## In this Issue

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message from the Chair</td>
<td>2</td>
</tr>
<tr>
<td>Faculty</td>
<td>3</td>
</tr>
<tr>
<td>Promotions &amp; Awards</td>
<td>3</td>
</tr>
<tr>
<td>Interview with Yuanjia Wang, PhD</td>
<td>4</td>
</tr>
<tr>
<td>Interview with Jeff Goldsmith, PhD</td>
<td>5</td>
</tr>
<tr>
<td>Interview with Gen Li, PhD</td>
<td>6</td>
</tr>
<tr>
<td>Interview with Caleb Miles, PhD</td>
<td>7</td>
</tr>
<tr>
<td>Interview with Linda Valeri, PhD</td>
<td>8</td>
</tr>
<tr>
<td>Grant Awards</td>
<td>9</td>
</tr>
<tr>
<td>Postdocs</td>
<td>12</td>
</tr>
<tr>
<td>Postdocs</td>
<td>12</td>
</tr>
<tr>
<td>Students</td>
<td>13</td>
</tr>
<tr>
<td>Student Data</td>
<td>13</td>
</tr>
<tr>
<td>Statement of Teaching Philosophy by Jihui Lee, PhD</td>
<td>14</td>
</tr>
<tr>
<td>Student Perspective: Chirag Shah</td>
<td>16</td>
</tr>
<tr>
<td>BEST Diversity Program</td>
<td>17</td>
</tr>
<tr>
<td>Student Awards &amp; Defenses</td>
<td>18</td>
</tr>
<tr>
<td>Staff</td>
<td>19</td>
</tr>
<tr>
<td>Staff Spotlight: Georgia André</td>
<td>19</td>
</tr>
<tr>
<td>Staff Spotlight: Richard Buchsbaum</td>
<td>20</td>
</tr>
<tr>
<td>Staff Spotlight: Justine Herrera</td>
<td>21</td>
</tr>
<tr>
<td>Alumni</td>
<td>22</td>
</tr>
<tr>
<td>Alumni Spotlight: Rose Linton</td>
<td>22</td>
</tr>
<tr>
<td>Faculty Publications</td>
<td>25</td>
</tr>
<tr>
<td>Follow Us</td>
<td>32</td>
</tr>
</tbody>
</table>
Welcome to Significant Moments. I am delighted to take this opportunity to share news of our Department of Biostatistics with you. Our Department is at the forefront of health improvement through creative biostatistical innovations in research, education, and the health sciences. This issue highlights some of our accomplishments in 2018, and, I hope, conveys my excitement about the Department’s future and the significance of our collective efforts.

As you know, biostatistics is a branch of statistical data science that focuses on biomedical research and public health. Biostatisticians provide the rigorous analytical methods needed to understand and properly interpret data in order to convert them to actionable knowledge. The statistical theory that underlies these methods enables us to distinguish signal from noise with confidence, reliably identify bias, and critically evaluate assumptions underlying our inferences about causality. In a data-intense era characterized by massive amounts of information from complex networks, it is difficult to overstate the importance of our rigorous data science approach to the public health issues faced by our society, and the clarity it offers.

Every bit as pressing as the need for research of the highest quality is the need to train the next generation of data scientists. Our Department boasts one of the largest graduate programs in biostatistics in the country, with many students from diverse backgrounds. We take pride in our innovative curriculum, which provides new course offerings and streamlined coursework. Our students learn technical and quantitative skills and are trained in collaborative leadership and communication skills. We strive to deliver an education that equips graduates to excel. We expect that our students will continue to find outstanding job placements at the same high rates as we’ve seen in recent years. To our many talented alumni and friends reading this newsletter, thank you for your support. Please stay in touch!

Our highly successful BEST (Biostatistics Enrichment Summer Training) Diversity Program completed its 11th year, enrolling 14 students this past summer. BEST introduces biostatistics to undergraduates from underrepresented or disadvantaged backgrounds. I am proud of our vision in creating this pipeline program and am committed to continuing to support diversity in our field.

We have seen several changes to our faculty this year. We congratulate Dr. F. DuBois Bowman, our departing Chair, on his Deanship of the University of Michigan School of Public Health. We all have benefited immensely from his outstanding vision, leadership, and innovation. This year we welcome Linda Valeri and Caleb Miles, two outstanding scholars with expertise in causal inference. Look out for interviews with both of them in this issue. Jianhua Hu also joined the Department and leads the biostatistics core resource at the Herbert Irving Comprehensive Cancer Center.

We congratulate Gen Li, the inaugural recipient of Sanford Bolton Faculty Scholars Program, which supports innovative translational methodological research by junior faculty. Dr. Bolton was trained in biostatistics by John Fertig, our School’s first head of Biostatistics. Dr. Bolton was a distinguished alumnus (M.S.) when Biostatistics was headed by Dr. John Fertig. Dr. Bolton was also an early benefactor of the BEST Diversity Program. I thank Dr. Bolton and his estate for his generous contributions and his visions on Diversity.

