In this Issue

Message from the Chair 2
Faculty 3
Arrivals, Promotions, and Awards 3
New Faculty Interviews 4
  Professors Jianhua Hu, Ying Liu, and Xiaoyu Che
Promoted Faculty Interviews 7
Professors Bin Cheng and Cheng-Shiun Leu
Grant Awards 9
The New PhD Curriculum
Professor Jeff Goldsmith 12
Innovations in Genomics
Professor Iuliana Ionita-Laza 13
Data Science and Psychiatry
Professor Melanie Wall 14
Reflections on Transitions
Professor Emeritus Bruce Levin 15
The Postdoctoral Program 16
Students 17
Student Data 17
The BEST Diversity Program 18
Student Awards & Defenses 21
The Biostatistics Computing Club 22
The Graduate Student Research Seminar 22
Staff Spotlight 23
Interviews with Katy Hardy and Alfredo Bautista 23
Alumni Spotlight 25
Tai Chen, Glaxo Smith Klein 25
Morgan de Ferrante, Regeneron 26
Nick Williams, Weill Cornell 27
Julia Wrobel, University of Colorado 28
Faculty Publications 29
Follow Us 40
Acknowledgements 40
It is my privilege and honor to welcome you to the January 2020 edition of Significant Moments. After 22 years at the University of Southern California, I am looking forward to the excitement and privilege of leading such a great department filled with accomplished faculty, extraordinary staff and talented students. My excitement stems from a multitude of reasons, including the rich history of the department as a methodologic powerhouse, its deep-rooted emphasis on solving the most pressing public health issues, the ever increasing focus at the Mailman School of Public Health on multi-disciplinary research to exploit opportunities and challenges during this exciting era of big data, its attention to diversity through pioneering enrichment initiatives such as the Biostatistics Enrichment Summer Training (BEST) program, the global focus of Columbia and the dynamism due to its location in the vibrant city of New York. One of the most impressive initiatives at Columbia is its forward-looking investment in forming a university-wide Data Science Institute (DSI). From my discussions with leaders across Columbia, including the director of DSI - Dr. Jeannette Wing, it is evident that we are well positioned to seize the many opportunities that the data science revolution provides for advancing public health research and training. The Department will continue to play a leading role in advancing Health Data Science at Columbia in partnership with DSI and other key partners (e.g., Statistics, Computer Science, (Bio)informatics, etc.).

Over the next several years, we will invest strategically to build on the great foundation already laid by past investments in solidifying the department’s reputation not only in data science, but also in “traditional” areas of biostatistics that the department is well known for.

As we embark on an ambitious mission to propel the department to the next level, I would like to take this opportunity to pay tribute to several past and current leaders of the department on whose shoulders we proudly stand. This issue of Significant Moments has a nice article of reminiscence and words of institutional wisdom from Dr. Bruce Levin, a past chair of the department who has transitioned to a well-deserved retirement and emeritus position after a long illustrious career as a leader. I would like to pay tribute to Dr. DuBois Bowman, the last permanent chair that I am replacing, for his leadership and strategic investments in hiring great new faculty that have positioned us well in several key areas including data science. I would also like to thank Dr. Ken Cheung for his unreserved willingness to provide insightful advice during the last few critical months and for his remarkable leadership as an interim chair during the transition period. Finally, I would like to thank the many senior and junior faculty and departmental leaders for being so generous with their time and for sharing their insights and vision for how the department should further invest in research, education and scholarship in order to cement its position as a leading department in the nation and, indeed, the world.

This issue includes several articles and interviews that will give you a good glimpse into the many areas that the department excels in, fascinating personal and professional stories from members of the department, and how the department is planning to advance its core missions of cutting-edge scholarly research and the training of the next generation of leading scholars. Specifically, I would like to draw your attention to the interview with Dr. Jeff Goldsmith on the revamped PhD curriculum with attention to the emerging competency areas in data science and several foundational topics in biostatistics, the interview with Dr. Iuliana Ionita-Laza on her exciting new CUIIMC campus-wide program on multiomics data analysis, and the interview with Dr. Melanie Wall on her personal journey as a scientist and her vision for the Division of Mental Health Data Science that she leads. My hope is that the articles and news items in this issue will be enticing enough to get you interested to know the department better. We hope that we will have the privilege of hosting you as visitors to share your knowledge and wisdom and, hopefully, as potential collaborators, new colleagues and/or future trainees at all levels.
New Faculty
Welcome new faculty!

JIANHUA HU, PhD is joining us as Professor of Biostatistics (in Medicine and in the Herbert Irving Comprehensive Cancer Center) and Director of the Cancer Biostatistics Program.

YING LIU, PhD is joining us as Assistant Professor in the New York State Psychiatric Institute

XIAOYU (JASON) CHE, PhD is joining us as Assistant Professor of Biostatistics (in the Center for Infection and Immunity) at the Columbia University Medical Center

Faculty Promotions
Congratulations to the following faculty on their recent promotions!

BIN CHENG, PhD has been promoted to Professor of Biostatistics at the Columbia University Medical Center.

CHENG-SHIUN LEU, PhD has been promoted to Professor of Biostatistics at the Columbia University Medical Center.

Faculty Awards
Calderone Research Prize for promising, high-impact proposals from Junior Faculty, 2018
LINDA VALERI, PhD, Assistant Professor of Biostatistics, proposes to develop causal inference approaches and automated software for estimating the joint causal effect of environmental and behavioral factors over time.

Tow Faculty Leadership Scholars, 2018-2020
JEFF GOLDSMITH, PhD, Associate Professor of Biostatistics

Sigma Xi Scientific Research Honor Society Inductee, 2019
GEN LI, PhD, Assistant Professor of Biostatistics

Web of Science Group listing of the world’s most Highly Cited Researchers 2019
MELANIE WALL, PhD, Professor of Biostatistics
What made you decide to pursue a career in biostatistics?

My training background was in mathematics (college) and statistics (master’s degree). I learned from my studies and training that I enjoyed more solving practical problems, especially in biomedical science, using quantitative methods. This is why I switched to biostatistics to pursue my PhD degree. And I am still continuing on the same path and love what I do, after almost 20 years since then.

What are the main challenges and opportunities for a biostatistician like yourself who is the Director of Biostatistics in a Cancer Center?

The main challenge is to build up collaboration with cancer researchers and clinicians from the scratch. Also, developing a well-functioning biostatistics group to help cancer researchers and improve quality of research conducted at HICCC takes a lot of time and effort. But in the meanwhile, it also provides opportunities for biostatisticians to foster new collaborations, which can motivate new ideas for statistical method development. For example, collaboration with a group of brain cancer researchers involves the cutting-edge technology of single-cell RNA-Se- quencing. Depthful discussion on this topic made me realize that the existing analytical tools have a lot of room to improve. Currently I and my trainees are developing a more statistically systematic modeling framework to improve the analysis quality for scRNAseq data. Also, the studies involving brain imaging data provide very good applications for deep learning methods that we are developing.

In addition, the rich collaborative activities provide biostatisticians opportunities to educate clinicians and basic scientists in cancer research the importance of performing proper statistical practice, including designs and data analysis. Graduate students are also given great opportunities for conducting their practicum projects.

What are your first impressions since joining Columbia?

The collaboration atmosphere is welcoming and open at Columbia. New colleagues are friendly. But it did take quite self-involvement to get things set up at the beginning.

