement scales. Joe introduced the idea of intraclass correlation coefficients (ICC) to me. On this basis, Joe and I collaborated and published multiple articles on how the MLS approach can be applied to construct confidence intervals on various forms of ICC. I later helped to supervise the dissertation of Kyle Gilder who also conducted research on that topic, which we published and co-authored, and who earned his Ph.D. in statistics from the URI.

3) Teaching statistical applications in clinical trials and research on dose finding

In the late 1980's, there were many workshops to train new clinical statisticians. I participated in one of those trainings and dreamed that some dayI can be one of the instructors to train other new clinical statisticians. With this dream in mind, I began to collect materials I learned from projects I worked on. My first opportunity came in 1993when Professor Choudary Hanumara of University of Rhode Island (URI) approached me to see if I can teach a course about clinical trials. Although I was not completely ready, I agreed to teach this course. My plan was to invite other experts from Pfizer to help cover those topics that I did not have much experience with. This strategy worked, and I was able to teach my first course about statistical applications in clinical trials.

Ten years later I received another opportunity. This time the Department of Statistics of University of Connecticut (UConn) invited me to teach such a course. By 2003, I was ready to teach such a course. Over 80% of materials I prepared by myself. I borrowed a few presentation slides from other speakers I met from various workshops or conferences. I have been teaching this course annually for UConn since 2003. In 2015, I began to teach this same course at Department of Biostatistics of Columbia University, with the course name later changed to "Pharmaceutical Statistics."

Another important development of my career at Pfizer also took place in 1993. I did research on the analysis of dose-ranging in clinical trials and concluded that the pharmaceutical industry can do a better job on testing and recommending the right dosages. Falling short of that could result in a very expensive failure in Phase II drug development. I got interested in study designs for dose-ranging clinical trials. I began to learn this subject. In 1998, Professor Hanumara of URI asked me to co-advise a Ph.D. student, Anthony Hamlett. I asked Anthony to work on the problem of designing dose-ranging clinical trials. We developed the concept of binary dose spacing. Anthony graduated with his Ph.D. from URI and published his dissertation. His publication "Dose spacing in early dose response clinical trial designs" (with me as co-author) was published in the Drug Information Journal in 2002. After this publication, Springer approached me to see if I wanted to publish a book about dose-response clinical trials. Subsequently, my first book Dose Finding in Drug Development was published in 2006 by Springer.

These two opportunities –teaching statistical applications in clinical trials and research on dose finding –established my career in pharmaceutical statistics. I received both opportunities while working at Pfizer. I am truly grateful to Pfizer for providing me with a wonderful learning environment and talented people who encouraged me to learn about clinical development of new drugs.

While losing my job at age 56 was initially a traumatic experience, I had the good fortune to join B.I. in September 2009. Obtaining an offer of employment gave me much hope. For this reason, I am very thankful of B.I.

In 2003, I received an invitation to teach a few subjects for a workshop in training new clinical statisticians. This training was sponsored by the Pharmaceutical Education and Research Institue (PERI). Between 2003 and 2009, I taught at PERI annually. In those years, I got to know Dr. Scott Evans (who is now leading the biostatistics group at George Washington University). Scott told me that he had a contract with Taylor and Francis/CRC to publish a book based on our training materials for PERI. I joined Scott and the two of us co-authored the book "Fundamental Concepts for New Clinical Trialists," published in 2015, and it serves as the text book for my clinical trial courses.

In 2009, I received an invitation to serve as a committee member for the Deming Conference. Starting 2010, I hae been actively involved in this conference. During the 2013 Deming Conference, I met Professor Ding-Geng (Din) Chen. Din approached me and said "Naitee, I hope to co-author a book with you." We began to collaborate, and later Shuyen Ho and Joe Cappelleri joined us. The book co-authored by the four of us "Phase II Clinical Development of New Drugs" was published by Springer in 2017. In addition, I also co-edited two other books –"Applied Statistics in Biomedicine and Clinical Trials Design" (Springer, 2015) and "Design

and Analysis of Subgroups with Biopharmaceutical Applications" (Springer, 2020), the latter also involging another partnership with Din, Shuyen and Joe.

As I am getting older, I enjoy more of teaching and mentoring. These activities bring me a very deep joy and satisfaction. Being able to see young talents developed, to see their contributions to the society, to see their hard work turn into success – these are very fulfilling experiences for me and keep me young in spirit. In 2017, I received the opportunity to train new hires for the B.I. biostatistics group and deliver a lecture series to young statisticians, programmers, and scientists.

Today, I believe that fostering intellectual capital is one of the most important activities of any organization, which is only as good as its employees. As mentioned previously, I am thankful to B.I. when I received the employment opportunity. I hope to do my best to pay back to B.I. In addition to my daily project work, I believe I can use my teaching and mentoring to help the company develop young talents, which is a very satisfying and rewarding experience for me.

In summary, my overall experience with Pfizer can be thought of as learning –I learned about clinical trials and drug development, especially about dose finding, and to apply my learning for the benefit of patients worldwide. My experience at BI allows me to apply my learnings to solve issues that B.I. is facing and to mentor and teach young talents.

6)What are some of your major accomplishments? Please list your achievements that you are most proud of?

In this interview I have previously highlighted several of my major achievements. Up to now, though, I consider my most important contribution to the field of pharmaceutical statistics is to have mentored you, Qiqi. It has brought me tremendous satisfaction to see Qiqi's papers published, to see important B.I. trials designed under her guidance –these study designs take fewer unnecessary risks, save investments and resources for B.I., and are more efficient. Qiqi made very critical contributions to the company, to the field of clinical development of new drugs, and to the pharmaceutical industry –she fulfilled my dreams. Every time she makes an important contribution, receives a promotion, or gets recognized for an achievement, I feel very proud of her.

I first met you, Qiqi, in 2009, when you began your career at B.I. after your just earning your Ph.D. degree. I started mentoring you, Qiqi, since I joined B.I. After over eleven years. including the present, I am proud to have been her mentor. Throughout this experience, I realized that I very much enjoy mentoring younger and talented colleagues like Qiqi and many others. Furthermore, her success implies that B.I. as an organization benefits from her direct contributions, along with my efforts to mentor her.

After so many years in industry, I finally understand the happiness of my friends in the academic world. Many times when I talk with my professor friends, they speak of their students' successes with great satisfaction. I now understand most intimately why teaching and mentoring are so rewarding.

7)You are internationally recognized as an expert in dose finding. Can you share your journey in it?

In new drug development, we know that if a drug works, it is because the drug changed the biological system in our body. If it does not change our system, it is a placebo or no better than one. If it changes our system, it can be toxic. This is why all drugs are potentially toxic. While all drugs can be toxic, how come so many drugs are approved and marketed for patient use? Because every drug has a non-toxic dose. In other words, if the dose is too high, it is toxic; if the dose is too low, the drug does not work. Hence, in my opinion, drug development is about finding the right range of doses –not too high, so that it is safe, not too low, so that it works. Nonetheless, dose finding is very difficult in the clinical development process.

One of the main reasons is that clinical trials are very expensive. In order to add one dose in the clinical trial, a reasonable number of patients will need to be randomized to receive that test dose. The cost to study that dose can easily go up to hundreds of thousand dollars or could be in the millions. Therefore, when the wrong doses are studied, the amount of waste in resources and investments can be prohibitive. That said, in early Phase II clinical development, there is only limited information available regarding the study drug. Hence I consider dose finding as the most challenging task in drug development. At every step of drug development (non-clinical, clinical Phases I, II, III, and IV), the understanding of efficacy and safety associated with each study dose is very crucial.

As described previously, my first exposure to dose finding was in 1993. My first book on dose finding was published in 2006. This book helped establish my career in dose finding. My interest in

this research topic continues up to now.

8)We know you love teaching. And your courses on statistics in clinical trials are very popular. Can you tell me about your teaching experience? What drives you to continue teaching?

When working as a computer programmer in Chicago, I was hoping to get some extra income. I realized that with a master's degree in statistics, I may be able to teach at a community college. I began to teach in the early 1980's, before entering my Ph.D. program. During my Ph.D. years, I worked as a teaching assistant (TA) for a few semesters. After starting Pfizer, I taught part-time at Eastern Connecticut University and URI. Starting 2003, I became an adjunct lecturer at UConn and began teaching (and continue to teach) about statistical applications in clinical trials.

As mentioned earlier, when I was first exposed to clinical trials and drug development, I fell in love with this profession. As early as the late 1980's, I was hoping that I would be able to teach this topic sometime in the future. URI offered me the first opportunity to try it out and then UConn provided me with an enduring relationship with academia. To me, it is an honor to be able to teach materials I organized based on my work experiences at a prestigious university like UConn. With the extensive teaching experiences from UConn, I was able to obtain another opportunity to teach courses offered by the Department of Biostatistics at Columbia University. Currently, I teach three courses for Columbia University each year: 1)Pharmaceutical Statistics (the course based on the book I co-authored with Scott Evans) in the spring semester, 2) Categorical Data Analysis in the summer semester, and 3) Regression Analysis in the fall semester. As such, I am proud to be an adjunct professor at an Ivy League school with a long and respected history in statistics and biostatistics. It is a double honor for me to able to teach at UConn and Columbia.

Shortly after I began to teach at URI, I mentored summer interns while at Pfizer. During the mentoring process, I lectured on my teaching materials and those lectures were well received. In 2008, a junior colleague at Pfizer asked me whether I would teach him, and I did. That brought me the idea to teach and mentor new hires. I have been mentoring since then.

PART III: Networking and ICSA

9)What motivates you to actively engage in activities of ICSA and related activities?

I remembered that I went to one of my first statistical conferences in the late 1970's and felt a strong feeling that I did not belong to the group. One reason was that my English was not that good, and I felt myself as a foreigner, an outsider. Another reason was that I was new to the field of statistics, and most of the presentations I did not understand. I thought that all I had to do was y to complete my master's degree, find a job, settle down, and that I would never have to go to another statistical conference in the future. Clearly I had a proffessional identity crisis at that time point, as a Chinese in a foreign country and not sure what my career would be.

I first heard about a Chinese statistical organization while working on my master's degree at MSU. The name of this organization was the Chinese Statistical Association in America (CSAA). Later, I became a demographer and a computer programmer. I slowly lost contact with the statistics profession. In 1987, the year I obtained my Ph.D. degree, is also the year ICSA was founded. In the mid-to-late 1980's, before ICSA, the only regular meeting of CSAA members was the Wednesday dinner during the Joint Statistical Meetings (JSM). By 1987, I have already participated in several such Wednesday dinners during JSM. At that time, I was looking for some identity and level of security. I felt ICSA provided both. In addition, I also had much curiosity about this new organization and decided volunteer for ICSA.

After getting more involved with volunteering for ICSA activities and events, I learned that there are so many members working very hard to make this organization successful. Their devotion to ICSA had a tremendous positive impact on me. During the early 1990's, I made many wonderful friends and colleagues from ICSA and developed an emotional link to the people who contributed to this organization. This link naturally extended beyond ICSA to related organizations and activities. After deeply involved in many ICSA activities, I learned how a professional organization operates and how individuals connect with each other while playing specific roles within this organization. This understanding largely broadened my view of my profession and career and also how I work with other people.

In order to maintain and grow a strong organization, we need to constantly recruit and develop young talents. Throughout my years volunteering in

ICSA, we successfully recruited and developed many excellent officers and board members. This experience provided a positive feedback loop: I learned from them and they learned from me. ICSA offered opportunities for me to develop myself; in turn, I offered opportunities for young talents to develop themselves. They then created opportunities for me to make contributions to this professional organization. Today, I enjoy the benefit of knowing many founding members of ICSA (the earliest cohort), my own generation of members and officers, as well as the younger cohort of active members and future officers.