In the last two months I have had the privilege of working closely with our immensely talented faculty and staff. I have been deeply humbled by your dedication to the Department and our research and education goals. Thank you!
New Faculty
Welcome new faculty!

CALEB MILES, PhD, Assistant Professor, is joining us from University California, Berkeley, CA

LINDA VALERI, PhD, Assistant Professor, is joining us from McLean Hospital and Harvard Medical School, Boston, MA

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

YUANJIA WANG, PhD has been promoted professor

JEFF GOLDSMITH, PhD has been promoted associate professor with tenure

Faculty Awards
Calderone Research Prize for promising, high-impact research proposals from Junior Faculty, 2017

ADAM CIARLEGLO, PhD, Assistant Professor of Clinical Biostatistics, for “Statistical and Machine Learning Methods for Integrating Clinical and Multi-modal Imaging Data to Select Optimal Antidepressant Treatment.”

YIFEI SUN, PhD, Assistant Professor, Biostatistics, for “Tree-based Methods for Dynamic Risk Prediction.”

Sanford Bolton Faculty Scholar Award, 2018

GEN LI, PhD, Assistant Professor of Biostatistics

Tow Faculty Leadership Scholars, 2018-2020

JEFF GOLDSMITH, PhD, Associate Professor of Biostatistics
What do you most enjoy about working at Columbia School of Public Health?

The research environment is great, the continuous interaction with collaborators is inspiring. There are even too many options of collaborations to choose from. I manage to stay focused by following my instinct and choosing the topics that I am more interested in, which at first sight might not necessarily seem the most impactful. I am interested in developing data-driven approaches for exploratory research questions, when data available to discover new targets are not necessarily in a large scale but with great quality.

Here at Columbia we have the chance to work with all kinds of data. Helping with medium or small but high quality data, where investigators pay careful attention, can matter a lot. For example, I am proud of one project where I helped assessing the penetrance for a causal gene of Parkinson’s disease. Previous studies yielded conflicting results (20%-100%) but with our study and method we established a solid estimate (25%).

You devote a lot of your work in improving Personalized Medicine. How do you define it and why, in your opinion, it’s important in Public Health?

My research angle on the topic is optimal treatment selection based on individual patients’ characteristics. I think this is very useful when applied to mental disorders where, due to the complexity of the diseases and substantial heterogeneity, one size fits all type of solutions do not work. Clinicians approach treatment selection by trial-and-error. We can do more than tailoring only based on previous drug response. I am working towards changing this narrow paradigm by developing statistical methodology to integrate clinical trials information beyond previous treatment course, adding more contextual factors. The biggest challenge in this pursuit is large variability. We get signal, but variance dominates, so we integrate trials and remove symptoms that purely add noise. We are realizing that at times less is more, and we need new statistical tools to eliminate noisy and redundant variables. I am also exploring which other factors we should add into the picture, for example biological targets, such as neuroimaging markers are of interest (binding potential at brain regions). It’s a balance between good quality and dimension and it’s hard to identify novel characteristics in small samples, so I proposed to learn from patients’ electronic health records in a recent grant.

What attracts you about New York and made you stay for such a long time?

It was Karma to get into biostatistics and working closely to the fields of neurology and mental health. In New York there is a wealth of institutions strong precisely in these fields, Columbia in particular. My grandfather was the founding chair of the Neurology Department in a medical university in my hometown. No one studied medicine in the third generation of my family, but I ended up doing something related and here at Columbia is where I believe I can do it best. On top of this, I really wanted to stay in New York to keep experiencing the diversity I encountered during my doctoral studies at the Statistics Department at Columbia University. The diversity within our Department of Biostatistics and at Mailman School of Public Health is a mirror of our city and I love it. Now I moved to NJ for family reasons, especially for good public schools. However, we are very connected to all New York city fun. Every once in a while we even come and stay in a hotel to watch shows and enjoy great food. Finally, I love the people of New York, despite the reputation. They are not so cold, rather they are open minded and not judgmental, which is good for young people to experience.
Interview with Jeff Goldsmith, PhD

Jeff Goldsmith is an associate professor in Biostatistics at the Columbia University Mailman School of Public Health. Dr. Goldsmith joined Columbia after receiving his PhD in Biostatistics from Johns Hopkins in 2012, where his dissertation focused on statistical methods for high-dimensional structured data. Dr. Goldsmith has research interests in scientific domains including neuroimaging, physical activity monitoring using accelerometers, motion kinematics and motor learning, and urban environments. In these domains, he develops statistical methods that examine relationships between complex data structures and patient-level information; examples include using MRI scans to predict patient function and examining effects of aging on daily activity patterns.

What inspired you to pursue a career in Biostatistics?