How do you like living in New York City?

Both weather and living style are very different from Houston where I came from. It took me some time to adjust. But I do enjoy much clearer four seasons and a large variety of activities in the city.
Interview with Ying Liu, PhD

What brought you to choose a career in Biostatistics?

I felt that statisticians are different, and have different strengths across the spectrum of math, computer science (an algorithmic/engineering mind set), and enthusiasm for science.

At the time I chose to do a PhD in biostatistics, ‘big data’, ‘data science’ and ‘machine learning’ were not buzzwords. I studied math as an undergraduate, but my interest in science has always been greater than my interest in math alone, and I am also very good at applied and computational statistics. So I felt that a career in biostatistics can make good use of my strengths.

What are your main research interests?

In short, I want to integrate machine learning methods in knowledge discovery to facilitate decision making and make the algorithm/methods practical enough to change clinical practice. This is not an easy path. It is much less straightforward than adopting machine learning in other sectors such as advertising and logistics. Some obstacles are limited samples and the need to make the black box models transparent. My recent projects include a deep generative model to solve selective inference problem, which enables False Discovery Rate control in variable selection for any supervised learning algorithm. I have become very interested in generative models and Variational Bayesian algorithms. I am using these to address a range of problems in projects from drug addiction to finding patient subtypes with multiple modality brain imaging data. I am also interested in adopting reinforcement learning and randomization designs in mobile health.

Why did you choose to join Columbia and the New York State Psychiatric Institute?

I got my Ph.D. in Biostatistics at Columbia in 2016. I felt this was a very nourishing environment that helped me to reach my potential as a student, and I trust it can be a good environment for me to grow further at this stage of my career. Psychiatry is seeking new computational methods and new ways of adopting machine learning to integrate ‘big data’ in diagnosis and treatment decisions. I felt that this is a field that matches well with my research interests.

What is your typical day like?

I get to the office, have a coffee. Then I might meet some collaborators, work on a grant proposal, or revise a manuscript. I might also run an analysis in R or code a package in Python, or talk with people from a wide range of departments and disciplines (such as biostatistics, statistics, computer science, or neuroscience) for fun and new ideas.
Interview with Xiaoyu (Jason) Che, PhD

Assistant Professor of Biostatistics

What brought you to choose a career in Biostatistics?

I’ve always liked math since my childhood. That’s why I picked math as my major in the Chu Kochen Honors College at Zhejiang University and chose to pursue PhD in math at Claremont Graduate University. My PhD adviser, Prof. John Angus, was an expert in reliability theory the analysis of which overlaps greatly with survival analysis. With the booming of medical data since the new millennium, I decided to develop new methodologies to analyze various types of medical data as my PhD dissertation in which I modeled the relationship between recurrent event process and time to terminal event. Therefore after my graduation, devoting myself to a career in biostatistics was an obvious choice.

What are your main research interests?

My interests of research include counting process, survival analysis, association analysis, machine learning, and their applications in the data analysis regarding public health.

Why did you choose to join Columbia

After I graduated from the PhD program, I was recruited by my current boss, Dr. W. Ian Lipkin, to the Center for Infection and Immunity (CII) as the lead biostatistician. I was impressed by the enormous amount and the high diversity of medical data that CII managed to collect from all sources. This would be the perfect job for a biostatistician. Being in Columbia University and NYSP also ensures that I would get maximal support from my colleagues in biostatistics and opportunities to collaborate with clinicians.

How is your typical day like?

Most of the time, I do research at CII designing and implementing statistical methods in research projects studying the pathogenesis of autism spectrum disorder (ASD), myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), and emerging infectious diseases. After being appointed as the assistant professor in biostatistics, I also teach MS-level probability for more than 100 students.
Can you tell us about your methodological research?

My primary research area is statistical methods in design and analysis of clinical trials, including phase I dose finding, phase II selection, and phase II/III seamless adaptive trials. Specifically, with my collaborators I propose novel dose finding methods that account for multiple toxicities, toxicities obtained from both clinicians and patients, toxicities in the presence of early disease progression, and toxicities in the presence of drug inadherence. We develop adaptive methods for dose selection and confirmation in a seamless fashion to increase the chance of success in drug development. I also work on precision medicine, clustering analysis, and mediation analysis.

Why choose to work at Columbia?

The Columbia medical center provides collaborative opportunities that often lead to methodological innovations in statistics. This enhances your sense of fulfillment because you know you are doing more than creating theorems that may never be used in practice.

What is a Biostatistician’s role in the emergence of data sciences?

First, a modern biostatistician must face the reality that not all data-related problems can have a “valid” solution in the traditional sense, that is, one that is currently theoretically justified. Second, a modern biostatistician should stand his or her ground in terms of educating people on the limitations of an unjustified solution, and the need to provide confirmation of its validity as much as possible. Exciting findings are emerging from the explosion of data, but only time can tell which of them are genuine. The modern biostatistician should have an open but critical mind as they face this data explosion.

You are actively involved in training PhD students. What is your educational philosophy?

I emphasize balanced training. A PhD student should be exposed to both theory and application in a balanced way. A student can choose a more theoretical topic or a more applied one as his or her thesis topic, but the training must address both areas.

What are your interests outside of work?

I enjoy reading books, including mathematics, history, politics, and world literatures (in English or Chinese).
Interview with Cheng-Shiun Leu, PhD

What are your main areas of methodological interest?

My methodological research focuses on adaptive designs in clinical trials, specializing in sequential ranking and selection problems. In earlier-phase research, rather than trying to declare that one treatment is better than another, we usually need to select treatment(s) from multiple candidates to bring forward for Phase III confirmatory testing. Sequential selection procedures not only require fewer patients to correctly identify the best (subset) from the candidate treatments with high probability, they also allow for sequential elimination of inferior treatments and sequential recruitment of superior treatments before the trial ends. I am delighted that after two decades of joint efforts with Professor Bruce Levin to develop a family of such procedures and establish their theoretical properties, they are being implemented in an ongoing trial and have been proposed for other planned studies.

Given your 20 years of experience, in your opinion, what is the most critical role for a Biostatistician in clinical trials; and what part of this work do you enjoy the most?

Since graduating I have been engaged in over 60 studies, many of them randomized clinical trials. In addition to conducting sample size calculations and statistical analysis, I think it is critical for experienced statisticians to get involved from the outset in the trial design, to help the investigators specify a trial protocol and design that meets the research goal(s) and will withstand subsequent challenges. I particularly enjoy addressing, and helping everyone involved to understand, statistical issues that arise while studies are ongoing, and interpreting unexpected findings. This is very fulfilling.

You work with many Masters students. What is your approach?

The primary goal of my teaching is to introduce students to an in-depth understanding of the terminologies and concepts we frequently use in our daily practice; the principles of data analysis, and methods to achieve scientific rigor in clinical research. I strongly recommend students to be aware of the importance of thorough thinking even before they start any analysis. For example, analyzing a dataset that is dirty without noticing this will eventually lead to disaster. I also like to bring students’ attention to the post-analysis thought process. While it is important to see that the results are statistically significant, cautiously interpreting the results and understanding what can and can’t be said from them is even more critical.

When did you come to Columbia, and why do you choose to stay?