It is only natural that I extend my experiences towards developing young talents from ICSA to day-to-day life in teaching and mentoring. Today, many of my students and mentees have become successful in their own ways. It is very fulfilling to watch how they climbed their career ladder and how they have overcome their challenges. Of course, I will continue to teach and mentor.

10)Tell us more about your special engagement with ICSA

In 1987, I joined ICSA as a lifetime member when the organization first allowed it. I participated in ICSA activities and began to make friends. As mentioned above, before the official organization of ICSA, the only regular meeting of Chinese statisticians was the Wednesday dinner during JSM. Through those dinner meetings, I got to know some of the founders of the newly formed ICSA. I looked for opportunities to volunteer myself in ICSA events and activities. My first opportunity came from its Applied Statistics Symposium.

In 1990, Dr. Jia-Yeong Tsay chaired the first Applied Statistics Symposium at Bethesda, Maryland. It was a very successful event, and we were hoping it would become an annual event. For some reason, there was no symposium in 1991. Then, in 1992, Dr. Shein-Chung Chow chaired the next symposium. During those years, some of us discussed ways to continue this symposium in the future and make it an annual event. One suggestion was to formulate a Biometric Section within ICSA and make this section responsible for future symposia. According to the Constitution and By-Laws of ICSA, a section can be formed if over 10% of members endorse a section proposal. Around that time, the ICSA active membership was about 500. To obtain 50 or more signatures was quite doable and then achieved.

I volunteered myself to prepare the section proposal. I also participated in the organizing committee of the 1993 and 1994 ICSA symposium. We set up face-to-face meetings to discuss the Biometrics Section, put the proposal together, and obtained sufficient signatures. The proposal was submitted to ICSA, and it was discussed and passed at the 1993 Board Meeting. Now that this section is official, and a new tradition of ICSA to sponsor annual Applied Statistics Symposium became formalized. I became the first secretary of the ICSA Biometrics Section.

In 1997, I was considered one of the candidates for the next executive director of ICSA, which required serving a three-year term between January 1998 and December 2000. Later I received this honor and began my key role in this organization. As executive director, I was able to communicate with each member directly –to update membership and to communicate between members and the ICSA Board. I worked closely with the ICSA's past and present presidents as well as its board members. These three years were the highlight of my involvement in ICSA. I was able to develop a lot of close friendships during that time. In 1998, under Dr. Gordon Lan's encouragement, he and I co-chaired the Applied Statistics Symposium, which took place at Connecticut College, New London, CT.

One of the reasons for the founding of ICSA as a professional organization was to launch Statistica Sinica as its official journal. Until then, Statistica Sinica has been the flagship journal of ICSA and a well-established journal in theoretical statistics. As applied statisticians working in industry or government, however, many of us found it difficult for our papers to be published in Statistica Sinica. In 2003, Dr. Frank Shen, then ICSA president-elect, approached me and asked to see whether I can start an applied journal for ICSA. I told Frank that this is not likely to happen for the following two reasons: 1) this attempt failed in the past, and 2) I am not good at working on journals.

There were at least two times in the past that the ICSA president tried to start an applied journal. Both times were voted down by the ICSA Board. Hence the first thing I need to do was to find out why the board members did not think an applied journal in ICSA was a good idea. I began to call senior members and board directors of ICSA. After about 30 or 40 phone calls, I realized that the concern was that ICSA tried very hard to develop Statistica Sinica into a top journal in statistics and that there may not be sufficient manpower to support two journals. In 2004, I spoke with Professor Xiao-Li Meng of Harvard University who was then co-editor of Statistica Sinica. Xiao-Li told me that Statistica Sinica was well established by then, with a

very high turndown rate of submitted manuscripts. On this basis, the timing was right –it was about time to start an applied journal within ICSA.

I approached Professor Xihong Lin of Harvard University to discuss the establishment of an applied journal for ICSA. Xihong and I first formed a committee to evaluate the feasibility of a new journal. Xihong proposed the name of the new journal to be Statistics in Biosciences. She also drafted the mission statement of this journal. The committee put together a proposal and submitted it to ICSA. During the 2006 ICSA symposium, the Board approved this proposal and a new committee was organized to establish this new journal. We first debated the pros and cons of working with a publisher. Then we approached a few publishers and eventually chose Springer which offered a very nice contract with ICSA in 2007. The new committee agreed to appoint two co-editors for this journa, one from academia and another from non-academia. Statistics in Biosciences published its first issue in 2009, with Professor Xihong Lin and Dr. Gordon Lan as the founding co-editors.

I chaired the 2008 Graybill Conference at Fort Collins, Colorado. Because this conference went well, I was encouraged to chair another Graybill Conference, this time in 2015. Given that the upcoming Graybill Conference and ICSA were both scheduled for June 2015, I suggested having a Joint ICSA and Graybill Conference. As a result, this joint conference took place in 2015. Chairing the 2015 Joint ICSA and Graybill Conference was my last major activity for ICSA and among the high points of my involvement with ICSA.

PART IV: Future

11) You have devoted yourself in pharmaceutical statistics for years. Nowadays, there are many options for statisticians to choose from. Can you talk about which part of pharmaceutical statistics is most attractive to you?

I obtained my Ph.D. degree in 1987 and decided to join industry to apply my statistical training in solving real-world diseases that could potentially affect millions of people worldwide. I was fortunate to have many on-site interviews with pharmaceutical companies. I accepted the offer from Pfizer and started in September 1987. Before joining Pfizer, I had no exposure to clinical trials. Shortly after joining Pfizer, I found that clinical development of new drugs was a fascinating career –there was so much to learn and the impact was so huge. If we can successfully develop a good drug, it could potentially benefit millions up to billions of patients globally. This fact provided me with a mission and sense of responsibility of being in a profession of helping patients on a very large scale.

In our daily work, we have close interactions with medical researchers and practioners trained as medical doctors (physicians). These interactions helped me to learn and understand medical conditions. Over the years, I have become familiar with how statisticians work efficiently with physicians. In order to explain statistics to non-statisticians, I have been forced myself to think hard and go into the depths of the fundamentals of statistics. Only after developing a good understanding of the basic concepts of statistics can I then talk about statistical methods in layman's terms. After this practice, I realized that, although I hope to help patients, I do not work with patients directly. The best way for me to make a useful contribution is to work closely with physicians and, the better they understand statistics, the more they can help or communicate with patients. Hence close communication between statisticians and physicians is absolutely critical in clinical development of new drugs.

The more I am in this profession, the better I can appreciate the importance of clinical trial design. Most of the statistical work is about data analysis. But no matter how good the analysis is performed, nothing can save a bad study design. The business model in pharmaceutical industry is that, for new drug development, it is high investment, high risk, and high potential return. Bad study designs not only waste valuable investments and resources, but also subject the sponsor to unnecessary high risks. One of the most important contributions a statistician makes to the project team is to help quantify risks as part of the overall benefit-risk assessment. Good clinical trial designs help non-statisticians understand the risks and benefits being taken as crucial elements in a benefit-risk assessment. Therefore, my research interest in the past two decades has been about clinical trial designs. My focus has been, in particular, on Phase II study designs.

12) How would you like to spend the rest of your career?

I will continue to do my best for BI. As mentioned earlier, BI offered me an employment opportunity at a crucial time. I am truly grateful to BI and would like to continue making valuable contri-

butions to the company. For now, I am still with very good health and really enjoy my daily work. Therefore, I plan to continue working for BI in the upcoming years.

Recently Dr. Jingjing Ye from BeiGene indicated to me that the Biopharmaceutcal Section under American Statistical Association (ASA) initiated a YouTube channel activity. This was set up in early 2019 to create an online platform to benefit the membership in addition to the Section's webinar series, podcast series and online training program. Jingjing suggested that I record my lectures on this program. I hope to complete this program in the near future.

The Department of Statistics of Colorado State University will celebrate its 50th anniversary in 2021. I was invited to teach a clinical trial course for this department remotely. This is a great honor for me to teach at the department where I received my Ph.D. With this opportunity, I will become an adjunct professor for three fabulous schools –Columbia University, University of Connecticut, and Colorado State University.

Over time, I gradually realized that my passion is about teaching and mentoring. Currently I have the pleasure of teaching at the three universities and mentoring younger colleagues at BI. I really cannot ask for more beyond these wonderful opportunities. For the rest of my career, I will continue to teach and to mentor.

13) What advice would you give for younger or junior statisticians?

My first advice is to build a dream. Be practical and be realistic in pursuing your dreams. Think of yourself five years in the future and then 10, 20, 30, and 40 years in the future. Whom do you want to be? What kind of person do you hope to develop yourself into? Who is your role model? What makes your life meaningful? What dreams are high enough that deserve your complete effort with a lifelong pursuit, yet practical enough that will be achievable? Dreams can be big or small, can be long-term or short-term. No matter what it is, always have a dream. Of course, in order for the dream to be practical, we must know ourselves. Knowing oneself is one of the most important things to do in life. We need to know ourselves and to build a dream for ourselves. People went to school because they want to graduate. Whether it is a graduation from high school or college, these accomplishments are all practical dreams of students at that time.

However, after graduation, people typically focus on making a living and may gradually may lose their dreams. Therefore, my advice is to consider your daily work not as part of a job but as building blocks of your career. One of the differences between a job and a career is about ownership –imagine the work you do is not for the team, not for the organization, but for yourself. Think of it as your own business. Try to think far ahead of your life. Again, whom do you want to be? What steps will you take to prepare yourself in order to realize your dream?

My second advice to younger or junior statisticians is to practice speaking and to practice writing. For most applied statisticians, the daily work is mostly achieved by reading, writing, listening, and speaking. Yet our statistical training emphasizes only reading and listening. For most of us with a graduate degree in statistics or biostatistics, our education involves very little speaking and writing. In the real-world setting, most projects are accomplished by teams. A typical project team has only one statistician, and this statistician has to communicate with other team members. These situations can be simplified as "statistical consulting." In the current environment, an effective applied statistician in industry is not someone who can derive formulas or prove theorems. An effective statistician is one who can communicate statistical thinking with non-statisticians.

In fact, the practice of speaking and writing is not only about communication; it is also the best way to organize your thoughts. Many times we study a problem and believe that we have a good understanding of it. When such is the case, consider speaking about the topic in public or writing about it to educate readers who are not familiar with it. Once we prepare to speak or write about it, we realized that there are many details we did not fully grasp. The best training for speaking and writing is to publish your ideas whenever possible as well as to talk about them. Some people are intimidated about writing a publication because they thought it is difficult to come up with a good idea. In fact, ideas do not have to be great or a certain size; everyone has lots of ideas. The more difficult part is to organize your idea so that you can talk and write about it. In other words, coming up with ideas is not difficult; putting them in words and sharing them with the public is the challenge.

Hence it is important to practice speaking and to practice writing. It is not about practicing to come up with ideas. In public speaking and writing, practice is very helpful. Practices will organically help to coalesce and elevate the cumulative experiences. We can improve from repeated practices. Speaking

and writing are very useful tools to help you succeed. For a very talented person with a lot of great ideas, if this person cannot express it, then nobody knows his or her talent. On the other hand, for an ordinary person like me, if I can speak and write about my idea or my understanding of the problem, then people may believe that I may have something to offer. Thus, for a successful career as an applied statistician, it is not only about how good your statistical knowledge is, but also about how good your speaking and writing skills are.