I majored in mathematics as an undergraduate, and really enjoyed it as an intellectual challenge—mathematics is rigorous, detail-oriented, and logical. However, especially in more advanced topics, the material is disconnected from real-world problems, and applications of “pure” math become increasingly hard to find. As I prepared to apply to graduate school, I came across biostatistics (largely by chance). After a bit of digging, I realized that the field was a perfect fit for my professional goals and quantitative background. In my position in Biostatistics, I get the intellectual engagement that comes from working on challenging problems, along with the satisfaction of working to advance public health.

What is the most exciting project you’re working on right now?

For several years, I’ve worked with collaborators on stroke recovery and motor control. These are hard problems that require understanding the components that go into “motor control”, determining how these are affected by stroke, and evaluating possible therapeutic interventions. The good news is that we have innovative experiments, rich data, and promising avenues for benefitting patients; even better, the interventions we try can be made available to patients almost immediately.

In your perspective, what opportunities for biostatistics arise from the emergence of data science?

Biostatistics is central to data science, especially in public health and medicine. We work closely with collaborators in a variety of fields and can have an impact based on our understanding of the domain; we also understand sources of variability and can identify complex patterns in data; and we have broad expertise in inference, traditional statistical methods, and new tools for statistical learning. We now have the opportunity to hone particular skills, especially related to the best practices for writing code to work with data, and to engage with datasets of unprecedented scale and complexity and that require novel approaches.

How do you take a break from work?

One of the best things about New York is that, even if you’ve been here a few years, there’s always something new. On the weekends, I explore new neighborhoods, try new restaurants, see new museum exhibits, and try to take advantage of what the city offers. I also like to revisit some favorite spots, especially in Central Park, or even take a break from the city and head upstate!
Congratulations on this great honor, Dr. Li! Can you tell us about the new methods that you plan to develop with the support of this award?

First of all, I’m really honored to be the inaugural recipient. The award provides tremendous support by securing my time to conduct methodological research and establish new collaborations. I plan to develop tensor methods for RNA-Seq data. High-throughput RNA-Seq data are frequently measured under multiple conditions from the same set of samples, naturally forming a multi-way tensor array. Critical genomic questions such as co-expression network analysis can be effectively addressed via tensor analysis. However, existing methods cannot be directly applied to sequencing data, which typically have special features such as count values, zero inflation, and/or block-wise missing data. I am committed to developing new statistical methods to tackle these problems.

How will this work translate to real-world public health impact?

The application of the methods which we will develop to large genetic data, sets such as Genotype-Tissue Expression (GTEx) multi-tissue data, will allow us to address fundamental biological questions (e.g., the nature and source of variation between tissues) and to provide critical new insights into human genomics and their relation to health.

What role have students played in your research?

I always enjoy the privilege of working with our students. Depending on the student’s interest and training, I identify an applied or methodological project to start our collaboration. I have worked with several PhD and more than 30 masters students. We have finished some statistical methods manuscripts and dozens of domain-specific conference abstracts in applied fields such as Parkinson’s Disease, Urology, EHR, and Sleep Studies.

What do you like most about being a biostatistician?

Learning domain-specific knowledge from every collaboration. It is like going to medical school without paying tuition or taking exams! John Tukey once said, “The best thing about being a statistician is that you get to play in everyone’s backyard.” I couldn’t agree more.

What do you enjoy doing outside of work?

I enjoy exploring things that I’ve never done before: learning new skills, going to new places, watching new movies, eating new food, etc. After the birth of my daughter, I enjoy keeping her company and having fun with her.
What made you decide to pursue a career in biostatistics?

I’ve loved math since I was in preschool, and always knew I wanted to do something involving math. However, I also really wanted to go beyond just satisfying my own intellectual curiosity and do something that would serve people suffering from disease and poverty. I first learned about biostatistics from an announcement for a summer program in an email from my college math department, and it seemed like the perfect intersection of these two interests. I decided to go to graduate school to study biostatistics, and several years later, here I am!

What is your main research area?

In terms of statistical methodology, it is developing semiparametric methods for causal inference. Causal inference is basically all about learning causal relationships from patterns in data. Using semiparametric models allows us to do this under minimal assumptions about how observed variables are related to one another, e.g., we don’t necessarily have to make assumptions about linearity or normality. Most of my methodological work has been inspired by and applied to HIV research, such as learning about what drug adherence can tell us about why some HIV drug treatments outperform others, and the effectiveness of task-shifting public health programs, in which low-risk HIV patients have health care tasks shifted from physicians to nurses in resource-limited settings.

Why did you choose Columbia University, Mailman School of Public Health?

I chose to come to Columbia for a number of reasons. The main one was the strength of the department and the quality of students it attracts. Another was the strength of the school of public health as a whole, which provides ample opportunity to collaborate with leaders in any field within public health (and beyond). Finally, I was very drawn to live in New York City where I already had friends, and which is obviously one of the most exciting cities in the world to live in.

With all the recent excitement around data science, how do you see biostatistics fitting in?