New York City was the first foreign city I visited. In 1992 I came to Columbia University, where I was extremely fortunate to be able to work with Professors Wei-Yann Tsai and Bruce Levin, and completed my Ph.D. in 1997. After graduating, the privilege of collaborating with these two world class statisticians and world leading clinical experts at the HIV Center on a daily basis for twenty years is really a dream come true. These experiences have helped me grow immensely in many ways. There is absolutely no doubt that working at Columbia has been highly productive and exciting!

Aside from work, I like to cook and eat, watch sports, and travel around the world. New York City offers options for food from all countries, opportunities to experience professional sports, and convenience for traveling all over the world. Although I love exploring new places, it remains the place I enjoy the most.
NEW GRANTS

**Shuang Wang**

R01LM013061 funded by the National Library of Medicine (Role: PI)  
(Award: 02/12/19-01/31/24)  
“Big Data Methods for Comprehensive Similarity based Risk Prediction”

The project focuses on developing a novel data science pipeline which includes a clinical data processing pipeline to format comprehensive patient health determinants from a variety of sources of clinical, genomic, socioenvironmental data, and a clinical-outcome-prediction framework that optimally fuses relevant patient health determinants to define patient similarity for improved clinical risk predictions.

**Qixuan Chen**

R21 ES029668 funded by the National Institute of Environmental Health Sciences (Role: PI)  
(Award: 07/15/19-06/30/21)  
“Bayesian exposure-response analysis for immunoassays data with measurement”

The proposed research will introduce new Bayesian approaches for estimating nonlinear exposure-response relationship between a continuous environmental exposure and a binary disease outcome and for assessing the combined health effect of environmental exposure mixtures, in which the exposures are measured with errors but external calibration data are available to correct the errors. The proposed methods will be applied to the New York City Neighborhood Asthma and Allergy Study to assess the effects of indoor allergens on asthma morbidity among asthmatic children, with the indoor allergen concentrations measured using immunoassays. The findings of this study will provide important insights for intervention and prevention of asthma morbidity among inner-city children with asthma.

**Yifei Sun**

U19 AG033655 (subcontract to CU) funded by the National Institute of Aging (Role: Subcontract PI)  
(Award: 09/01/19-05/31/24)  
“Novel Measurement Approaches to Preclinical AD”

Dr. Sun will develop and implement the tree-based statistical learning methods in Aim 1. Dr. Sun will discuss with Dr. Wang regularly via conference call on the methodology development, and will work on simulation studies, software packages, and data analyses. Dr. Sun will also participate in the preparation of manuscripts based on the results of the proposed research.

**R. Todd Ogden**

R21 MH120534 (subcontract to CU) funded by the National Institute of Mental Health (Role: Subcontract PI)  
(Award: 08/16/19-06/30/2021)  
“Nonparametric depth-based methods for analyzing high-dimensional data. Applications to biomedical research”

This research proposes to develop robust nonparametric methods for the analysis of complex high-dimensional and functional data that are increasingly being used in many biomedical fields. Aim 1 focuses on developing methods to detect outliers in such data and Aim 2 focuses on developing a global envelope test that can compare different groups of multidimensional and functional data to provide global p-values.

**Shing Lee**

Funded by SWOG/Hope Foundation (Role: PI) (Award: 01/01/20-12/31/21)  
“Analysis and visualization of adverse events and patient reported outcomes that reflect overall treatment toxicity Burden”

Analyses of adverse events data for cancer clinical trials have focused solely on the clinician’s assessment of the presence of one or more severe or life-threatening events. We propose to develop new methods for analysis and visualization of adverse events that better reflect overall treatment burden and patient tolerability.
CONTINUING GRANTS

Melissa Begg
- NIH/NHLBI (R25 HL096260), “BEST-DP: Biostatistics & Epidemiology Summer Training Diversity Program,” 2009-2024 (Contact PI; other Multi-PIs)

Ying Kuen (Ken) Cheung

Min Qian

Yuanjia Wang
- NIH/NIDA (R01 DA035846), "Impulsivity in Cocaine Abuser: Relationship to Drug Taking and Treatment Outcome," 2014-2019 (Role: Multi-PI)
- NIH/NIMH (R21MH117458) “Integrative Learning to Combine Evidence for Personalized Treatment Strategies” 2018-2020 (Role: PI)
- NIH/NIDCR R03 DE027773 “Multivariate analysis of microbial absolute abundance in population-based studies” 2018-2020 (Role: PI)

Iuliana Ionita-Laza
- NIH/NIMH (R01 MH106910), “Integrative methods for the identification of causal variants in mental disorder,” 2016-2019 (Role: Lead PI)
- NIH/NIMH (R01MH095797), "Novel Statistical methods for DNA Sequencing Data, and applications to Autism," 2012-2018 (Role: PI)
- NIH/NIMH (R01MH095797), “Novel Statistical methods for DNA Sequencing Data, and applications to Autism” 2018-2022 (Role: PI)

Gen Li
- NIH/NIDCR R03 DE027773 “Multivariate analysis of microbial absolute abundance in population-based studies” 2018-2020 (Role: PI)

Ian McKeague

Qixuan Chen
- NIH/NIDA (R01 DA034634), “Impact of health reform on outpatient substance abuse treatment programs,” 2013-2018 (Role: Sub-contract PI)
- NIH/NIEHS (P30 ES009089), “Bayesian Multilevel Models for Poststratification and small Area Estimation,” 2016-2018 (Role: Career Award PI)
R. Todd Ogden


- NIH/NIBIB (R01 EB024526), “Advance Modeling Techniques for Brain Imaging Data with PET” 2017-2021 (Role: PI)

- NIH/NIMH R01MH099003 (subcontract to CU) “Biotyping placebo and treatment-specific responses for precision medicine” 2018-2022 (Role: Subcontract PI)

Martina Pavlicova

- NIH/NIDA (R01 DA035707), “Evaluating ART for All HIV Seropositives: Can it work with the hardest cases?” 2013-2018 (Role: Sub-contract PI)


- NIH/NIMH P30 MH43520 (subcontract to CU) “HIV Center for Clinical and Behavioral Studies” 2018-2023 (Role: Subcontract PI)

John L. P. Thompson

- FDA (R01FD005407), “Phase 3 Trial of DCA in PDC Deficiency IND 028,625,” 2016-2021 (Role: Multi-PI)

Linda Valeri

What skills would you like every doctoral-level biostatistician to have? How is that reflected in the PhD curriculum changes?

Success in biostatistics requires a broad collection of strengths. These of course begin with firm grounding in statistical inference and traditional methods, but also include fluency in more recent tools in statistical learning and data science. Critically, doctoral-level biostatisticians are expected to be excellent collaborators and leaders in interdisciplinary teams, which requires an understanding of the substantive domain (or the willingness and ability to gain such an understanding) as well as strong communication skills. The new curriculum seeks to build these strengths through new or revamped courses in probability, inference, statistical methods, and interdisciplinary collaboration. This makes training in each of these domains explicit and formal.

What existing strengths does the new curriculum build on?

As with many top biostatistics departments, we have maintained a strong but fairly traditional curriculum, firmly rooted in statistical inference and probability theory. Our students have always been able to enter the workforce and lead very productive careers, and changes to the curriculum sought to build on the strength of our existing training program, particularly in the areas of statistical theory and inference. To this, we have added coursework in response to the challenges posed by a world characterized by access to massive and complex data, advances in computing resources and tools, and challenging public health and medical problems that call for in-depth interdisciplinary research.

What are the major changes to the curriculum?