My third advice is to be flexible and to be adaptive. As the British naturalist Charles Darwin stated "It is not the strongest of the species that survives, nor the most intelligent; it is the one most adaptable to change." For our generation and future generations, one thing that can be certain is that changes will happen. Some changes happen based on our request –to apply for a school, to apply for a job, to move to a better place, to get married, to have children, and so forth. Most other changes happen out of our control. In industry, such change frequently occurs in the form of reorganization or restructuring. When reorganization happens, some of us may change our roles; some may lose our jobs. In these challenging times, adaptability would help us to live through the tough times. When learning to be more adaptive, you become more flexible, and it will be easier to adjust to and flourish in a new environment.

Role of Biostatistics in the Big Data World of Modern Biomedical Research

Jeffrey S. Morris

In recent decades, biomedical research and population health have been completely transformed by the development of hypersensitive technologies taking automated measurements of quantities previously unmeasurable, including but not limited to genetics and multi-platform genomics, imaging, and wearable devices, and the assembly of large population-level databases facilitated by the broad adoption of electronic medical records across health systems. Such developments produce volumes of complex, often structured and high-dimensional data. These data contain a treasure trove of information about the underlying diseases and patient health with potential to lead to transformational breakthroughs, but raise significant informatic and analytical challenges, many not handled by traditional and commonly-used approaches. The ability to integrate information across these data as well as other more traditional data types, and to extract and translate reproducible knowledge from them, is central to modern efforts to devise impactful and transformational new risk modeling, prevention, and precision therapeutic strategies, and if not done effectively may be the limiting factor holding back progress in biomedical research.

Quantitative sciences have always been important in biomedical research, but the explosion of available data have placed quantitative scientific areas into an even more central and strategic role in biomedical and population health research. The group of quantitative scientists working in biomedical research is multidisciplinary, including biostatisticians, bioinformaticians, epidemiologists, biomathematicians, computational biologists, and computer scientists, who together comprise a broader field that might be called *biomedical data science*.

Data science has become a major entity in modern society with the advent of enormous internet-level data and the transformation of society into a digital world. This field exploded with the advent of targeted advertising, which has transformed the business world and raised technology companies to become the most valuable and lucrative in the world. These opportunities and developments have accelerated the development and dissemination of big data analytical tools in the fields of machine learning and artificial intelligent, and made data science a fundamentally important field.

The big data of biomedical research bring additional analytical and inferential challenges not present in the internet-level data sparking the data science revolution. Making treatment decisions for patients has a lot more at stake than decisions of

what ad to flash on a person's social media page, so the cost-benefit decision is more complex and level of evidence required to make data-driven decision is higher. This makes decision theory, uncertainty quantification, and statistical inference crucial; point estimates and predictions do not suffice. The inherent biases present in observational data make it difficult to infer causal insights from them, and clinical trials, while enabling causal inference if randomized, are often not representative of population level care so can have limited usefulness. And important biological mechanistic information is only available for deep study in animal or cell line models, whose relevance for human health is unclear. These make integration across disparate data types a crucial, unique problem. The disparate medical data systems and requirement for privacy make assembly of large population-level data sets a daunting and sometimes impossible task, fraught with missing data, measurement error, and raising numerous other complex analytical challenges. These challenges and others make *biomedical data science* its own unique field, containing some of the elements of data science but also other challenges that define the scientific endeavor to understand and improve human health.

The skill set and perspective of biostatisticians makes them central players in biomedical data science, especially in the big data world.

- Biostatisticians are experts in *inferential thinking*, which enables scientists to collect and analyze data in a careful way to enable the discovery of scientific truth and population insights from sampled data while accounting for potential sampling-induced biases and providing careful uncertainty quantification, and ensuring reproducibility, replicability, and generalizability.
- The underpinnings of biostatistical training and inferential thinking is an *understanding of randomness and variability*, which enables biostatisticians to be uniquely equipped to lead in areas of experimental design, uncertainty quantification, and building models able to account for design-induced systematic biases, handle structure inherent to complex data including correlation induced from subsampling, cluster designs, and longitudinal or spatial sampling, and adjust for multiplicities that produce false discoveries when not properly taken into account.
- Biostatisticians are inherently *collaborative*, spending a large amount of time working in team science environments with numerous scientists and clinicians, with experience and skill in discerning the key elements of the underlying science to elicit and guide the quantitative aspects of hypotheses, study design and analysis.

Based on these core expertise areas, biostatistics plays a fundamental role in working with biomedical and population scientists in hypothesis refinement, experimental design, analysis, and uncertainty quantification to ensure their studies produce carefully-interpreted reproducible insights. In this process, biostatisticians identify and provide novel solutions to unsolved quantitative problems, including innovative designs and analytical approaches that are applied to collaborators' data and disseminated into the broader biomedical research and population science communities. We also play an important role in educating trainees in other areas of science and medicine in quantitative and inferential principles so they are equipped to properly understand the insights revealed by their data, as well as the limitations and uncertainties so they can precisely communicate what is known and what is yet to be learned.

These roles are accentuated by the emergence of big data in biomedical research, including multiplatform genomics, imaging, and wearable device data, sometimes all present together, and increasingly available in large biobanks enabled by electronic medical records. These produce rich data containing many valuable insights about the underlying biological and medical processes, but also many spurious associations that might appear to be clinically relevant but would not be reproduced. Careful analytical thought is necessary to efficiently extract the translatable knowledge contained in these data, with new methods sometimes necessary to extract insights missed by simple standard approaches that are often reductionistic and discard information contained in the big, complex data. The various types of data contain complementary information, and innovative integrative methods are necessary to synthesize information across these modalities and reveal the complete picture they provide. With big data, it is especially crucial that these analyses adjust for complicated issues such as multiple testing, batch correction, and bias adjustment for causal inference in order to ensure the insights are reproducible. Efficient learning from these big data is the key challenge in devising modern precision therapeutics that can be customized to the particular characteristics of the individual's disease.

Given the centrality of our expertise to biomedi-

cal research, biostatisticians should see themselves as quantitative scientific leaders, not just technicians. While biostatistics as a field provides broad quantitative support to scientists and clinicians, biostatisticians will increasingly be recognized by collaborators as scientific co-leaders and leaders in the crucial quantitative elements of biomedical and population research. This leadership is realized through proactive promotion of sound experimental design, best analytical procedures, and reproducible research, the development and application of innovative new design and analytical techniques motivated by the substantive scientific questions, and the discovery and validation of new translatable biomedical knowledge through careful secondary analyses of existing big data resources. We are crucial players in efforts to help ensure research is done in the most rigorous, robust, and reproducible manner, and we often need to take a proactive, not reactive, role in interdisciplinary research teams to exercise this leadership.

Biostatisticians have always had a central role in biomedical and population health research, but as biomedical research has entered the big data era, our role has become even more essential. In order for medicine to take the next step forward, it needs to leverage the rich information contained in these big, complex data sets, which they cannot do without us. We need to step up to meet this challenge, and identify and solve the most important unsolved quantitative problems, engage as quantitative leaders on multidisciplinary research teams, and disseminate our work and provide tools to ensure our impact.

Jeffrey S. Morris, Ph.D.
Professor,
Director, Division of Biostatistics,
Department of Biostatistics, Epidemiology and Informatics,
Department of Statistics,
University of Pennsylvania

Data, Data, Everywhere⋯

Nicole Lazar

Editorial: This is a reprint from the Harvard Data Science Review in the *HDSR website* https://hdsr.mitpress.mit.edu/pub/yrzpcrus/release/2) with permission.

It is a common saying that we live now in a data-driven age. But what does this actually mean? There are many buzz words to describe this new reality: data science, big data, machine learning, even artificial intelligence-which is not exactly new but has gained new life in recent years. It's true that there are more data at our fingertips than ever before, with smartphones (and other so-called 'smart' devices, everything from watches to refrigerators) collecting information about our location and activities; web browsers that store (and pass on to others) our search history; and social media that track our interactions-and this is just the tip of the iceberg!

Indeed, it is an interesting time to be a data scientist or statistician (and some may ask: what is the difference between these two? More on this later.). In addition to the personal data that most of us generate on a daily basis, advances in medicine, psychology, biology, and the social sciences have all added to the steady stream of opportunity and challenge for people who are 'data literate' (another buzz word) and 'uncertainty literate.' We are truly awash in data in a way that would not have been imaginable not that long ago; along with that comes an increased need to make decisions under uncertainty, something that is a part of everyday life whether we like it or not. Even within the last five years, the data landscape has shifted significantly. Such rapid change is unprecedented in the history of statistics and data. For most of the past hundred years, the span of time in which many statistical methods were developed, data were small and relatively structured. That is no longer the case. Data are more varied and vaster than ever before, requiring the development of new ways of analysis-hence some of the buzz words that are bandied about in the popular press and scientific literature alike.

This is both exciting and overwhelming, especially if you are new to data science as a field, or have heard about it but are not really sure what it is. What, then, do people mean when they talk about 'data science'? For a simple starting point, you can think of data science as a big term for the fields that work with data, in an effort to learn what they can tell us about a phenomenon of interest and how they can guide us through its inherent uncertainty. This may include statistics-a central player

in data science-as well as computer science, applied math, and some areas of engineering, depending on whom you ask.

In this column, and future columns in the Minding the Future series, the aim is to introduce the big picture of data science, particularly to younger readers, parents, and teachers, and to demystify what can, at times, seem like a confusing array of terminology and concepts. Along the way, we hope to help you develop into more data literate and uncertainty literate consumers of information, a key part of being an educated citizen of the modern world.

Let's Talk About Data⋯ An important first step is to understand the different types of data available, their meanings and uses, and also to be aware of how data can be abused and misused. 'Data for good' is another term that has gained popularity in recent years. A group of students at the University of Georgia, where I teach, started an organization dedicated to using their data analysis skills for the betterment of society, completely unaware that this is a bigger, worldwide movement. The data for good movement has several aims: to increase the ethical use of data in the public sphere; to raise awareness of how data are used to make decisions that may be biased against certain groups of people (as in the popular book Weapons of Math Destruction); and to point out the ethical implications of big corporations having access to all of our personal information, much of which we freely, though often without realizing, give up to them. After all, who carefully reads the terms of agreement before signing on to a new type of social media, online game, or convenient website?

Governments also routinely collect data on their citizens. For the most part, this is done to assess the needs of the populace and hence to direct services provided by the government to the places where they are needed most: schools, medical care, or infrastructure such as roads and bridges. Yet there is also the concern that, in the wrong hands, such seemingly innocuous data could be used in ways that might cause potentially disproportionate harm to certain groups. An example that received a lot of publicity in the United States this past year dealt with the addition of a citizenship question to the 2020 census. Census data, collected every ten years in the United States, are used to allocate resources as well as seats in the House of Representatives. The concern was that inclusion of a citizenship question would deter some people-specifically non-citizens in this instance-from responding to the census, resulting in undercounting and misallocation of resources. The census is supposed to count all residents, not just citizens, so a severe undercount of that group would have been a real problem.