Biostatistics is a key component of data science. Of course, data science goes beyond just classical statistics, but without sound statistical reasoning, “data science” can quickly veer into what some call the “danger zone”. So while on the one hand, we are making important contributions in, for instance, developing and incorporating machine learning into the health sciences, we also provide the skills to interpret input and output scientifically, quantify uncertainty, and generally root data science in the necessary statistical theory so that we aren’t swept away by trends or ad-hocery.

What is a typical day like for Dr. Caleb Miles?

I don’t know that I’ve really established a “typical” day yet. Most days consist of some combination of writing, meetings, reading, proving theorems, and writing code. After work I like to go rock climbing or running in Central Park for exercise. It feels like most weekends I’ve been lucky enough to have friends or family in town visiting—another benefit of living in New York!
Interview with Linda Valeri, PhD

What made you decide to pursue a career in biostatistics?

After I gained my degree in Economics and Social Sciences, I followed my interest in causal inference, which emerged after I attended an econometric class, and my passion for improving the quality of life for all. My doctoral studies in Biostatistics confirmed these intuitions.

What is your main focus on research now?

My interests now include developing new approaches to causal inference; understanding complex mechanisms such as mediation and interaction; improving statistical methods for analysis of biases such as measurement error and unmeasured confounding; and the application of these approaches to environmental health and psychiatry. I am also very interested in harnessing digital technologies to better capture environmental and social exposures, and to transform complex data streams into actionable information using machine learning and causal inference to foster our understanding of mental health.

What are your first impressions since joining Columbia?

Since I arrived I have been amazed by the sense of community in the Department of Biostatistics and the Mailman School of Public Health more broadly. The Dean has really made us feel welcome, and it’s exciting to be part of this special place. I am also very happy that in the Department we have an excellent group of Assistant Professors who are eager to help each other, which is not a small thing. The more senior faculty and in addition the staff are also always available. These are the things that have impressed me the most.

From your perspective, what are the major current challenges faced by biostatisticians?

Well, a practical challenge is funding. So far, for me it has been fun to bring people together and to join new teams to write proposals. I focus on learning about the science from my collaborators. This takes away the pressure and makes grant writing a great experience.

How do you like living in New York City?

I love it. So far the most beautiful experiences for me have been...having access every day to Central Park: greenness really matters for me. Music is life and New York is full of it. I have never in my life met so many artists (composers, directors, painters) as in the three months I have been living in New York City. I hope to bring this energy and creativity in my own work.
NEW GRANTS

Iuliana Ionita-Laza (Competitive Renewal)
R01MH095797 funded by the National Institute of Mental Health (Role: PI)
(Award: 07/01/18-03/31/22)
“Novel Statistical methods for DNA Sequencing Data, and applications to Autism”

Autism Spectrum Disorders are common diseases with major impact on public health.

Although coding variation has been extensively studied for its role in affecting risk to autism, the analysis of non-coding variation poses substantial statistical challenges. The proposed statistical methods and their applications to nearly 20,000 whole genomes from three large autism whole-genome sequencing studies will improve our understanding of the biological mechanisms involved in autism with important implications for disease treatment strategies.

Gen Li
R03 DE027773 funded by the National Institute of Dental & Craniofacial Research (Role: PI)
(Award: 09/15/18-8/31/20)
“Multivariate analysis of microbial absolute abundance in population-based studies”

The human oral microbiome plays an integral role in many oral and extra-oral diseases. By developing novel statistical methods for association analysis and applying to an ongoing longitudinal study, the project will further reveal the relationship between oral microbiota and metabolic disorders, and inspire potential therapeutic approaches.

R. Todd Ogden (Competitive Renewal)
R01MH099003 (subcontract to CU) funded by the National Institute of Mental Health (Role: Subcontract PI)
(Award: 03/15/18-01/31/22)
“Biotyping placebo and treatment-specific responses for precision medicine”

Using complex data collected at baseline, Dr. Ogden is working on methods to characterize placebo response, to identify placebo responders, and to help make patient-specific decisions about treatments.

Martina Pavlicova (Competitive Renewal)
P30 MH43520 (subcontract to CU) funded by the National Institute of Mental Health (Role: Subcontract PI)
(Award: 03/23/18-01/31/2023)
“HIV Center for Clinical and Behavioral Studies”

Dr. Martina Pavlicova is StAD Core Director and Subcontract Principal Investigator. Dr. Pavlicova will work with Dr. Tross to ensure that the Core meets its overall goals. She will lead monthly Core Leadership Meetings with Drs. Tross and Campbell, to review and plan ongoing Core activities; co-lead quarterly full StAD Core meetings for problem-solving and strategic planning with Core members; and participate in Cross-Core Meetings. She will take the lead in overseeing the Core’s data analysis and management teams with Dr. Wall and Mr. Bauschbaum, respectively. Dr. Pavlicova will consult with HIV Center established and early stage investigators on grant proposals, helping them to establish statistical expertise in their research teams and choose appropriate rigorous analytic strategies. She will contribute to the development and implementation of workshops and seminars in her (above-described) areas of expertise and interest. She and Dr. Tross will work closely with the Core Scientific Coordinator in implementing the Core’s functions and activities. As Core Director, Dr. Pavlicova will serve on the HIV Center Core Committee