There are several new or restructured classes in the PhD curriculum:
— Probability for Biostatisticians, which provides training in the elements of doctoral-level probability that are most relevant to biostatistics
— Advanced Methods, which covers topics in traditional statistical methods at the doctoral level
— A two-semester sequence in Statistical Practices and Research for Interdisciplinary Sciences (SPRIS I & II), which focuses on collaboration and quantitative leadership in interdisciplinary teams.

How does the new curriculum address the emergence of data science?

The department’s response to the emergence of data science, particularly in how we train the next generation of doctoral-level biostatisticians, is critical. To that end, incoming doctoral students will take Data Science I and II for an introduction to programming and statistical learning, followed by Advanced Statistical Computing and Topics in Statistical Learning and Data Mining for doctoral-level training in these important areas. Our students are also encouraged to pursue coursework and other opportunities at the Data Science Institute.

What change to the new curriculum are you most excited about?

I’m excited about all of it! Seriously — I’m thrilled that we can maintain the areas of strength that have long distinguished our department, while broadening our training overall and introducing some truly innovative elements like SPRIS. I think our students will be among the best-prepared as a result of the new curriculum and our faculty’s engagement.
What is the goal of the genomics program?

The goal of the genomics program is to bring together an interdisciplinary group of people from multiple departments across Columbia University with diverse research expertise in statistical/computational genomics and other omics, computational biology, biomedical informatics, and interested in understanding biology and human health. Through various activities including regular seminars and working groups we aim to increase research and learning opportunities, and benefit students and postdocs through interdisciplinary training in quantitative genomics. More information can be found on the program’s website: http://www.columbia.edu/~ii2135/genomics/.

What are the activities planned in the near future?

In the coming months we have planned regular seminars with speakers from different departments at Columbia, as well as academic/research institutes in the New York area. We are also planning a short course in June 2020, meant to provide an overview of statistical methods (along with hands-on tutorials) for the analysis of genomic and other omic data. There will likely be other educational activities in conjunction with the CTSA. The Irving Institute for Clinical and Translational Research, funded by a National Institutes of Health Clinical and Translational Science Award (CTSA).

What are your broader perspectives and thoughts on the importance of genomics in public health?

I think genomics plays a fundamental role in public health. Advances in understanding the genetic basis of complex traits and diseases can lead to substantial improvements in public health and disease prevention. We already know that drug targets that are informed by genomics have a higher chance to be successful. Also, methods such as mendelian randomization that use genetic variants as instrumental variables can help predict the efficacy of potential therapeutic targets in drug development, leading to potentially large cost savings for the pharmaceutical industry. Genomics also offers unique opportunities to identify individuals at high risk early on in their life, and these individuals can benefit from more careful screening, and adjustments of lifestyle factors that can lead to disease.
Interview with Melanie Wall, PhD

I am a first-generation college student from St. Louis. My dad was an apartment building maintenance man; my mom stayed at home to care for my two siblings and me while working part-time. Attending public school in a blue-collar suburb during the 1980s desegregation mixing black city children into predominantly white suburban schools impressed on me the values of diversity in education and society. With help from a high school guidance counselor in scholarship applications, and part time jobs (fast food restaurants, a paper factory), I paid my way through Truman State University, Missouri to a BS in mathematics, minoring in physics. After one semester in the math graduate program at Iowa State University I transferred to the statistics department. My PhD dissertation (“On nonlinear structural equation modeling”, 1998) was with Yasuo Amemiya. I became assistant professor of biostatistics, then tenured, in the school of public health at the University of Minnesota, where the fast-paced world of NIH grants and collaborations with epidemiologists, health policy researchers, and medical doctors was very stimulating. I collaborated on projects examining predictors of obesity and substance abuse using surveys and administrative databases, was a frequent statistical reviewer for NIH, and started a new graduate biostatistics course on latent variable and structural equation modeling for health sciences.

When did you come to Columbia?

In 2010 I joined Columbia University (Psychiatry) and the New York State Psychiatric institute, with its own large research portfolio and dedicated biostatistics group. This fit very well with my expertise in latent variable models and psychometrics, and enthusiasm for applied collaborations. Since becoming director of this group (2012) I have grown it to include 14 full-time biostatisticians.

What are the some of the research areas in psychiatry you collaborate on?

We participate in extensive statistical collaborations with over 50 psychiatry researchers in any given year, providing statistical expertise and analysis to projects, including: biomarkers from brain imaging and neurocognitive tasks for mood and anxiety disorders and psychosis; clinical trials of new pharmacotherapies and psychotherapies for psychiatric and neurological disorders, including building treatment decision rules; implementation studies of support service programs for mental health treatment and prevention; measurement studies for improving psychometrics of diagnosis instruments in substance use disorders, depression, and biological aging; cohort studies of child-adolescent development of psychiatric and substance use disorders; causal analysis of prescribing practices monitored from medical claims records; momentary assessment studies of cardiovascular response to emotions and stress markers related to suicide; and much more.

Why change the name of the division to Mental Health Data Science?

In 2018 I “rebranded” the psychiatry biostatistics group as the division of Mental Health Data Science (columbiapsychiatry.org/mental-health-data-science). This exemplified our mission to embrace new data collection and processing technologies (e.g. multimodal neuroimaging, wearable devices, electronic health systems) and our expertise in developing analytic tools for making sense of these complex data domains.

What do you enjoy doing in your time outside work?

I have a 10 year old kid and as much as possible we try to explore the city, go to plays, events, and be outside in the parks including taking trips to the ocean. Having spent all of my life before 2010 in the Midwest, I take every opportunity to go to the ocean, now just a car ride away.
REFLECTIONS ON TRANSITIONS

Transitions occur all the time and all around us, from the microcosm of quantum particles to the evolution of the cosmos. When transitions occur on the human level, though, they can be challenging to deal with. Here in our own Department of Biostatistics we are poised for a new transition of leadership as we welcome incoming Chair Kiros Berhane. So when Professors Thompson and Valeri asked if I would share some reflections on my own transitions, I was happy to agree. Transitions? I’ve had a few of them, going from being a professor to becoming chair, then back again from chair to happy professor (if also a recovering administrator), and then into retirement as professor emeritus. Before I begin, though, I want to thank Seamus for reminding us that “retirement” is something of a misnomer, because one doesn’t stop breathing or thinking after the paychecks stop, nor do I feel “drawn back” from doing anything interesting. I can do things now that I want to and have time enough to do them, which is a transition I can live with!

The Department was in a low spot in 1997 when I became Acting Head of Biostatistics. We had been described by one reviewer as “a sleepy little department” and morale was down because we faculty were expected to raise 100% of our salaries on funded research projects and teach courses for no compensation, that being the long-standing model of the Medical Center. When Dean Allan Rosenfield asked me to head the Department in April of 2000, I was eager to try to turn things around a bit, but what completely surprised me was how easy it was—all I had to do was respect faculty by listening to their needs, being open and transparent about how we would all move forward, and, to be sure, paying faculty to teach while they carried out their other research and academic responsibilities. Almost overnight faculty meetings turned from nasty, political bickering to stimulating discussions of biostatistics and science. This sea change allowed us to triple the size of the Department from about 10 to over 30 faculty members and achieve recognition by one scholarly assessment as the fourth most productive department of biostatistics in the nation. I am forever grateful to Dean Rosenfield for allowing me to break the mold and help our extraordinary teachers and researchers do what they do so well.