Who owns your personal data: your shopping habits, your online activities, and your social network? These are data that somebody else collects about you, maybe through a website or a free app. Current laws do not fully address such questions, though some countries have started to consider issues of confidentiality and data collection. For example the European Union (EU) instituted the General Data Protection Regulation (GDPR) in May 2018. The regulation protects data privacy within the EU and the European Economic Area, giving individuals more control over their personal data and regulating the transfer of such data outside the EU and EEA. A similar law, the California Consumer Privacy Act, took effect in California on January 1, 2020. The law allows users to see what sorts of data companies collect on them, and gives them the ability to refuse companies the rights to use or sell those data. Also in the US, former Democratic presidential candidate Andrew Yang-noting that data are more valuable than oil-has stated that tech companies such as Facebook and Amazon should pay people for the use of their data. Who might take advantage of such an offer, and who might value their rights to privacy and to own their own data, more than any monetary gain? While people will weigh the options differently and hence reach different conclusions, it is not hard to imagine that some groups taken as a whole might be more willing to sell their data compared to other groups.

It is important for modern data analysts to think about questions of this type, and that is before one even gets to the data themselves. There is a myriad of interesting questions that can be explored using approaches from statistics, computer science, and the basic sciences: what we are broadly calling data science.

⋯And Collect Some Data Examples A few examples show how wide the span is. Imaging technologies have given neuroscientists new insights into how the human brain functions. Techniques such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), and others generate massive amounts of data on the brain in action. Such data are not only big (terabytes in size) but complicated in ways that pose challenges for the data analyst. The interesting brain images that you might see on the cover of a popular science magazine purporting to show differences between healthy individuals and those with

some disease, are the output of sophisticated statistics or machine learning algorithms that take years to fully master.

Similar advances in genetics have led to a revolution in genomics data, whereby scientists can study in previously-impossible detail the very makeup of the human (or other) being. Services such as 23andMe, which give you a peek at your genetic background, are relatively recent arrivals on the scene; they also rely on data science to build their stories.

The field of digital humanities-which brings statistical analysis, machine learning techniques, and computational methods to the humanities-has a surprisingly long history. Already in 1963, well before the modern computing revolution, statisticians Fred Mosteller and David Wallace used statistical methods to analyze the 12 disputed Federalist Papers. It was long surmised that these 12 essays had each been written either by James Madison or Alexander Hamilton. Mosteller and Wallace used differences in word choice to attribute authorship of the papers to Madison. The discipline of digital humanities has particularly flourished and gained visibility in the last decade. The Stanford Literary Lab, for instance, is a collective of researchers who use computational tools to study literature, focusing on topics as varied as fanfiction and the effects of translation on a text. Digital humanities labs and centers are now found in universities worldwide.

Businesses have also of course been quick to jump on the data bandwagon, as have sports teams, who use increasingly sophisticated data collection and analysis tools to improve their records (think the book and movie Moneyball). Many of these corporations have dedicated data science groups. The prevalence of such was brought home to me recently, when our department hosted a visitor from The Home Depot, a senior member of their data science team. For many of us, it was news that The Home Depot even had a data science team, although it probably shouldn't have been. His talk was enlightening in a number of regards.

First, he described the many ways that a chain that specializes in home repair, home projects, appliances, and supplies uses data science. Though his particular examples were from The Home Depot, most people these days have experience with online shopping, whether for patio furniture, clothes, school supplies, music, or games. The recommendations that pop up after a search, as well as the advertisements that you subsequently see when you go to other websites, are the results of algorithms developed and tested by data science teams.

Second, our guest discussed several algorithms that few of us in the audience had heard of, indicating that new developments continue to unfold at a rapid pace. As data scientists, there are always going to be new things for us to learn, since new types of data and new data analysis problems require different solutions to what worked in the past. The field has seen this before, with machine learning, for example, and we can anticipate that it will continue to be the case.

Third, he emphasized the interdisciplinary nature of many data science teams. The people he works with come from a variety of backgrounds: statistics, computer science, and machine learning, to name a few. Data science teams may also include applied mathematicians, engineers, and domain experts. There is a lot of overlap among the different disciplines, and even to those of us working in the worlds of data, the distinctions between data science and statistics in particular are not always clear.

A popular description goes something like this: 'A data scientist knows more computing than a typical statistician and more statistics than a typical computer scientist, but can communicate better than either of them.' This somewhat tongue-in-cheek summary does, however, highlight important ideas for those who are interested in learning more. Mathematical skills are crucial, as is knowledge of computing (programming, database management, and the like). Communication refers to being able to explain what you did, be it in writing, orally, through the use of data visualizations, or other means. In addition, data scientists often have a serious, in-depth engagement with a data domain: neuroscience, genetics, politics, engineering, business, literature, or meteorology, to name a few.

What about Uncertainty? Clearly, an understanding of data and how they are used is important for the modern citizen. Schools and universities hence are increasingly emphasizing the idea of data literacy for all their students. Equally important is to think about uncertainty, both in the data (which are almost always collected with some noise) and in how we use data to make decisions. This issue is playing out in front of all of us in real-time, as the world confronts the COVID-19 pandemic. Governments (and individuals) have had to make decisions about appropriate actions in the face of great uncertainty: how widespread is the virus? how easily does it spread? how lethal is it really? There are many unknowns in the current situation, in part because good data are hard to obtain in the midst of

the crisis itself, and indeed they may never be fully available.

On a more prosaic note, we all need to make decisions with incomplete data sometimes.

As a student, you might be applying to universities early decision. How do you know which is your 'best' choice? What do you mean by 'best' for that matter? How do you decide which date to ask to the prom? How do school districts decide whether to close because of bad weather? In the Southern U.S. where snow is relatively rare and cities are unprepared to deal with it, even a hint of the possibility of snow may lead to school closures. It is easy to laugh when the snow fails to materialize and children are sitting at home for no reason, but think of the other side: schools do not close and a big snowstorm comes to a city with no plows. Sometimes the right decision (what one should have done) is clear in retrospect; other times, it will never be known. Grappling with these facts, understanding them, and not expecting or demanding solid answers in every circumstance is the sort of uncertainty literacy that is important in a world where (often incomplete, messy) data play a critical role.

Minding the Future With all of these different types of data and data-related questions, it is no surprise that data scientists are in high professional demand. People who can analyze, interpret, and present data in meaningful, understandable ways will play an ever-increasing role in science, business, education, government agencies, and industries of all variety - the horizons are near-limitless. Likewise, those who can help society more generally work through the ethical implications and complications of our data-rich world should have an important voice in protecting the public from abuses and misuses of their personal information. The realms of data science and its constituents are vast, and you are invited to explore them. The goal of Minding the Future is to be a guide on this journey: to demystify and decode some of the buzz words; to inform about important and often confusing concepts; and to provide fun activities that you can try at home or in school to deepen your own data science knowledge. We welcome questions from readers, and especially suggestions for topics you would like to see covered in future columns.

References

1. Lewis, M. (2003). Moneyball: The Art of Winning an Unfair Game, New York: W.W. Norton & Company.

2. Mosteller, F. and Wallace, D.L. (1963). Inference in an authorship problem. Journal of the American Statistical Association, 58, 275-309.

3. O'Neal, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, New York: Crown Books.

4. Sonnemaker, T. (2019). Andrew Yang wants you to make money off your data by making it your personal property. Business Insider, Nov. 14 2019. https://www.businessinsider.com/andrew-yang-data-ownership-property-right-policy-2019-11

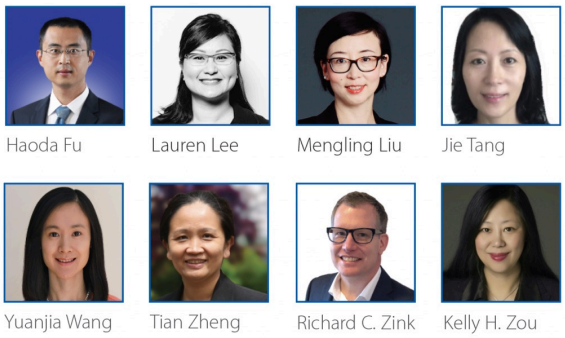
5. Yang, A. (2020). Data as a property right. https://www.yang2020.com/policies/data-property-right/

Nicole Lazar, Ph.D.
Professor of Statistics,
University of Georgia, USA.

ICSA Symposium Panelists Offer Leadership Advice

Haoda Fu, Eli Lilly and Company; Lauren Lee, Pfizer Inc.; Mengling Liu, New York University; Jie Tang, Lotus Clinical Research LLC; Yuanjia Wang, Columbia University; Tian Zheng, Columbia University, Richard C. Zink, Lexitas Pharma Services, and Kelly H. Zou, Viatris

Editorial: Permission: This article is reprinted with permission from Amstat News, April 1, 2021. URL of the original article:https://magazine.amstat.org/blog/2021/04/01/icsa.



As invited session panelists for the 2020 International Chinese Statistical Association (ICSA) Applied Statistics Symposium, eight statisticians and data scientists were asked to share their thoughts and advice on leadership, with a special focus on statistics and data science. The following is a summary of their answers to the questions about how to lead, negotiate, and deal with a pandemic.

Why and how is leadership about empowerment?

Empowerment is a key determinant of effective leadership in both manager-team dynamics and self-empowerment. First, empowerment from a leader to the team often takes the form of substantiated support to team members who are taking initiative on high priorities in the form of resources and mentoring.

Second, empowerment comes from within. Leaders who are confident in their abilities can lead with assurance, kindness, generosity, and aspiration. Leaders who are not threatened by others can trust their leadership journeys.

Third, empowerment is trust. The first job of leaders is to inspire trust, because trust is the single most essential element of our ability to deliver extraordinary outcomes in a long-lasting way. Consequently, trust likely leads to success because it enables an organization to work in a psychologically safe environment and in its culture.

How do you negotiate professionally while showing respect?

We use negotiation in our daily lives, whether knowingly or subconsciously. Winning a debate is not a good way to think about and approach negotiation. There are several insightful books that can help shed light on this topic. We would like to recommend Getting Past No: Negotiating in Difficult Situations by William Ury. The author was a co-founder of Harvard's program on negotiation and one of the world's leading experts on negotiation and mediation. One technique he writes about is reframing the problem from face-to-face confrontation to side-by-side problem solving. Showing respect is important, but it doesn't mean you need to be soft. But where is the line? The book showcased a key principle: soft on people and hard on questions.

Another recommended book is Crucial Conversations: Tools for Talking When Stakes Are High by Kerry Patterson and Joseph Grenny. Often, we have too many assumptions before conducting a conversation or negotiation. Thus, it is important for us to first listen and understand before conveying our ideas to others.

Just like in all career sectors, negotiations happen frequently in academia. In many cases, compromises must be made so everyone benefits. When there are shared interests, it is easier to reach agreement. It is helpful to bear in mind what everyone can offer as well as what their constraints are. The perspective of viewing the negotiating parties as complementary, rather than conflicting, members is helpful. For junior investigators, it takes courage and patience to become seasoned at negotiation. Constantly building one's strength will lead to successful negotiations.

What are effective leadership skills in statistics/data science?

According to the Effective Statistician Leadership Course Program, one of the most useful activities to spend time on is developing a personal leadership statement and list of principles by a leader. First,

any interaction should be approached with humility. This humility should be applied to direct reports since they are immersed in the details of their tasks and their success leads to the success of the department. It also applies to nonstatisticians, so we can effectively communicate both simple and complex concepts.

For a leader, communication should be applied early and often. Often, people make assumptions about what others know, understand, or agree with. One reason could be that people hesitate to ask questions about the basics so they do not look timid or foolish. However, these questions can lead to meaningful and necessary discussions and actions. Statisticians should seek responsibility and take responsibility. Do not be the individual on the side of the room who does not ask questions or offer advice. Come to meetings prepared with at least one question and, if there is the opportunity to volunteer for a task or to find a solution, take that opportunity! An obvious skill is to use data to guide decisions.