Yuanjia Wang
R21MH117458 funded by the National Institute of Mental Health (Role: PI)
(Award: 09/01/18-08/31/20)
“Integrative Learning to Combine Evidence for Personalized Treatment Strategies”

Treatment responses for mental disorders are inadequate and considerable heterogeneity is observed, in part because an individual patient’s, clinical, psychosocial, and/or biological markers are not accounted for when selecting treatments among available options. This research develops analytic methods to discover new powerful, yet interpretable personalized treatment strategies and integrate evidence of strategies identified in multiple prior studies to increase robustness and reproducibility.
R01GM124104 funded by the National Institute of General Medicine Sciences (Role: Multi-PI)
(Award: 04/01/18-/03/31/22)
“Efficient Statistical Learning Methods for Personalized Medicine Using Large Scale Biomedical”

This project develops scalable statistical learning methods to analyze electronic health records (EHRs) and proposes to use two real-world, high-quality EHR databases for personalized medicine research. The methods will handle the non-experimental nature of data collection process, along with heterogeneous data types, dynamic treatment sequences, and the trade-off between benefit and risk outcomes. The results will complement the current knowledge base for individual patient care using evidence generated from patients in real-world clinical practices.

Linda Valeri
K01MH118477 funded by the National Institute of Mental Health (Role: PI)
(Award: 09/19/18-08/31/22)
“Statistical methods for the assessment of social engagement in psychosis using digital technologies”

Complex psychiatric diseases, such as chronic psychotic disorder, are major public health issues in the United States. Combining causal inference and machine learning principles, the proposed research provides an innovative framework and develops powerful and computationally efficient statistical methods to integrate active (e.g. survey) and passive (e.g. GPS, text and call log) data streams from mobile sensors for the discovery of behavioral targets of treatment for chronic psychosis.

CONTINUING GRANTS

Melissa Begg
• NIH/NHLBI (R25 HL096260), “BEST- DP: Biostatistics & Epidemiology Summer Training Diversity Program,” 2009-2019 (Contact PI; other Multi-PIs include F.D. Bowman)

F. DuBois Bowman
• NIH/NINDS (R01 NS090677), “Customized Cortical Stimulation Therapy in the Rehabilitation of Stroke Patients,” 2015-2019 (Role: Sub-contract PI)

Ying Kuen (Ken) Cheung

Min Qian

Yuanjia Wang
• NIH/NINDS (U01 NS082062), “Identifying Huntington’s disease markers by modern statistical learning methods,” 2014-2018 (Role: PI)
• NIH/NIDA (R01 DA035846), “Impulsivity in Cocaine Abuser: Relationship to Drug Taking and Treatment Outcome,” 2014-2019 (Role: Multi-PI)
• NIH/NINDS (R01 NS073671), “Statistical Methods for early disease prediction and treatment strategy estimation using biomarker signature,” 2011-2021 (Role: PI)

Ying Wei
• NIH/NHGRI (R01 HG008980), “Develop Quantile Analysis Tools for Sequencing and EQTL Studies,” 2016-2020 (Role: PI)

Jeff Goldsmith
• NIH/NINDS (R01 NS097423), “Functional data analytics for kinematic assessments of motor control,” 2016-2021 (Role: PI)
• NIH/NIBIB (R21 EB018917), “Generalized, multilevel functional response models applied to accelerometer data,” 2015-2018 (Role: PI)

• NIH/NHLBI (R01 HL123407), “Statistical methods for biosignals with varying domains,” 2014-2018, (Role: Sub-contract 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 (R01 MH095797), “Novel Statistical methods for DNA Sequencing Data, and applications to Autism,” 2012-2018 (Role: PI)

• NIH/NIMH (R21 MH106888), “Applications of novel statistical methods to CNVs in autism and schizophrenia,” 2015-2018 (Role: PI)

Shing Lee

Bruce Levin
• NIH/NIMH (P30 MH043520), HIV Center for clinical Behavioral Studies,” 1987-2018 (Role: Sub-contract PI and Director of the Statistics, Epidemiology and Data Management Core)

Gen Li

Ian McKeague
• NIH/NIGMS (R01 GM095722), “Point Impact and Sparsity in Functional Data Analysis,” 2011-2019 (Role: PI)

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)

Zhezhen Jin
• NIH/NHLBI (R01 HL119485), “Novel viruses and viral dynamics in multiple transfusion recipients,” 2014-2019 (Role: Sub-contract PI)

R. Todd Ogden
• NIH/NIMH (R01 MH099003), “Characterizing Placebo Response,” 2013-2022 (Role: Sub-contract PI)

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


John L. P. Thompson
• FDA (R01FD005407), “Phase 3 Trial of DCA in PDC Deficiency IND 028,625,” 2016-2021 (Multi-PI)
Tian Wang, PhD

Tian Wang obtained his Ph.D. in Mathematics from Washington University in St. Louis under the mentorship of Professor Jimin Ding. He is currently working with Professor Shuang Wang. His major current project involves Microbiome, which has proven to be related to human health and disease. The tentative goal is to utilize the phylogenetic tree of operational taxonomic units to set up models for Microbiome data. In his leisure time, he enjoys swimming and hiking.