I stepped down as Chair in 2011 after 11 years (“Levin’s Eleven”) to make way for an energetic new cohort of leaders who would bravely bring us into the age of statistical data science. This was a challenging transition because the golden days of readily available NIH funding were over. Though I was no longer Chair, DuBois Bowman asked me to continue my leadership of the Departmental Committee on Appointments and Promotions, which I found buoyed me up because it allowed me to mentor faculty (many of whom I had hired), to watch their growth as scholars, and help guide them to successful promotions. This showed me how to keep the rest of my career rewarding—by prioritizing the important things in life, namely, family, friends, collegiality, and a mathematically stimulating intellectual life.

Now as I write these words in beautiful Bear River, Nova Scotia, swathed in fiery Fall colors, I feel my transition to emeritus status couldn’t be more auspicious. I continue to write on stimulating topics. I’m a pretty lucky guy, but I haven’t forgotten the Department of Biostatistics and the transitions that lie ahead. My fervent hope is that every student and faculty member will continue to treat one another with dignity and respect, to listen earnestly to each another, to be generous with one’s time and consideration, and to work hard to make all of our lives better. I know in that way, the Department of Biostatistics will thrive in the days and years to come.
Shiyang Ma, PhD

I obtained my PhD in Statistics from the Department of Biostatistics and Computational Biology at University of Rochester, supervised by Professor Michael P. McDermott and Professor David Oakes. My doctoral research was in the area of randomized clinical trials, especially of dose-response studies, adaptive designs, repeated measurements and optimal experimental designs. Currently I am a postdoctoral research scientist in the Department of Biostatistics at Columbia University, under the supervision of Professor Iuliana Ionita-Laza. I am currently working on integrating Hi-C (High-throughput Chromosome Conformation Capture) data with association tests for whole-genome sequencing data analysis to improve the power of these tests by incorporating the 3D human genome structure.

Zikun Yang, PhD

My research interests are mainly in the area of model selection within the objective Bayesian paradigm. For the past several years, I have been primarily working on developing innovative methods to perform model selection for regression models, such as proposing new non-informative prior distribution and proving the corresponding consistent theorems. Currently I am working with Professor Ionita-Laza on developing statistical models to understand the functional effects of variants in the non-coding genome based on the data from ENCODE and Roadmap Epigenomics projects. There are certainly several challenges to building meaningful models, and we are working on addressing them in order to produce not only accurate predictions but also interpretable statistical inferences.

Shanghong Xie, PhD

I obtained my PhD from the Department of Biostatistics at Columbia under the mentorship of Dr. Yuanjia Wang. My doctoral research focused on machine learning, network analysis, graphical models, mediation analysis, causal inference and variable selection, with applications to neurological and psychiatric disorders. I decided to stay here for postdoctoral training because of my amazing supervisors, Drs. Todd Ogden and Yuanjia Wang, and the great environment in our department. I am currently working on network analysis and precision medicine. Although the environment is as before, I am seeing the difference between being a postdoc and being a PhD student. Postdoctoral training provides a transition period in which I am becoming more independent and mature in my research. It provides more opportunities to develop my own ideas, and also to be the lead statistician with clinical collaborators. Another awesome thing is that I have more opportunities to meet people from different departments and universities to discuss research and new collaborations. I am now involved in more simultaneous projects than before. This is helping me to be more efficient in my research and in my time management. I think that these opportunities are invaluable for preparing me for a position as a junior faculty member in the future.
Department Data: Students

Number of 2019 graduates (including February 2019 and October 2018)
DrPH: 3
PhD: 5
MS: 86
MPH: 11

Number of Returning for Fall 2019
DrPh: 9
PhD: 20
MPH: 16
MS: 119
MPH Applied Certificate: 35

Number of Incoming for Fall 2019
DrPH: 1
PhD: 5
MPH: 21
MS: 108

Where our 2019 graduates have found jobs:
• Aetion, Biostatistician/Scientist
• Boehringer Ingelheim, Statistical Programmer
• Boston Consulting Group, Analyst
• Charles River Associates, Data Analyst/Associate
• CBPartners, Analyst
• Columbia University, Data Analyst
• CDC, ORISE-Epidemiology Fellow
• DIA Associates, Analyst
• Eli Lilly and Company, Senior Statistician
• Genesis Research, Analyst
• Icahn School of Medicine, Biostatistician
• IQVIA, Associate Consultant
• KMK Consulting, Junior Analyst
• Memorial Sloan Kettering Cancer Center, Research Biostatistician
• Michael Allen Company, Research Associate
• Milliman, Healthcare Data Analyst
• Mount Sinai Health System, Biostatistician I
• NYU School of Medicine, Data Analyst
• Precision Xtract, Associate Research Scientist
• Press Ganey, Data Scientist
• Regeneron Pharmaceuticals, Biostatistician
• Research Foundation for Mental Hygiene, Biostatistician
• Sphere Institute, Statistical Programmer
• Syneos Health, Biostatistician
• Syntactx, Clinical Data Specialist
• University of Chicago, Statistician
• University of Colorado, Assistant Professor of Biostatistics and Informatics
• Verified Clinical Trials, Biostatistician
• Weill Cornell Medicine, Biostatistician
• Yale Center for Analytics, Biostatistician

Where our 2019 graduates are pursuing a doctoral degree:
• Columbia University, Biostatistics
• Purdue University, Biostatistics
• University of Southern California, Biostatistics
• University of California, San Diego, Biostatistics
• University of Pittsburgh, Biostatistics
Biostatistics Epidemiology Summer Training (BEST) Diversity Program

Each summer, a highly selective group of undergraduates from across the country attend classes in introductory biostatistics and statistical computing, and are engaged in research under the supervision of a faculty member.

**Summer 2019**

**BEST: 14 students**

**Schools they came from:**
- Montclair State Univ
- Barnard College
- Macaulay Honors Hunter College
- Los Angeles Pierce College
- Georgia State University
- Columbia University
- Wayne State University
- Albany State University
- Winston-Salem State University
- Rutgers University
- Macalaster College
- University of North Carolina-Chapel Hill
- University of Florida

**Research projects:**
- Investigating How Environmental Chemicals Lead to Adverse Liver Functions in the United States
- Analyzing the Effectiveness of New and Old Antipsychotics
- Traumatic Life Events and Viral Suppression Outcomes for Children on Antiretroviral Therapy (ART) in South Africa
- The Impact of a Fatigue Relief Intervention on Pain and Sleep Disturbances in HIV+ Latinx Individuals
- Does Moderate-To-Vigorous Intensity Physical Activity Eliminate The Detrimental Association Of Sedentary Behavior With Mortality?
- Testing The Cognitive Reserve Hypothesis in Large Data Sets
- The Effect of Ridesharing on Injury Crashes in New York City
BEST Alumni Interview

Margaret Gacheru and Melanie Mayer are BEST alumni and both are in their second year of the PhD program. Both students entered the PhD program straight from undergrad.

Prior to BEST, did you know that you wanted to be in a Biostatistics PhD program? If not, what did you think you wanted to do after graduation and how did your experience in BEST influence your trajectory?