Finally, management does not automatically equal leadership, and people of all levels can exhibit leadership qualities. Leadership is a talent, but it is also a learned skillset through self-awareness, technical capability, and people-centricity.

Why is it important to embrace diversity, equity, and inclusion, including women?

Alondra Nelson, deputy director for science and society at the US Office of Science & Technology Policy, said, "[W]hen we provide inputs to the algorithm···we are making human choices, choices that bring our social world to bear in a new and powerful way. ···I've always sought to understand the perspective of people and communities who are not usually in the room when the inputs are made, but who live with the outputs." We are missing out on our chance to reach our goal of excellence if "people in the room" do not reflect the diversity of "people out of the room."

People's backgrounds, experiences, knowledge, skillsets, mindsets, and views are diverse. When we embrace these, we can essentially broaden ourselves in all these aspects. Without diversity, equity, and inclusion, we won't know what we don't know. To embrace diversity, equity, and inclusion, we should engage all members of our community in a conversation about our shared vision of excellence and examine how much of this vision has not been realized in some underrepresented parts of our community. Through such a conversation, we can recognize the need for more proactive efforts in the direction of diversity, equity, and inclusion.

What are some effective ways to foster and encourage collaborations?

Leaders should be willing to collaborate, roll up their sleeves, and dive into the weeds, no matter how tough situations become. A gritty mentality lets the team know we're working for them and are always willing to jump into the mix. It is most important to frame problems in a language everyone can understand, otherwise not everyone is starting out in the same place. One effective way to do this is to limit jargon as much as possible. Delegate and assign tasks that are enjoyable and completed on time with high quality.

Technology-oriented and methodology-savvy, we must be proactive in offering our solicited and sometimes unsolicited expertise. Although we may directly develop questionnaires and deal with health care providers, we are acutely aware of different types of survey instruments and patient-reported outcomes. Set up ongoing meetings to regularly check in or have face-to-face meetings with key individuals to develop better professional relationships on a personal level. The bottom line is communication is key to effective collaboration, and it makes everyone feel like a partner with something to gain.

Successful collaborations stem from getting people to agree on several common goals and work hard to reach them. On such initiatives, many decisions need to be made about the aims, choices of analytic methods, and interpretations of the results. Structured discussions can be helpful to facilitate collaborations. Keep organized and engage members. When there are competing interests, skills for conflict resolution are helpful. For large initiatives and collaborations in particular, breaking complex problems into subproblems and working in smaller groups is effective. Last, having a few highly reliable, active, and engaged members to lead can be the key to success.

How can leaders rise to challenges during a pandemic?

There have been work challenges and life challenges during the pandemic. Leaders who can provide critical resources to overcome work challenges and genuine support to overcome life challenges are necessary. To assume leadership roles and deal with the current and future pandemic, training in communication with the media and public to convey uncertainty and statistical concepts could be especially helpful.

Statisticians have contributed to major efforts on multiple fronts, including modeling the epidemiology and spread of infectious diseases such as

COVID-19, designing and analyzing vaccine trials and therapeutic trials, and studying risk factors and prognostic factors of the virus. We have seen groups pivoting their efforts to quickly turn to research during the pandemic, actively communicating with local governments to inform policy decisions, and organizing online symposiums and seminars to disseminate information. For example, a research group is part of the COVID ensemble modeling hub, which provides weekly forecasts to the US Centers for Disease Control and Prevention to inform the public. The mental health of health care workers due to COVID-19 is also evaluated.

How can we be change-agile professionally these days?

Given the global nature of workforces, a nine-to-five job may no longer be the norm. Working from home with a more flexible schedule has become the trend. As much as possible, take opportunities to practice! This is especially important when using new technologies (e.g., Zoom, Teams, RingCentral) or when using a major new feature from a familiar technology, particularly if important customers or potential customers will be involved. The laissez faire style can be useful to empower colleagues and mentees. For new and less experienced staff, coaching may work better. It is helpful to ensure the coaching and mentoring process is a two-way interaction, where careful and empathic listening comes before sharing or guiding.

With most interactions being virtual, there are several opportunities for training webinars and online meetups. Take the opportunity to learn new topics such as "estimands," even if the concept sounds unfamiliar or not for immediate use. Perhaps set up a mentoring group to take courses or discuss new topics together. Brief check-in meetings with colleagues may provide opportunities to offer expertise early on.

What are some good tips for communicating with teams virtually?

Digitalization has accelerated in a virtual environment. Nevertheless, there are pros and cons of virtual interactions. We need to realize that many of our messages are delivered instantaneously. Our gestures, facial expressions, and body language are harder to read via a two-dimensional screen. Be present, avoid multitasking, and turn on video during virtual conferencing to fully engage with others. Having a separate section in your home for virtual work helps, as well as communicating in ways other than web conference calls, such as Teams.

While on videoconferences, we may leverage technology for efficiency. The virtual background function can be used in multiple apps. It can help deliver messages in different situations. For example, setting up an agenda or providing the meeting objectives.

The virtual work environment may save a significant amount of time in terms of meeting transitions and commuting to work. On the flip side, it may make us more stationary. Instead of having back-to-back long meetings, consider shorter meetings with reasonable breaks in between.

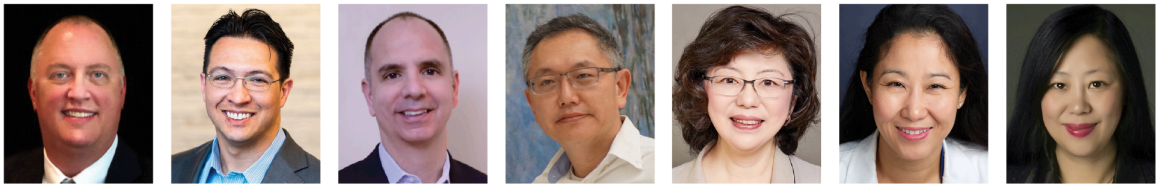
Panel Provides Insights on Real-World Evidence in Health Care

Joseph P. Cook, Viatris; Aaron Galaznik, Acorn AI / Medidata Solutions; Joseph S. Imperato; Max Ma, Johnson & Johnson; Jun Su, Astellas; May Yamada-Lifton, SAS; and Kelly H. Zou, Viatris

Editorial: Permission: This article is reprinted with permission from Amstat News, June 1, 2021. URL of the original article: https://magazine.amstat.org/blog/2021/06/01/panel-rwe.

Real-world data (RWD) can come from a variety of sources, including claims, electronic health records, biobanks, genomics tests, and imaging modalities. Increasingly, it is coming from digital data through electronic health and mobile health modalities. Recent events have pushed innovation in the use of RWD to maximize the value of real-world evidence (RWE) in an era of big data, data science, and artificial intelligence (AI).

Given the COVID-19 pandemic, digital innovation, patient preference assessments, and electronic patient-reported outcomes are likely to see increased use. RWE is also increasingly used to help identify patients for randomized clinical trials (RCTs), optimize RCT design, and help optimize evidence pack-



ages to accelerate the approval and reimbursement processes.

A panel was held during the 2020 International Chinese Statistical Association's Applied Statistics Symposium with the goal of providing insights into the current trends and future outlook for RWE. What follows is a summary of the key topics discussed.

Technology: Health Innovations

With respect to RWD, people most commonly recognize insurance claims or electronic medical records. Often used for health services research, they are administrative byproducts of care delivery repurposed for research. In addition to commonly used RWD such as pharmacy claims and EMR, there are other data sources that can be leveraged, including chart reviews, registries, patient-reported outcomes (PROs), biobanks, genomics tests, and imaging modalities. Specific database needs are driven by what study questions investigators aim to address, with each type of data providing unique value.

The study design and statistical methods considered should be those that best address the study objectives. One should consider if the approach should be descriptive, causal-comparative, quasi-experimental, or experimental, as well as the appropriate cohort, whether a prospective or retrospective observational study. Further, one should ask if case-control, cross-sectional, or case report/series should be used. What is the optimal statistical approach? Is there a role for meta-analytic or predictive modeling methods?

Electronic health records (EHR) data-linkage with claim databases, registries, PROs, and surveys are seen more frequently to address specific research questions. Increasingly, it is possible to link these sources of data to increase the richness of what they can provide. We are beginning to figure out how to use genomics, wearables, consumer data, and even social media as new data sources.

RWE is being widely used to gain an understanding of patient populations and subpopulations, as well as the patient journey and when and how treatments are used and any resulting gaps in care. Using this knowledge, RWE can be leveraged across the lifecycle of drug development, including planning and early development, as well as business and commercial activities, including market access, health technology assessments (HTAs), contracting, or tenders.

RWD can serve a crucial role in helping payors understand the financial impact of new treatments for their specific cost structures and in their populations. It can also help bridge the efficacy-effectiveness divide when seeking to understand how new treatments' clinical trial results generalize to their populations. HTAs routinely seek to estimate the incremental economic impacts to a health care system or insurer using models employing RCT data. Direct measurement, however, of the achieved cost burden and experienced cost-effectiveness for a given health care system is assessed using data from a real-world environment.

Anyone who has worked with RWD knows much of the work is in preparing the data and deriving meaningful variables. Companies have found success in employing AI to enhance data anomaly detection, standardization, and quality checking at this pre-processing stage. Rigor and transparency around how data is then transformed and in how machine learning (ML) is applied will help increase trust and understanding of where and how to employ ML effectively. Improved data linkage and interoperability will be needed to provide the real-time feedback loops in RWD necessary to unleash the potential of AI for clinical decision-making.

The key to data access and linkage is interoperability. No single health system has all the data, and there is an increased tendency to do federated analysis to deal with privacy issues. The difficult part is that not all analytics are adapted or suited to federated analysis. We need a shared system based on trust. For example, what SAS learned during the support for the opioid crisis in Massachusetts was that laws are sometimes required-Chapter 55 was put into place to interconnect multiple databases into law and that helped inform state policy and a program to manage overdose-based mortality.

Technology is already entering the health care industry robustly. Starting to integrate information from wearables and diagnostic devices has the potential to significantly increase the reliability and comprehensiveness of electronic health records. For

example, Apple has its health app and the Apple Watch, and Google is adding Fitbit to its holdings. Moreover, we see new ways of doing business with many stakeholders looking for ways to partner that will help encourage greater value for patients. There are many startups and other organizations working on this, but having groups collaborating on data sharing, quality, transparency, and collaboration can help bring structure and order to these innovations.

Applications: RCTs and RWE

RWE and RCT data complement each other. RCT data is generated within a controlled experiment to tease out the incremental benefit of a therapy in a defined setting. RWD is necessary to understand what happens when therapies are deployed in real life. It permits understanding effectiveness in a broader range of patient types, in a larger group than is achievable in an RCT setting. Thus, it complements the internal validity of RCTs with the external validity of RWD. We think by using the two side-by-side, one can better translate the clinical-trial-to-real-world divide, as well as contextualize the representativeness of trial data populations in a broader real-world context.

While RWD can be incredibly rich and varied, it can also be messy and challenging to tease out a signal from background noise. Mining this data, whether with old-fashioned data mining techniques or incorporating AI tools, can be quite challenging. AI is just a tool of data mining. The latest deep reinforcement learning and graph neural network developments show great potential to mine the data with good depth. Cutting-edge changes occur rapidly in technology spaces. However, devices of many types are being developed to help patients better track their health and, if shared, can provide additional richness.