Tianying Wang, PhD

Tianying Wang earned her PhD from the Department of Statistics at Texas A&M University in 2018, supervised by Dr. Raymond Carroll and Dr. Irina Gaynanova. Tianying’s research interests are at the intersection of applied, theoretical and methodological statistics. She collaborates with national and international domain scientists and statisticians. Methodologically, she works on measurement error analysis, misspecified models, gene-environment interaction analysis, multivariate analysis and high-dimensional data analysis such as variable selection and classification. With respect to specific areas, she is primarily interested in cancer genomics, case-control studies, and epidemiology studies. She has worked on a variety of applied problems such as misspecified model subject to measurement error, high-dimensional binary classification with dimension reduction, semiparametric analysis of complex gene-environment interactions in case-control study, and Bayesian analysis for the reproducibility of psychological science. Currently, Tianying is working with Dr. Ying Wei and Dr. Iuliana Ionita-Laza on quantile regression and genomics data. Now she is developing an efficient quantile test for group-wise association with application in sequencing data.

Danqing Xu, PhD

Dr. Danqing Xu earned her PhD in Statistics from the Department of Statistics and Applied Probability at University of California, Santa Barbara (UCSB). Under the supervision of Yuedong Wang, Danqing’s doctoral research developed approaches to fit smoothing splines to massive data sets. She has also worked on clustering chronic kidney disease progression using longitudinal measurements of multiple clinical, laboratory and functional indicators. Danqing’s research interests include smoothing spline, big data, scalable computation, functional/longitudinal data analysis, biomedical data analysis, machine learning and statistical genetics. Danqing works with Dr. Iuliana Ionita-Laza on functional annotation for genomics. She is currently focusing on statistical methods to identify functional regions in genome.
Department Data: Students

Number of 2018 graduates (including February 2018 and October 2018)
DrPH: 4
PhD: 5
MS: 77

Number of Returning for Fall 2018
DrPH: 11
PhD: 22
MPH: 12
MS: 97

Number of Incoming for Fall 2018
DrPH: 1
PhD: 6
MPH: 14
MS: 109

Where our 2018 graduates have found jobs:
• Acumen
• Analysis Group
• Analytica/LASER
• Axio Research
• Beth Israel Deaconess Medical Center
• Brightech International
• Charles River Associates
• Columbia University Medical Center
• EcoHealth Alliance
• Hospital for Special Surgery
• Icahn School of Medicine at Mount Sinai
• Genesis Research
• Kite Pharmacy
• Memorial Sloan Kettering Cancer Center
• Michael Allen Company
• Mount Sinai Health System
• National Consumer Panel
• Novartis
• Office of Inspector General of the U.S. Department of Health and Human Services
• Oregon Health & Science Health System
• Weill Cornell Medicine
• Zoetis, Inc.

Where our 2018 graduates are pursuing a doctoral degree:
• Columbia University, Biostatistics
• University of Washington, Global Health Metrics
• University of Florida, Biostatistics
• University of Connecticut, Statistics
• University of Maryland, Health Services

Class of 2017 Employment Stats:
- 90% Employed
- 9% Continuing Study
- 1% Seeking
Statement of Teaching Philosophy by Jihui Lee, PhD

What I love the most about teaching is to share my joyful experience of learning new concepts. I believe that joy of learning is contagious; I value how my enthusiasm for statistics could positively influence students’ appreciation of course materials. Especially, I enjoy the process of filling in the gaps between steps of proofs and demonstrating key points of theorems. I understand the frustration that students may encounter, and it is extremely rewarding when they appreciate how much thought I have put into how best to explain complicated concepts in statistics as clearly as possible. I also like to introduce multiple ways of approaching the same problem and connect key concepts to a bigger picture, both of which are valued by my students.

My principles as a teacher are threefold: be creative, be interactive, and be prepared.

Be creative: I do not hesitate to utilize various resources in class. First, it is always helpful to demonstrate complicated mathematical concepts on the board: dissecting a formula into pieces or visualizing a concept into diagrams or maps. Students find it easier to recall the maps I drew in class compared to memorizing a complicated formula. I also like to utilize the eyes in our mind; R. A. Fisher could successfully develop his geometric sense (despite his severe vision impairment) by learning to visualize mathematical problems in his mind. I find this approach exceptionally helpful in my programming class. Before executing codes, I ask my students what they think the code will produce. This step helps them understand the logic of the code. Lastly, I always seek for various online resources. For example, I have used online polls to engage students’ participation in class. Conducting real-time online polls in class prompts shy students to voice their thoughts and feel engaged in class. Also, I often recommend some online math games that help students understand abstract theoretical concepts. Entertaining resources help students get an intuitive sense of complicated concepts more efficiently.