Margaret Gacheru: In undergrad, I majored in Computational and Applied Mathematics and right up to the end of junior year, I had no idea what I wanted to do with my degree after graduation. I had explored various career paths (actuary, investment banking, consulting, etc.) but nothing piqued my interest. I knew I wanted to go to graduate school and someday get a PhD but I wasn’t sure in what field (certainly not mathematics). Luckily, the BEST program gave me the perfect answer! It was my first time being introduced to the world of biostatistics and I simply fell in love with it. Through the BEST program, I was finally able to articulate what I wanted to do after graduation. The decision to go straight from undergrad to a doctoral program was not made as quickly. After months of soul searching and conversations with mentors (especially people I had encountered during the BEST program) trying to decide whether I should pursue the master’s or PhD, I was finally able to make the commitment!

Melanie Mayer: Prior to BEST, I was not sure what career path I wanted to take. I knew I wanted to work with developing economies but what that looked like remained unclear. BEST introduced me to both the world of research and biostatistics. I really enjoyed the research portion of the program, motivating me to want to pursue a PhD. The program also showed me how I could use my interest in statistics on an interdisciplinary application such as public health, specifically the subarea of global health, where one works in large groups of experts in diverse fields to find a much more holistic and developed conclusion.

What have been some challenges that you’ve faced going straight from undergrad to a doctoral program?

MG: Last year, the main challenge I faced going straight from undergrad to a doctoral program was learning how to be a graduate student for the first time while adjusting to the level of rigor in my classes. As opposed to undergrad, there is an increased level of expectations and workload. Things that work in undergrad do not necessarily work in graduate school -- I had to change how I studied, how early I started working on assignments/projects, etc. In the first semester, through trial and error, I had to discover what worked in this new environment and how I could be most successful as a student.

MM: Going straight from undergrad to a PhD program can be difficult in terms of adjusting from undergrad to grad school. You expect it to be easy since you never stopped being a student, but I found there’s an unexpected learning curve. Especially coming from a large university where I was one of hundreds of students in my department to a much smaller program where you’re receiving a lot more attention. Although this is a good thing, it took me a while to start taking advantage of these benefits simply because I never had the option before. I slowly realized the opportunities available to me and my experience improved greatly once I took advantage of them. The expectation of you is very different once in a doctoral program. Although this can be intimidating, it also comes with more reinforcement and once I harnessed this, I began to feel a lot more comfortable.

What have been some highlights of your doctoral journey so far?

MG: This past summer, I had the opportunity to be a teaching assistant for the BEST program. It was such a unique experience to give back and contribute to
a program that altered the course of my academic/professional path. I had such a positive experience in the BEST program, so it was only natural to become a mentor and share what I have learned with other students. Talking to undergraduate students, who are from underrepresented backgrounds, about their options and watching some of them have that light bulb moment of “I can actually do this!” was truly rewarding.

MM: I have really enjoyed meeting and learning from many reputable biostatisticians/statisticians this past year. It is really great how willing professors you look up to are willing to meet with early doctoral students and encourage them. Listening to talks has also been a crucial part of my journey thus far. I remember attending talks early last year and not understanding most of what was going on. The difference from then to now is immense; my statistical knowledge has increased a lot in just one year and I feel I can bring more to the table. I’m really excited for my expertise to continue to grow.

What has been the best piece of advice that you’ve received from either our faculty member or classmate while in the doctoral program?

MG: The best piece of advice I have received is to accept that failure is a part of the process (from research mentor, Dr. Christine Mauro). Getting a PhD in any subject is a challenging process and it is impossible to know everything in your respective field. So, there are days when nothing goes according to plan, research or classes make no sense, and it might feel as if everything is falling apart. However, the bad days/failures are not an indication of what type of student or researcher you are. All you can do is try your best and move on.

MM: I have been told repeatedly “it’s not a sprint, it’s a marathon” and to be reminded of this from time to time, has been very valuable. It’s easy to get wrapped up in the moment and spread yourself really thin. But at the end of the day, you wear yourself out and end up overwhelmed. This is not sustainable in the long run, especially considering the length of a doctoral program. Being reminded to take a step back and breathe has actually made me more productive and impactful with my time.

What advice would you give a BEST alumnus who might be considering applying to a graduate program in biostatistics?

MG: One piece of advice I would give is to utilize the connections formed during the BEST program. As I was applying to graduate programs and making important decisions, I reached out to various people who I had interacted with during the summer to hear their perspective and receive guidance. I talked to previous BEST students who were currently in graduate school to hear about their experience and others who were already done with graduate school to learn about career options as a biostatistician. I talked to professors in the department to get advice on whether the master’s or PhD program was best for me. There is a large network of BEST alumni and people involved with the program who can serve as your advocates and guide you through your decision making process.

MM: If any BEST alumnus is considering biostatistics, I would tell him/her it’s a great decision! It is an amazing field with vast opportunities so it is really easy to cater to your specific interests. That being said, I would also advise them to explore different areas within biostatistics. There are so many different routes one can take, from pharmaceuticals to tech companies. Even if one comes in with a clear idea of their ultimate goal, you may be surprised by other options.
Student Awards & Defenses

Sanford Bolton-John Fertig Award for a distinguished doctoral dissertation in Biostatistics

HYUNG PARK, PHD for “Flexible Regression Models for Estimating Interactions between a Treatment and Scalar/Functional Predictors”

Joseph L. Fleiss Memorial Award

XIN QIU, PHD for “Statistical Learning Methods for Personalized Medicine”

Chair’s Award for Outstanding Master’s Students in recognition of outstanding academic and research achievements.

DITIAN LI, MS for “Analysis of risk factors of metabolic syndrome using a structural equation model: a cohort study”

KWANGMIN KIM, MS for “A Comparative Study of Dimension Reduction Techniques: The use of PCA, PLS, sPLS, and RCCA on Metabolomic Data from APOE4 carriers in Long Living Families”

MORGAN DE FERRANTE, MS for “Bayesian Adaptive Randomization in Clinical Trials”

Doctoral Defenses (May 2019-Dec 2019)

REBECCA ANTHOPOLOS, DRPH, “Bayesian Modeling of Latent Heterogeneity in Complex Survey Data and Electronic Health Records”


XINYU HU, PHD, “Personalized Policy Learning with Longitudinal mHealth Data”

WODAN LING, PHD, “Quantile regression for zero-inflated outcomes”

JULIA WROBEL, PHD, “Functional data analytics for wearable device and neuroscience data”


SHANGHONG XIE, PHD, “Statistical Methods for Constructing Heterogeneous Biomarker Networks”
Biostatistics Cohort

The Biostatistics Computing Club

The Columbia Biostatistics Computing Club hosts monthly seminars focusing on computational topics often encountered in biostatistics research. These tutorials provide hands-on learning at a variety of skill levels, with some focused on bringing new students up to speed and others addressing more advanced problems. In the past, we have covered a wide range of topics ranging from introductory Python, Illustrator, and LaTeX. Our overarching goal is to teach students potentially useful skills in their research or for work in the industry. The Computing Club was founded by Julia Wrobel and Ariel Chernosfsky with the help of Dr. Christine Mauro in 2017. This year’s organizers are Angel Garcia de la Garza, Margaret Gacheru and Christian Pascual.

The Graduate Student Research Seminar

The graduate student research seminar (GSRS) is a bi-weekly meeting for doctoral students to practice presenting their research in front of their peers. From conference talks to oral exams to defenses and job talks, 1-5 students share their work each meeting in an informal environment, gaining constructive feedback to sharpen their presentation skills and improve the quality of their talks. As an added benefit, the all-student audience gains helpful insights into the types of research problems being solved by their peers, and with which faculty members their peers are collaborating. This exposure is particularly beneficial for those who are exploring possible avenues for their own dissertation research.