For now, RWE complements RCTs. The holy grail is to not just extend medication indications and labels, but to also get an approval for new medications using RWE sources, as it has the potential to be cheaper and to better account for real-world practice than conducting an RCT. We can have a better understanding of efficacy and effectiveness of medicines in patients with this approach.

It would be ideal to create a feedback loop where "patients like this get treated this way/that way" for clinical decision support that optimizes the outcomes. For this, we would need to define clear goals such as quality of life or cost-effectiveness.

Synthetic control arms and RCTs with real-world data sources are also intriguing. SAS can support these applications, but to do so requires partnership between many stakeholders. Focusing on fundamentals such as data sharing, quality, transparency, standardized processes, and imputation methods for use of RWE would increase confidence in its usage.

The quality and availability of RWD are improving exponentially, providing more reliable data for analysis and generating RWE. In addition, advances in statistical algorithms continue to improve our ability to leverage RWE for the inferential statistics and hypothesis testing required by regulatory agencies around the world.

Outlook: Aftermath of the Pandemic

The COVID-19 pandemic has accelerated drug development and trial and manufacturing processes at our major customers, and RWD is being captured actively in many countries. Scientific breakthroughs in the future should be faster if we leverage this experience to tackle regulatory processes, incorporate new data sources, and leverage emerging analytics and technology infrastructure. There is a push for greater racial diversity in the patient populations studied and analyzed for COVID-19, but this has not been the case in many historical trials.

We know many people don't live in areas with access to clinical trials, and social determinants of health are key drivers of treatment success. We must pay attention to and eliminate biases from our data sources and models. Secondary data collected for other purposes can introduce collection bias. For example, using claims data already means you are working with a subset of patients who are working. Heightened awareness of this bias in data is a good thing for the future of RWE.

The pandemic has spurred dramatic changes to market access conditions for patients that will have lasting positive effects. For example, to help boost adherence with better access, patients were increasingly allowed to get 90-day prescriptions, use mail-order and home delivery, experience lower out-of-pocket costs at the register for insulin and COVID-19–related health care, and gain improved access to telemedicine and chronic care medications.

In terms of predictions in the aftermath of the pandemic, the health care markets will increasingly find ways to use the available RWE to shape the way markets work, but not without limits. There will be a blurring of the distinction between retrospective and prospective data gathering, both with respect to different RWD types and in linking RWD to RCT data. Innovation in integrated data collection and comprehensive evidence generation will be used to

gain insight into the real world, to inform stakeholders' understanding, and to improve patients' lives. The confidence in data quality and increase in data sharing, integration, and transparency will mean greater uptake of RWE. Automating access to RWD to gain RWE will continue to drive our decision-making from intuition to insight.

However, there may be both surprises and disappointments in our future, despite the hope to gain more accurate and reliable results using AI and algorithms. Thus, explainable AI may be important to understand and interpret the results, as well as provide forecasts. We are in early days of using AI for medicine; data standards are still being defined, governance of models could be improved, and users may not understand and trust the results. From this perspective, the failure rate of AI projects can be high. AI requires a combination of disciplines: science, engineering, statistics, math, and biology. Some even say data science is an art due to its exploratory nature and need to convince humans.

Ultimately, AI projects should focus on business value, not AI value. Life is not an AI reality talent show. AI solutions should be no-brainers for end users to incorporate as part of their workflow and not be standalone solutions. Thus, fit-for-purpose data, sound methodologies, and impactful applications go hand-in-hand when dealing with RWD and big data.

XL-Files: Time Travel and Dark Data

Xiao-Li Meng

Editorial: This is a reprint from a column article published in the *IMS Bulletin* (Volume 49, Issue 1: January/February 2020; https://imstat.org/wp-content/uploads/2019/12/Bulletin49_1.pdf) with IMS' permission.

"We should stop lamenting how other professions repackage our methods, and start doing it ourselves properly, to better engage the broader data science community and beyond".

As might have been anticipated (jinxed?) by my thesis title, "Towards complete results for some incomplete problems," self-pity for being incomplete has never left me. This is as true now as it was back when I accidentally reduced my almost complete thesis to merely its title, exactly 10 days before it was due. All 12 LaTeX files, one for each draft chapter, displayed zero bits on that almost fatal morning, after a 2 a.m. attempt at creating a backup copy reversed its direction. A painful lesson learned: data augmentation with sleepy or closed eyes should not be attempted. But God obviously had more lessons for me: a DVI file was left for me to build imputation. Imputation is never perfect, but I did graduate in time.

Since that time, imputation has become a source of self-help whenever my feeling of incompleteness fails to entertain itself, for reasons that are known or, otherwise, in need of rational imputation themselves. I imputed unobserved data, biased responses, latent variables, counterfactual fantasies, hidden agendas, implied ideologies, unspoken threats, suspicious motivations, and of course, the hardest of all, blinded wine labels. I imputed sometimes with deep satisfaction, and other times with deep regret. I even multiply imputed.

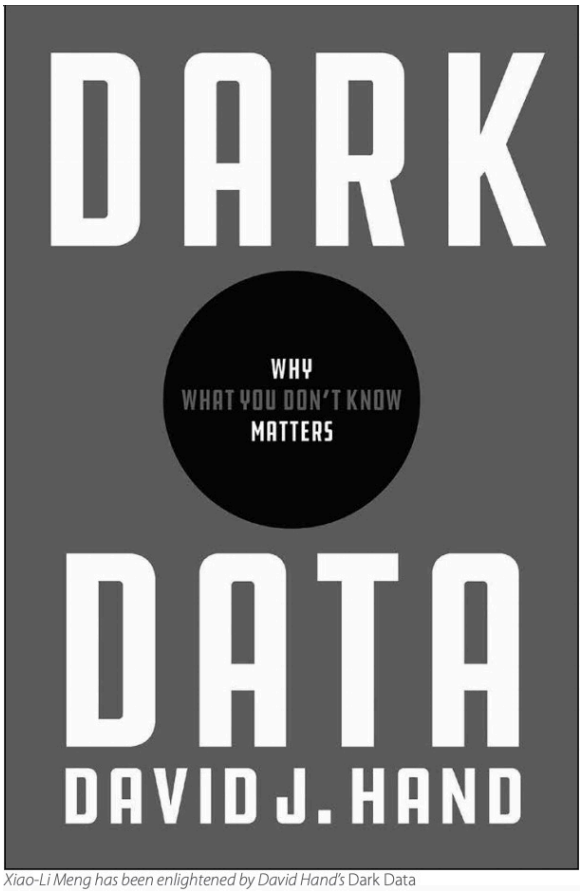
Never in my wildest imagination, however, had I contemplated the possibility that imputation could transform me into a time traveler—never, that is, until I encountered, a few years ago, an ingenious reporter who wrote about a comparative study on voting behaviors of politicians. The researchers couldn't make direct comparisons among all the politicians studied because very few voted on all the legislation: politicians obviously cannot participate in voting before they were elected, or after completing their terms. The researchers therefore built a model to impute what these impossible votes would have been had these politicians been in office at the time of voting. In effect, the reporter observed, the researchers were building a time machine, allowing the politicians to travel back and forth in time to cast their votes.

Regardless how skeptical we are (as we all should be) about the validity of such an imputation model, we should admire the reporter's creativity in coming up with a vivid analogy to arouse the public's curiosity about something rather technical, or at least to remind the reader that there is something here both remarkable and questionable. No statistician has ever used "time travel" to describe imputing counterfactuals, despite it being a rather effective and engaging analogy. Indeed, collectively, we statis-

ticians have done a regrettable job in coming up with rhetorically attractive means of engaging those beyond the already converted. The italic emphasis here is to remind ourselves that "regret" is even a technical term for us!

And this is not the only R-word in our vocabulary. We also have "regression", "risk", "rejection", "residual error", etc.; and speaking of error, we have another rich collection: "type 1 error", "type 2 error", "standard error", "standard deviation", "absolute deviation", "variance", "bias", "mean squared error", and the most depressing of all, "total error"··· I have to wonder how many other fields would knowingly adopt a term that may leave the impression of total wrongness?

Of course, I am as guilty as anyone, for I coined "uncongeniality" as a technical term (initially for describing a thorny issue for multiple imputation, now more broadly for pre-processing).



Xiao-Li Meng has been enlightened by David Hand's Dark Data

Science and statistics are serious businesses, and as such, we should resist any temptation of creating hype terms merely for their soundbite value. At the same time, we have to admit that no matter how much we complain about deep learning without deep understanding, the phrase "deep learning" is far more likely to attract our attention than, say, "multi-layer adaptive non-linear function compositions" or MLANFC.

We should stop lamenting how other professions repackage our methods, and start doing it ourselves properly, to better engage the broader data science community and beyond. This is not an easy task, because most of us are not trained to appreciate the important roles of branding and marketing in scholarly products and dissemination,especially in an era of progressively shorter attention spans. I am therefore particularly excited about my fellow columnist David Hand's (yet another) new book, Dark Data. Right away, without reading any text, you can tell that this is a book about data we cannot see but matter. Indeed, David was inspired by dark matter: "Since we can't see this extra mass, it has been called dark matter. And it can be significant (I almost said 'it can matter')." The first time I saw the title, my immediate reaction was to kick myself for being so incomplete –how could I have never thought about such a catchy and apt term, especially given my years of messing with missing data, non-responses, and latent variables, all forms of dark data???

I calmed my statistical ego down (sadly) by comforting myself with the thought that, "Well, this must be another CS term." I googled and found the term indeed has been used in the CS community, but it was used exchangeably with "dusty data." Hats off once more to my CS friends, for "dusty data" is another clever and vivid term, which describes data that are never processed or analyzed, effectively making their collection an expensive process for gathering dust.

However, "dark" and "dusty" are not exchangeable, semantically or visually. David's use of "dark data" is much more appropriate and comprehensive, despite his emphasis that his list of types of dark data is necessarily incomplete. David discusses 15 types of dark data, and why and in what ways they matter. He shows that they must be dealt with even if they are invisible (especially to untrained eyes). In David's taxonomy and notation, the various forms and conditions of dark data are as follows:

DD-Type 1: Data We Know Are Missing
DD-Type 2: Data We Don't Know Are Missing
DD-Type 3: Choosing Just Some Cases
DD-Type 4: Self-Selection
DD-Type 5: Missing What Matters
DD-Type 6: Data Which Might Have Been
DD-Type 7: Changes with Time
DD-Type 8: Definitions of Data
DD-Type 9: Summaries of Data

DD-Type 10: Measurement Error and Uncertainty

DD-Type 11: Feedback and Gaming

DD-Type 12: Information Asymmetry

DD-Type 13: Intentionally Darkened Data

DD-Type 14: Fabricated and Synthetic Data

DD-Type 15: Extrapolating Beyond Your Data

Because I initially thought that David's notion of "dark data" only covers the kind of missing observations or incomplete data to which statisticians commonly refer, I didn't fully appreciate some items on this list, for example, Type 11 or 12, on their own. I wouldn't be surprised if the list generates a similar feeling for you.

But this is why you need to read the book, and be convinced by David's reasoning and his examples of cases in which unseen or unreported data play a critical and sometimes even a fatal role. You are likely to walk away with the feeling that the term dark data is indeed a very effective one to arouse both curiosity and suspicion, mixed with happiness that finally a great term was coined by a statistician —and sadness that the statistician is not you.