Be interactive: The fact that students bring different levels of statistical maturity and familiarity of course materials can present a challenge. However, I as a teacher find such heterogeneous environment invigorating. In class, students with different field of expertise present diverse perspectives in comprehending the same concept. It is exciting to see how students can learn a lot from one another through active discussion in class. My goal in the beginning of a semester is to set up a friendly, supportive environment in class so that students feel free to ask questions and share their comments. As a result, I am approachable and students feel comfortable initiating a discussion in class. In a theoretical class, for example, having students present their way of solving problems is very helpful. Not only does it encourage student participation, but it also alleviates the seriousness of the class.

Be prepared: In classroom, things do not work as expected. I learned from my experience that it is important to be prepared for everything. When preparing a lecture, I try to view my lecture in the perspective of a student and make sure to demonstrate the materials succinctly and straightforwardly. It also helps to anticipate the questions that students may ask in class. Especially in the field of biostatistics, applying theoretical concepts to practical scenarios is a valuable quality, as it directly relates to the value of statistics in scientific discovery. Thus, I always prepare a series of examples outside of the textbook to illustrate the utility of theory in practice. I consider that my job as a teacher is to guide students to build a better picture that connects all concepts in statistics. In order to achieve that, I provide a solid sketch of the course in the beginning of a semester and clarify a goal of each lecture clearly.

My teaching experience at Columbia University Medical Center (CUMC) has greatly contributed to shaping my philosophy in teaching. Most importantly, CUMC has brought such a diverse audience to my class over the years. A majority of students in my statistical programming class are from the Mailman
School of Public Health, and have no or little analytic skills. Introducing the use of statistical software to conduct quantitative studies is, to them, ‘drawing on empty canvas’. It comes with a great responsibility, but is extremely rewarding to see them being more confident. Besides, taking a lead in summer programs has been extremely fulfilling. The Biostatistics Enrichment Summer Training (BEST) Diversity Program is an inspiring program that introduces the field of biostatistics to undergraduate students that are either underrepresented or with disadvantaged backgrounds. For the last four years, I have taught the fundamentals of biostatistics to the BEST participants who are eager to learn. Motivated to share my passion in biostatistics and help them maximize their experience, I voluntarily have taken the lead in arranging a variety of activities such as journal club, discussion sessions, and seminars. I view that my role as a teacher is not only to convey knowledge to students but also to motivate and mentor students.

Nothing is more exciting to see a sparkle of interest in students’ eyes.

Ultimately, I believe that a good teacher greatly influences students’ perspective in life. I met my inspiring professor in undergraduate, who showed her genuine passion and enthusiasm for statistics. The experience with her made a huge impact on my career trajectory and has motivated me to come this far. Since then, I have dreamed of becoming a teacher who can positively inspire students and help them reach their potential. I am proud to say that many of the students I have taught in the summer program have decided to pursue their postgraduate degrees in biostatistics or a related field, in part motivated by the time they spent at Columbia with me. My ultimate goal is to be a teacher who can not only introduce a fascinating spectrum of both educational and research-focused aspects of pursuing statistics, but also motivate students to go after their passion in life.
Being a guinea pig is never easy. However, despite going through the growing pains as the MPH in Biostatistics program develops, I see it more so as an opportunity to create the foundation and legacy for what Biostatistics MPH students represent at the Mailman School of Public Health, and within the greater context of public health. I have learned that beyond the obvious value in studying biostatistics there is a substantial need to apply those skills to reduce the inequalities that exist between communities around the world. That is where the real reward lies of being a biostatistics student at Mailman.

During this past summer, I spent the majority of my time understanding the disease burden of countries undergoing epidemiologic transition. The rise of so called “western diseases” in developing countries are especially problematic because the healthcare infrastructure that is built to support those populations are ill-equipped to handle the influx of non-communicable disease within that population. As such, it presents a unique opportunity for public health development within those countries to put into place policies that would prevent diseases such as Type II diabetes and cardiovascular disease among many others. It was my job to use regression techniques to identify where specific public health campaigns would be most effective and where those campaigns can yield the best outcomes for the least advantaged. This is a perfect example of how public health and biostatistics intersect. In a world with limited resources for public health, it is these insights that biostatisticians can provide which can help optimize the impact of public health interventions.

The skills that are being taught within the biostatistics department as well as the relationships that I have built with individuals who have similar goals are the keys to unlocking a meaningful career in healthcare. In the short term, I hope to continue my research trying to establish the relationship between glycemic loading and colorectal cancer by using pooling techniques using what I have gained from my MPH degree. Yet, that is only beginning to scratch the surface of what I feel I can accomplish after leaving Mailman. As my fellow peers would tell you, Mailman is just the springboard for lifelong learning as well as lifelong success. With that in mind, I hope to follow a career path that allows me to develop my knowledge and skills, have a deep sense of purpose by solving real world problems that face humans across the globe, whilst also allowing me to study aspects of healthcare that I am truly passionate about (namely cancer and the dietary/environmental factors that give rise to an increase in cancer incidence).