Throughout the year, we hold some special sessions with the objective of highlighting opportunities available to doctoral students. These include: 1) an annual session of speed talks from faculty members to showcase projects that have potential for student engagement, and 2) an annual session of speed talks on recently completed summer internship experiences. Future special sessions will also include speed talks from the department’s post-docs.

GSRS has become a powerful tool for success among the doctoral student body, and a strength of Columbia’s biostatistics department. We look forward to continuing to provide a supportive environment for students to become effective communicators of research.
Katy Hardy is the department’s hard-working, award-winning Administrator. She joined us from the Department of Pediatrics in 2011. Since her arrival, Katy has received not one, but two Staff Awards for Excellence from the Mailman School of Public Health; first in 2014, and again in 2017. As if that weren’t enough, she also received last year’s Baton Award from the leadership of CUIMC in recognition of the outstanding collaboration and teamwork with which she fulfills her responsibilities. She is a remarkable multi-tasker, overseeing all manner of department affairs, including financial matters, grants and contracts, human resources, faculty affairs, and facilities operations. While Katy’s dependability and organization keep the department a functional workplace, it’s her kindness and generosity that help to make it the warm, friendly environment it is.

When she’s not busy juggling responsibilities at work, you can usually find her chauffeuring around teens and tweens. But when she gets some time all to herself, she likes to read, relax, go to street fairs, and work on home projects.

If you could learn to do anything, what would it be?

I would like to learn how to blow glass… or how to chop vegetables properly.

What is your favorite outdoor activity? Indoor activity?

Outdoor: taking a walk or hiking
Indoor: watching Marvel movies with my kids

If you could meet anyone, living or dead, who would you meet?

Johnny Cash or Ernest Hemingway

Have you ever had something happen to you that you thought was bad but it turned out to be for the best?

Of course, tons of times. I have found that to be the case more often than not. And if it isn’t really for the best, I just assume that that’s the way things were supposed to go and I move on to damage control.

Ten years ago, where did you think you would be now?

I thought that I would still be working at CU. I really enjoy working here and am grateful, especially now, to be working with a really great group of faculty, researchers and staff.

What would you do (for a career) if you weren’t working in the Biostatistics Department?

I would love to run a café, but mainly so that I could hang out in the café all day. Otherwise, I would like to do something to help others

What is your favorite thing about working in the Biostatistics Department?

From my point of view, it seems that most of the faculty and staff really enjoy the work that they do and enjoy working here. I think that is my favorite thing.
Alfredo joined the Department in early 2019 as a software developer in the Statistical Analysis Center (SAC). He spent the previous decade working for a financial firm on Wall Street, developing technology solutions to business problems. Although the overall goals of business and clinical research differ, they share key challenges. High quality data collection is the cornerstone of all successful information management systems, whether for clinical research or for market-oriented businesses, and we cannot manage what we cannot effectively measure.

Alfredo brought with him extensive expertise in software development, and specialization in process automation and MIS (Management Information Systems). He is experienced in enterprise-wide web applications and services, back-end relational databases, and the development of big data stores which allow the collection and analysis of data in real time.

At Columbia, Alfredo has already developed several key systems. One is an administrative application for REDCap which has streamlined project and user management and increased the efficiency of the collection of billing information within the system. Another helps gather survey information for studies of Alzheimer’s disease. The application both presents the text of questions on the screen and reads them for the participant in the language of their choice (English or Spanish).

Alfredo is looking forward to many new opportunities as he seeks to leverage his experience in software development to help advance the various scientific studies currently under way at Columbia University.
Can you tell us about the various biostatistics roles that you have had since you graduated from the department?

My first position after I graduated from the department was Associate Biostatistician, a supporting role at the study level at Bristol-Myers Squibb. Within four years, I was named the lead statistician of an asset developed to treat leukemia. Subsequently, I was appointed the lead statistician of an immunotherapy, ipilimumab. The compound later was approved and became the first check-point immunotherapy in the world. Dr. James Allison was later awarded the Nobel Price in 2018 for his contribution to ipilimumab. The statistical methodology led by our team during the development of ipilimumab revolutionizes the immuno-oncology research and has since become the standard practice around the world. In 2014, I was appointed by the senior management to establish a team of biostatisticians and programmers in field of medical and market access, where we were involved in quantitative aspect of the pricing reimbursement and negotiation. Prior to joining GSK, I was named Vice President, Head of Biostatistics, Clinical Pharmacology & Pharmacometrics of the non-oncology unit at BMS, overseeing multiple functions and therapeutic areas, including immunoscience, fibrosis and cardiovascular. Since October of 2019, I joined GSK as Vice President, Head of Oncology Biostatistics, where I oversee a team of statisticians globally.

You have been an instructor in the department and even developed the Pharmaceutical Statistics course. Why was it important for you to remain connected with the department?

Years before I graduated from the department, I had told my parents that I would like to teach at the department one day. While I was given a fantastic toolbox by the department, I had to learn how to use the tools on my own in the context of drug development. I wanted the graduates to have a head start and be prepared when they step into the world. I had a clear vision of what I needed to do, which led to the creation of Pharmaceutical Statistics curriculum.

What advice was given to you as a student that still resonates with you now?

Garbage in, garbage out. Always understand your data before you do anything to it.

In your current role at GSK, you oversee an entire department of biostatisticians. What qualities are you looking for when hiring a biostatistician.

I am looking for a rare breed of biostatisticians who have the mindset of drug developers and who can translate their quantitative knowledge qualitatively in an environment full of non-statisticians.

What advice would you give to current students/future biostatisticians?

I always say that no one should be able to identify a truly good statistician in a room full of drug developers. A well-rounded biostatistician can make contribution across different functions, such as clinical development, regulatory discussions, operational or commercial discussions. Never view yourselves as biostatisticians. This can only limit your potential.
I thoroughly enjoyed my experience at Mailman School of Public Health. One aspect that surprised me about being at Mailman was the sense of collaboration with other students in the program. During my undergraduate experience, I didn’t encounter the same willingness from other students to work together on projects that I did at Mailman. It was really nice to be able to work on practice problems or group projects with other students in an environment where everyone’s main concern wasn’t getting a better grade than you. Coursework was challenging but also very manageable, and office hours made it easy to get help when needed.

In the first year of my Master’s program, I had the opportunity of working with Dr. Thompson on a pragmatic trial for hypertension coordinated through Columbia University Medical Center. I really enjoyed working on a cross-functional team of clinicians and data managers as a statistician, and Dr. Thompson was always there to provide guidance on statistical methodology, as well as general career advice. I highly recommend taking advantage of the wealth of knowledge surrounding CUMC to explore some research positions, even if the position is only a few hours a week. I also highly recommend taking advantage of New York City – its like no other place I’ve ever lived. Exploring Central Park and having picnics in Sheep’s Meadow is a must (weather permitting). During my second year at Mailman, I lived in Little Italy, which to me was well-worth the commute. There are never-ending things to do downtown, including eating your way through Chinatown and Little Italy, riding city bikes through Greenwich village, brewery hopping in Brooklyn and seeing comedy shows in West Village. It is so easy to get around by taking the subway, but it’s also really nice being able to walk around for hours in the city and never be bored.