Oh, whereas I probably don't want to be relabelled as a Dark Data Scientist, I'm enlightened by David's dark data, and believe my years of imputation practice can shed some light on the dark matter revealed in David's book. And I am sure you can too, unless, of course, you prefer to be a dusty (dark?) statistician···

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

Teaching

Hans Rudolf K'unsch

We all started as pupils and students and then at some point switched to the role of a teacher, either as teaching assistant or as professor. So you know in principle both sides of the coin and can build on your own experience: Avoid mistakes that some of your teachers made and follow the example of the good teachers you encountered. However, all human beings are different. A teacher who was ideal for you maybe failed to connect with other students, and you cannot emulate easily someone else's teaching style, ignoring your personal strengths and weaknesses.

In this column I want to draw your attention to some of the issues in teaching, discussing why, what and how to teach and how to assess the achievements of your teaching efforts. In addition to my own experience it is based on chapters from Paul Halmos' book "I want to be a mathematician" (Springer, 1985) and on "Ungrading: Why Rating Students Undermines Learning (and What to Do Instead)", S. D. Blum, ed., West Virginia Uniersity Press, 2020.

The obvious answer of why to teach is that in academics most jobs require you to teach. Beyond that, teaching is a chance to pass your insight, experience, intuition and enthusiasm about the subjcet on to younger people. However, not everybody is a born teacher and teaching requires in addition to knowing the subject well also other abilities, e.g. to see that something which is obvious to you can be a stumbling block for others. In order to become a good teacher you have to develop an intuition for different ways of thinking and to anticipate difficulties and misunderstandings of your students. This requires time and effort, but as in research you shouldn't give up when you don't succeed immediately.

I also had some experiences with students that seemed to expect their teacher to be an entertainer or an animal tamer rather than a scholar. If this happens to you, meet with a delegation of students, listen to their complaints without immediately defending yourself and then decide what you can change without giving up you convictions about what you think they need to learn. Another option is to ask an experienced colleague whom you trust to sit in your class and give you feedback.

The answer to what to teach is more complex, as it includes not only knowledge and skills, but also attitudes. Knowledge means things like formal definitions, rules and methods, theorems and proofs; examples of skills are communication or the use of statistical software, and the category of attitudes contains for instance motivation, curiosity, perseverance, creativity, work ethics or critical thinking. The art of teaching is to transmit not only knowledge and skills, but to develop at the same time also such attitudes.

Turning to the question of how to teach, lectur-

ing is the most common and easiest way, but it has also been called sarcastically the "best way to get information from teacher's notebook to student's notebook without touching the student's mind" (George Leonard, Education and Ecstasy, 1968). You have to find ways to touch the mind of your students, to get them involved beyond passive listening, to encourage them to ask questions and to challenge them with problems at the right level of difficulty.

In discussions about teaching methods, often the quote "I hear, I forget; I see, I remember; I do, I understand" is mentioned which is attributed to Confucius, although this is not correct (see reference). That visual communication is usually more effective than verbal is also reflected in the statement "A picture is worth a thousand words", but I believe John Tukey once said something like "100 words are needed to fully understand the meaning of a picture". And E. Dijkstra, a pioneer of computer science, said "A picture may be worth a thousand words, a formula is worth a thousand pictures" although formulae need maybe even more words than pictures to be understandable. To me it makes no sense to create a hierarchy between pictures, words and formulae, or between oral and written communication. They should all be used in teaching because what is best depends also a lot on the student's personality and the level of understanding. And "I do" should be interpreted as active involvement and reflection instead of passive reception which is possible also while listening and looking.

Exams and grades are the most common way to assess what the students have learned which is relevant not only for the student, but also for the teacher whose aim is to make his teaching more effective. In the foreword to the book Ungrading mentioned above, Alfie Kohn writes "... research showing that grading has three predictable effects - less interest in learning, a preference for easier tasks, and shallower learning". As an example of these undesirable effects let me mention the story of one of my colleagues who required that students can not only state, but also prove the main theorems in the oral exam. As proofs are often long, he was satisfied when a student described the first steps of it correctly. But this had the effect that students learned only the first steps of the proofs by heart!

Alternatives to grades range from descriptive labels over individual narrative judgements and feedback to letting the students propose or even decide their grade. If you are sceptical if such alternatives work in a conventional system, take a look at the book "Ungrading".

Much more could be said about teaching than can fit in a single column. As I have emphasized here the importance of active involvement and critical reflection, I end with a few questions instead of a summary: Who was the best teacher you had and what made him or her stand out? With which points in this column do you disagree? Can you imagine yourself teaching a course where students decide themselves which grade they should get?

Reference https://urldefense.com/v3/__https://english.stackexchange.com/questions/226886/origin-of-i-hear-and-i-forget-i-see-and-i-remember-i-do-and-i-understand__;!!Ls64Rlj6!iBmm42wQ096vWHptw8BN3EmanTRrgMZMlqGQL9dbRHRz4iqsrsV-p4J0qzkNWMeC$

Hans Rudolf Künsch, Ph.D.
Professor Emeritus of Mathematics
Seminar für Statistik
ETH Zürich
Switzerland

Yi' s FDA Story: When Statistics Met Regulation 1995

Yi Tsong

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

Remember the important historical events happened in 1995? Historically, it is the year of the trial of O. J. Simpson murdered his wife Nicole Brown Simp-

son and her friend Ron Goldman on June 13, 1994. It started with Jury swearing in on November 9, 1994 and acquitted on October 3, 1995. It is the "most publicized criminal trial in American history". Given all the problems in the evidences presented by the prosecuting team, a seemingly clear case became so fuzzy in the 11 months of the trial. However, the opinions of the white and black citizens are so divided. The black vs. white discussion were carried out in office across all states.

It is also the year that the Operation Desert Storm officially ended on Nov. 30. Though a ceasefire had been declared years earlier for the Gulf War, the conflict that arose when Iraq invaded Kuwait, the operation didn't officially end until 1995. It is little did we know, troubles in Mideast became a never-ending story. It led to 9-1-1 and Iraq war and ···

April 19 of 1995 is the date marked the beginning of domestic terrorism with The Oklahoma City Bombing. In an attack on a government building in Oklahoma City, 168 people were killed —including more than a dozen children at a day-care center at the site —and many more were injured. It was then the deadliest terrorist bombing in the nation' s history. One of the bombers, Timothy McVeigh, was executed about six years later. The domestic terrorism has been escalated throughout the years. It reached its new high on January 6th of 2021, with the invasion of United States Capital. Remember the day, eBay went live? It was on Sept. 3 of 1995. Then it was called AuctionWeb, the site was started by Pierre Omidyar to help his then-fiancée find Pez dispensers for her collection. It has a revenue of 27.6 billion USD in 2020. Other tech milestones of that year were the launch of IMDb and the release of Windows 95.

Other historical 1995 news in music areas are the deaths of Mexican - American singer-song writer Selena and Jerry Garcia of the Grateful of the Grateful Dead. I describe the two events below.

Selena Quintanilla-Pérez (April 16, 1971 –March 31, 1995) was an American singer, songwriter, spokesperson, model, actress, and fashion designer. Referred to as the "Queen of Tejano music", her contributions to music and fashion made her one of the most celebrated Mexican-American entertainers of the late 20th century. Billboard magazine named her the top-selling Latin artist of the 1990s decade, while her posthumous collaboration with MAC cosmetics became the best-selling celebrity collection in cosmetics history. Media outlets called her the "Tejano Madonna" for her clothing choices. She also ranks among the most influential Latin artists of all time and is credited for catapulting a music genre into the mainstream market.

Selena was shot and killed on March 31, 1995, 16 days before her 24th birthday, by Yolanda Saldívar, her friend and the former manager of her Selena Etc. boutiques. Saldívar was cornered by police when she attempted to flee, and threatened to kill herself, but was convinced to give herself up and was sentenced to life in prison with possible parole after 30 years. Two weeks later, Texas Governor (and future US President) George W. Bush declared Selena's birthday Selena Day in Texas. Her posthumous crossover album, Dreaming of You (1995), debuted atop the Billboard 200, making Selena the first Latin artist to accomplish this. In 1997, Warner Bros. released Selena, a film about her life and career, which starred Jennifer Lopez as Selena and Lupe Ontiveros as Saldívar. Selena has sold around 30 million records worldwide, making her one of the best-selling female artists in Latin music.

Selena was shot and killed on March 31, 1995. https://www.youtube.com/watch?v=XBCDvINm0Vo&list=PLsvD5ibAnWy_BrS_nD5mcWbTLfUiCHN-I "I could fall in love" sung by Selena.

The Grateful Dead played their last concert (July 9): It was a bad year for Deadheads. Earlier that summer, over 100 fans were injured when a deck at a concert site collapsed, and another concert' s cancellation led to rioting. Then, mere weeks after their final show in Chicago, lead guitarist Jerry Garcia died of heart failure. The band would announce their official break-up later that year.

"Ripple" by The Grateful Dead: https://www.youtube.com/watch?v=671AgW9xSiA

The Grateful Dead was an American rock band formed in 1965 in Palo Alto, California. The band is known for its eclectic style, which fused elements of rock, folk, country, jazz, bluegrass, blues, gospel, and psychedelic rock, for live performances of lengthy instrumental jams; and for its devoted fan base, known as "Deadheads". "Their music", writes Lenny Kaye, "touches on ground that most other groups don't even know exists." These various influences were distilled into a diverse and psychedelic whole that made the Grateful Dead "the pioneering Godfathers of the jam band world". The band was ranked 57th by Rolling Stone magazine in its The Greatest Artists of All Time issue. The band was inducted into the Rock and Roll Hall of Fame in 1994 and a recording of their May 8, 1977 performance

at Cornell University's Barton Hall was added to the National Recording Registry of the Library of Congress in 2012. The Grateful Dead has sold more than 35 million albums worldwide.

In around the spring break of 1993, I ran into an FDA chemist reviewer, Dr. Ting Eng Ong. She was looking for Karl Lin, a statistical reviewer who was responsible for statistical review of animal study data and drug product quality. Because of the spring break, Karl was taking a vacation. Ting Eng seemed to be in hurry needs of statistical help to interpret some statistical sampling acceptance plan for the drug quality. I volunteered to look into the problem. It somehow got me involved in researches in statistical methods of drug product quality for the next thirty years. I recalled that when I read USP (United States Pharmacopeia), I was totally confused about the statistical interpretation of these. I also realized that the need of more methodological statisticians involved to develop new methods or revise the methods available at the time. After I worked with Dr. Ong and a few of my statistical colleagues, I decided to write a statistical article about dissolution test and published it in Journal of Biopharmaceutical Statistics in order to generate more statistical interest in the pharmaceutical quality control area.

Dissolution testing was introduced during World War II as a military procurement tool to minimize the number of incidents of the passage of intact tablets through the full gastrointestinal system and their elimination in the feces. A dissolution test was first introduced in U.S. Pharmacopeia (USP) XV in 1955. Due to a few cases of therapeutic failures with readily disintegrated tablets recognized in the middle 1960s, and while dissolution testing appeared to be predictive for in vivo dissolution, the importance of the development of a dissolution procedure were recognized. In 1970, an official dissolution procedure was first adopted in USPXVIII (1,2).

By 1973, bioavailability differences between brands of a numbers off drugs were recognized and demonstrated through intensive screening of dissolution profiles. Hence, the Food and Drug Administration (FDA) set performance limits on dissolution in terms of that was a specific percentage of the label claimed quantity of drug required at a specific time. Based on inferences from correlated bioavailability studies, these dissolution specifications were intended to assure that at least 80% of the labeled quantity of drug was available.