The practicum allowed me to work as Biostatistics intern at Regeneron Pharmaceuticals, which paved the way for my post-graduate position at Regeneron. I was able to fulfill the practicum requirement the summer after my first year and presented my research at a Statistics conference shortly after the end of internship. Since starting at Regeneron almost 6 months ago, I’ve had the opportunity to work as a Statistician on multiple studies for three different investigational products. The study I have found most interesting, and the one that takes up most of my time, is for our investigational treatment for a high-fatality viral disease. The project has been challenging but very exciting, and being able to see the impact this drug has made from a global perspective has been exceptionally rewarding. Coursework in epidemiology, longitudinal data analysis and survival analysis was beneficial in preparing me for this role. As the study takes place in a third world country, communicating with sites has been a challenge, and data management skills as well as techniques for handling missing have been important. Altogether it has been a fascinating study to be a part of. I am very much looking forward to continuing my career as a Biostatistician and am thankful for the experience I had at Columbia in helping me prepare for this career.
Alumni Spotlight

NICK WILLIAMS, MPH
Research Biostatistician
Weill Cornell Medicine

What were some of your memorable moments as a student in the department?

My most memorable moments were those as a teaching assistant. Working as a teaching assistant exposed me to the other side of education and gave me the opportunity to develop professional relationships with the faculty in the department. In my experience, the adage “the best way to learn is to teach” is also very true.

Can you tell us a little about your role at Weill Cornell?

As a Research Biostatistician at Weill Cornell I assist with the statistical side of a variety of long-and short-term research projects that are taking place across the medical college and at a couple of outside institutions. Depending on the project, I might be designing statistical protocols, conducting analyses, developing randomization schemes, etc. It’s a great first job out of graduate school because I’m exposed to a wide range of healthcare related domains; as an example, I’ve been involved with projects focused on new cancer therapies, orthopedic surgery, rheumatology, and health disparities.

How did Columbia Biostats experience prepare you for your career?

When I first started my job at Weill Cornell, I remember the fear of realizing the training wheels were coming off. I no longer had my mentor to lean on and others were going to expect me to have the answers to some of their problems. But, the biostatistics department at Columbia gave me that necessary foundation to lean on. As a student, I was challenged in a diversified manner that instilled the confidence necessary to translate lessons in the classroom to real-world public health application.

What skills (other than quantitative) do you think are important for a biostatistician to possess?

I think effective communication skills are the most important “soft” skills for a biostatistician to possess. This includes both written and oral communication. I think some of the best biostatisticians are those that can switch between explaining a topic between different audiences (principal investigators, journals, colleagues, and lay audiences).

As a member of the first 2-year MPH Core cohort to graduate, what advice would you give to current and future MPH students in the department?

I guess I have a couple of pieces of advice. 1. Keep learning about statistics outside of school; it’s impossible to really dig into every topic in biostatistics in two years and you will be a better professional statistician if you develop a passion for learning about the field beyond your formal education. 2. Don’t just go through the motions—R.A. Fisher wrote that statistics is the “study of methods of reduction of data”; while Fisher contributed so much to statistics, I have to disagree, statistics isn’t just a collection of methods and answers don’t just lie in the data. Proactively think about the problem at hand instead of blindly applying the methods. 3. When you graduate with an MPH in biostatistics you aren’t just a statistician, you are also a public health professional. Keep that dual-role in mind as you move on to your career.
Can you talk about your experience as doctoral student at Columbia? What were you expecting coming in and what surprised you?

Since I went through the MS program at Columbia first I had some idea of what to expect going in. I knew that I would have supportive faculty to work on projects with, as well as an abundance of biomedical research areas to draw from across the Mailman and CUMC campuses. I was surprised by how challenging the PhD classes were in comparison to the MS classes. In spite of the difficulty, some of these became my favorite classes (shout out to advanced computing in particular).

Congratulations for your move as assistant professor in Colorado! How did the doctoral program prepare you to make this transition?

Thanks! I’m ecstatic to be here. It was a thrill the first time I walked into my office and saw “Assistant Professor” etched on a plaque next to the door. Luckily, I feel that my training at Columbia prepared me incredibly well for this transition. As a student, the graduate student research seminars (GSRS) gave me invaluable practice sharing my research. My advisor, Jeff Goldsmith, was great about introducing me to his mentors and collaborators and inviting me to speak at conferences. This taught me how to communicate my work clearly and effectively to all kinds of people. Since I spend much of my time establishing relationships with scientists in other disciplines and translating their ideas into statistical analyses, these experiences really set me up for success.

What are the things that excite you most about your new position?

I have so much new data to work with! I’ve enjoyed fostering relationships with new collaborators and determining where my area of statistical expertise fits into their scientific areas. It also doesn’t hurt that my department has an official ski-during-work day.

We all know in the Department that you enjoy the outdoors how did you do it in New York?

I started climbing at a rock gym regularly during my second year as a PhD student, then started climbing outside up in New Paltz by my fourth year. In Colorado I now have a group of colleagues I climb with in the gym and outdoors (pictured). As a student I also took advantage of “free” travel and went to conferences in neat locations whenever possible. I encourage current students to come join me for WNAR 2020 in Anchorage, Alaska!
Significant Moments

Publications


S Leng, S Lentzsch, Y Shen, WY Tsai, JD Wright, DL Hershman (2018) Use and impact of herpes zoster prophylaxis in myeloma patients treated with proteasome inhibitors. Leukemia & lymphoma 59 (10), 2465-2469.


Bellavia, A., Centorrino, F., Jackson, J. W., Fitzmaurice, G., & Valeri, L. (2018). The role of weight gain in explaining the effects of antipsychotic drugs on positive and negative symptoms: An analysis of the CATIE schizophrenia trial. Schizophrenia research.


Significant Moments


Ya Wang, Min Qian, Peifeng Ruan, Andrew E. Teschenendorf, Shuang Wang (2019) “Detection of differentially methylated genes using weighted epigenetic distance-based methods” Nucleic Acid Research, 47(1) e6


And we have fun!
We Want to Hear from You!

We would like to share information about your accomplishments and activities. If you have changed jobs, received a promotion or award, earned a new degree, or if anything new has happened to you, please let us know. Send information to Justine Herrera: jh2477@cumc.columbia.edu along with any updates in your mailing and email addresses.

Follow Us

facebook.com/Columbia.University.Biostatistics
facebook.com/ColumbiaPublicHealth
linkedin.com/groups/5116323/profile
twitter.com/ColumbiaMSPH
instagram.com/columbiapublichealth
youtube.com/ColumbiaMailman

Acknowledgements

The many contributors to Significant Moments include members of the Department of Biostatistics Communications Committee (Seamus Thompson, Linda Valeri, Cheng-Shiun Leu, Corey Adams, Caleb Miles, Prakash Gorroochurn, Cody Chiuzan, and Melanie Wall); the Department Administration (Justine Herrera, Jessica Jimenez, and Anthony Guerrero); the Mailman Communications Office (Diana Gazzia); other faculty members (Christine Mauro); and graduate students (Angel Garcia de la Garza and Erin McDonnell).

Support Biostatistics

The Mailman School of Public Health is funded by individuals, foundations, corporations, and organizations who share the School’s unwavering commitment to addressing critical and complex public health issues that affect millions.

TO DONATE: MAILMAN.COLUMBIA.EDU/GIVE