When carried out appropriately, dissolution analysis of pharmaceutical drug has been identified as an important test to assure the quality of a drug product (3). To ensure the manufacturing procedure is yielding product equivalent to that which formed the basis of new drug application approval, dissolution testing is required by the FDA to confirm that the dissolution profile is consistent with the declared dosage (4). The USP dissolution test is provided as a standard to determine compliance with the dissolution specification(s) stated in the individual monograph for a tablet or capsule dosage form. The dissolution apparatus is typically constructed so that the dissolution testing may be performed on six tablets or capsules simultaneously, and so that samples may be performed on six tablets or capsules simultaneously, and so that samples may be drawn at various predetermined time intervals to generate a dissolution profile. Each lot of tablets or capsules must be the USP requirement as set forth in the compendium in order to be released as a satisfactory drug product (1,2,4).

The objective of the acceptance sampling procedure is to reject lots with large number of tablets or capsules with dissolution values less than Q, a pre-specified required percentage of the label claim dissolved at the pre-specified timepoint. On the other hand, to accept lots with small numbers of tablets or capsules dissolved less than Q. The acceptance rules proposed by the USP(5-7) have an important function in the process. Although the rules were proposed for one sample consisting of six to twenty-four tablets depending on the number of steps required for acceptance decision on the sample), typically the manufacturer will accept or reject the lot based on the rest of the tested sample and taking more testing sample if any technical problem is detected and identified. The acceptance rules are based on acceptance sampling by attributes. Modifications of USP procedures have also been proposed and used by both the manufacturers and FDA reviewers (8,9).

Many of the modification were proposed based on the capability of the manufacturers. When published as a revised chapter in USP, it often stated without the original statistical base of the procedure nor the evolvement throughout the history of USP. Furthermore, Japan and Europe published their own compendia that were not exactly the same as the USP procedure. Although statisticians of USP claimed the USP procedure is a compendium which meant that the procedure could not be used for inference beyond the sampled data. It could not be used for lot release justification due to its extremely small non-random sample size. However, many manufacturers used it exactly for lot release purpose. After we completed the project in 1994, I proposed also a new procedure based on maximum-

likelihood estimation (10). My co-authors and I documented the work into a manuscript submitted to Journal of Biopharmaceutical Statistics (JBS). The manuscript was titled "Dissolution test acceptance sampling plans" and published in 1995 (11). Please refer to that article for a clear description of the compendia used by USP and the statistical procedure proposed using the operating characteristic curves. *(Note: This is edited by Dr. Ming Wang. Due to the format issue, the content of this article is omitted here).*

References

1. Banakar UV: Pharmaceutical Dissolution Testing. Marcel Dekker, New York, 1991.
2. Wood JH: Dissolution rediscovered. In: Pharmaceutical Dissolution Testing (Banakar UV, Ed). Marcel Dekker, New York, 1991, pp. 413-421.
3. Bisaillon S, Tawashi R: Influence of fluid motion in dissolution-rate determinations. J Pharm Sci 60:1874-1877, 1971.
4. Cox DC, Douglas CC, Furman WB, Kirkchhoefer RD, Myrick JW, Wells CE: Guidelines for dissolution testing. Pharm Tech 2(4):40-53, 1978.
5. The United States Pharmacopeia XIX. Mack Printing Company, Easton, PA, 1975.
6. The United States Pharmacopeia XX. Mack Printing Company, Easton, PA, 1980.
7. The United States Pharmacopeia XXII. Mack Printing Company, Easton, PA, 1990,
8. Pheatt CB: Evaluation of U.S. Pharmacopeia sampling plans for dissolution. J Qual Tech 12(3):158-164, 1980.
9. Givand TE: An evaluation of the dissolution test acceptance sampling plan of USP XX, Pharmacopeial Forum, Mar-Apr.:186-190, 1980.
10. Grant EL, Leavenworth RS: Statistical Quality Control, 6th ed. McGraw-Hill, New York, 1988.
11. Tsong Y, Hammerstrom T, Lin K, Ong TE. Dissolution test acceptance sampling plans. J Biopharm Stat. 1995 Jul;5(2):171-83. doi: 10.1080/10543409508835106. PMID: 7581426.

Yi Tsong, Ph.D.
Division Director,
CDER/OTS/OB/DBV,
U.S. Food and Drug Administration

ICSA 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. Mengling Liu, the ICSA Executive Director (executive.director@icsa.org).

ICSA Sponsored Meetings:

ICSA 2021 Applied Statistics Symposium

September 12-15, 2021
The 2021 ICSA Applied Statistics Symposium will be held virtually on September 12-15, 2021. For detailed information, please click https://symposium2021.icsa.org/.

ICSA 2022 Applied Statistics Symposium

The 2022 ICSA Applied Statistics Symposium will be held in Gainesville, Florida. More detailed information will be shared later.

ICSA 2021 China Conference Postponed to 2022

The ICSA 2021 China Conference is postponed to 2022 and will be held at Xi'an University of Finance and Economics, Xi'an, China.

ICSA 2022 China Conference

July 1 - 4, 2022
The ICSA 2022 China Conference will be held at Xi'an University of Finance and Economics, Xi'an, China, and the date will be announced later. For information, please contact Scientific Program Committee Co-Chairs Professor Yingying Fan at fanyingy@marshall.usc.edu and Professor Chunjie Wang at wangchunjie@ccut.edu.cn.

ICSA 2023 China Conference

The ICSA 2023 China Conference will be held at Chengdu, co-sponsored by Southwest Jiaotong University (SWJTU).

12th ICSA International Conference

December 18-20, 2022
The 12th ICSA International Conference will be held at the Chinese University of Hong Kong from December 18 to December 20, 2022.

ICSA Co-sponsored Meetings:

77th Annual Deming Conference on Applied Statistics

The 77th Annual Deming Conference on Applied Statistics will be held from Monday Dec. 6 to Wednesday Dec. 8, 2021, followed by two parallel 2-day short courses on Thursday Dec. 9 and Friday Dec. 10 at the state-of-the-art Tropicana Casino and Resort, Havana Tower, Atlantic City, NJ.

The purpose of the 3-day Deming Conference on Applied Statistics is to provide a learning experience on recent developments in statistical methodologies in biopharmaceutical applications. The conference is composed of twelve three-hour tutorials on current topics in applied biopharmaceutical statistic and FDA regulations, and a one-hour distinguished keynote speaker on each of the 3 days of the conference. The books, on which these sessions are based, are available for sale at an approximately 40% discount. Attendees will receive hard copy program proceedings of the presentations.

There will be poster sessions. Early registrants who submit a poster presentation will receive a $150 discount. For poster submission, please contact "Deming Poster Chair": Dr. Pinggao Zhang at email: pinggao.zhang@takeda.com.

There will be student scholar presentations. For a student scholar application, please contact "Deming Scholar Chair": Dr. Sofia Paul at email: sofia.x.paul@gsk.com.

The conference is sponsored by the American Statistical Association Biopharmaceutical Section and the International Chinese Statistical Association. Walter Young has chaired this conference for 52 consecutive years. The program committee include: Alfred Balch, Joseph Borden, Ivan Chan, (Din) Ding-Geng Chen, Kalyan Ghosh, Satish Laroia, Sofia Paul, Manoj Patel, Naitee Ting, Bill Wang, Wenjin Wang, Yibin Wang, Li-an Xu, Walter Young and Pinggao Zhang.

Registration is expected to open by mid-August and one page program should be available by that

time. For more information about the conference, please visit https://demingconference.org/ or email Din Chen, Deming Publicity Chair, at din@demingconference.org.

IMS Asia Pacific Rim Meeting

Postponed to January, 2023
The sixth meeting of the Institute of Mathematical Statistics Asia Pacific Rim Meeting (IMS-APRM) will provide an excellent worldwide forum for scientific communications and collaborations for researchers in Asia and the Pacific Rim, and promote collaborations between researchers in this area and other parts of the world. The meeting will be held in Melbourne, Australia and please see http://ims-aprm2021.com/ for details. Firm dates will be announced at a later date.

China Biostatistics Summit

July 16-18, 2021
China Biostatistics Summit (CBS) is an annual national conference focusing on statistics in drug development and regulatory decision making. This conference is organized by the China Tripartite Coordination Committee (CTCC) which comprises statisticians from government (Center for Drug Evaluation or CDE of the National Medical Product Administration or China's FDA), academic, and industry. The purposes of the CTCC are to draft new and/or revise existing statistics guidelines for use in the industry and by the regulatory agency. Therefore, the annual conference (similar to the Regulatory-Industry Statistics Workshop or RISW in the US) aims at bringing statisticians working in drug development and regulation to discuss new technical guidelines, share experiences, and learn new methodologies. This year's CBS (the 4th CBS) will take place in Guangzhou on July 16-18, 2021 with the theme of "Practice and Consensus - A Joint Journey for Development and Regulation" (实践与共识–研发和监管共发展). As usual, many CDE officers, statisticians, and department heads will be invited to give speeches at this conference. Details can be found here https://mp.weixin.qq.com/s/bmhnTZF79cQxwYyXhBasgg.

Online Training and Seminars:

ICSA Online Training

Online training serves as a viable alternative to traditional continuing education options, e.g., to short courses offered at biostatistical conferences. Over the past year, the ASA Biopharmaceutical Section has been working on creating an online training program aimed at clinical trial statisticians and set up a pilot program, which includes half-day and full-day courses on key topics in biopharmaceutical statistics:

- Analysis of Longitudinal and Incomplete Data
- Multiplicity Issues in Clinical Trials
- Analysis of Surrogate Endpoints in Clinical Trials
- The section has received much positive feedback from industry and academic statisticians. Clinical trial statisticians who took advantage of the online training program emphasized that this program is convenient, inexpensive and quite flexible.

A similar online training program has been set up for ICSA members. As a member of the ICSA, you will receive a 50% discount when you sign up for any course included in the program. The online training courses are based on professionally recorded videos using a format similar to that used in YouTube videos. The videos can be accessed 24/7 on a computer or even on a smartphone. The cost of online training is low compared to traditional training, and it can be further reduced by using a group-training format. Up to 25 people can view an online training course with a single registration, which lowers the cost of online training to about $20-25 per person for full-day courses and $10-15 per person for half-day courses.

For more information about the online training program and to sign up for the individual online courses, please visit this web page: http://sprm.com/icsa/.

Healthcare Innovation Technology: The Pod of Asclepius

Looking to stay up to date on developments in health care technology around the world? The American Statistical Association is sponsoring "The Pod of Asclepius", a new podcast where data scientists, statisticians, engineers, and regulatory experts discuss the technical challenges in their healthcare domain.

We have over 20 episodes published and available on YouTube, Podbean, iTunes, Stitcher, Podchaser, Tune In Radio, and Google Play. Looking for a good place to start? Check out the following episode links:

ICSA Upcoming Events

- Risks and Opportunities of AI in Clinical Drug Development with David Madigan and Demissie Alemayehu
- Kidney Injury - Biomarkers for Prediction and Prognosis with Allison Meisner
- NHS Digital Health Initiatives with Emma Hughes
- Data Platforms to Monitor Animal Health with Shane Burns
- Bayesian Approaches in Medical Devices: Part 1, Part 2, Part 3 with Martin Ho and Greg Maislin

You can catch up on all episodes on our YouTube playlists for Season 0 and Season 1. The easiest way to catch new episodes is to subscribe via our channels:

- Youtube: https://www.youtube.com/channel/UCkEz2tDR5K6AjlKw-JrV57w
- Podbean: https://podofasclepius.podbean.com
- You can see our full schedule on the website: www.podofasclepius.com

For detailed information, please visit: https://www.podofasclepius.com/philosophy-of-data-science.