I am extremely proud of being a member of the Biostatistics department because I can see, tangibly, how work that staff, faculty, and my peers are doing can improve the health of people everywhere, every single day. Through our research and projects, we are at the forefront of the push to create a ‘healthy and just world’ as we develop groundbreaking techniques and ask novel research questions that progress the field of public health forward. In spite of the tumultuous political and economic circumstances that we find ourselves in, I am confident that this dedicated community will not falter in their will to make a meaningful contribution to benefit society.
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 2018**
**BEST: 14 students**

**Schools they came from:**
- Amherst College
- Brown University
- California State University-Northridge
- City College of New York
- Hunter College
- Johns Hopkins University
- Rutgers University
- Spelman College
- University of Pennsylvania
- University of Southern California
- University of Virginia
- Vassar College

**Research projects:**
- Characterization of Ovarian Morphology for Diagnosis of Polycystic Ovarian Syndrome
- Assessing Physiological Reactivity to Trauma in Patients after Stroke
- Quantifying Any Potential Association Between In-shoe Heel Sensor and Ground Reaction Force
- The Effects of Continuous Positive Airway Pressure on Hunger and Satiety in Individuals with Obstructive Sleep Apnea
- Opioid Treatment Program Moratoriums & Medication Access
- Validating MRAS as a Candidate Gene in the Pathogenesis of Cardiomyopathy
- Examination of the Change in Chemical Exposures Over Time in the LEGACY Girls Study
Student Awards & Defenses

Graduate Student Instructor Presidential Teaching Award Winner 2018
JIHUI LEE, PhD candidate, Biostatistics

Columbia University downtown data science competition ($20,000) 2017
SHANGHONG XIE AND WODAN LIN, PhD candidates, Biostatistics, October

Society for Clinical trials Thomas C. Chalmers Student Scholarship, 2018
XIAOBAO ZHONG, PhD candidates, Biostatistics, for “Gate-keeping Test for Selecting Adaptive Interventions under General SMART Designs” invited presentation at the 39th Annual Meeting of the Society for Clinical Trials, Portland OR, May

ENAR student paper and JSM Biometrics section student paper awards
JULIA WROBEL, PhD candidate, Biostatistics, for “Registration for Exponential Family Functional Data” presented at ENAR 2018 in Atlanta and at JSM 2018 in Vancouver

ASA Korean International Statistical Society (KISS) Outstanding Student Paper Award, 2018
ANNIE LEE, PhD candidate, Biostatistics, for first authorship of “Efficient Statistical Methods for Genome-Wide Association Studies with Disease Family History Data.” Joint Statistical Meetings (JSM) 2018, Vancouver, July

Sanford Bolton-John Fertig Award for a distinguished doctoral dissertation in Biostatistics
TZU-JUNG HUANG, PhD

Fleiss Award for a Biostatistics student whose outstanding dissertation advances statistical methods and their applications to biomedicine and public health
DANIEL BACKENROTH, PhD

Chair’s Awards for Outstanding Master’s Students in recognition of outstanding academic and research achievements
VICTORIA COOLEY, MS for “An Initial Analysis of Patient Genotypes and Phenotypes in the North American Mitochondrial Disease Consortium Registry.”
JIAYI JI, MS for “Identify Prognostic Biomarkers for Patients with Stage II-III Melanomas.”
YIYI ZHAO, MS for “An Improved Statistical Model for Taxonomic Assignment of Metagenomics.”
Staff Spotlights

GEORGIA ANDRÉ

Georgia is an Administrative Coordinator with the Department of Biostatistics. She joined the Department in 2004 as a receptionist and was promoted in 2012 to Administrative Coordinator. In between she was also the GCD Program Coordinator with Dr. Susan Hodge at the NY Psychiatric Institute. For the past six years Georgia has supported the financial operations of the department and coordinates our facility operations. She has steadily grown in her finance role and is a real stickler about enforcing our internal procedures, which helps the Department to maintain compliance with CU policies. Besides managing the day to day headaches associated with the facility operations, which can be anything from getting a light-bulb changed to dealing with the after-math of a flood, Georgia also manages our ongoing “design on a dime” projects. These are projects that the Department identifies from time to time in order to brighten up some of its common use areas. Georgia brings a lot of energy to these projects, has great ideas, keeps us on budget, and takes great pride in ensuring that the Department looks good. The Department is grateful to have a staff member as dedicated as Georgia and truly values her contributions. Her positive attitude and willingness to be a team-player are essential to its administrative operations.

In her spare time, Georgia loves to be with her friends and family. You can catch her going to Broadway shows with friends to visiting family in Europe. She is a great cook and often treats co-workers to her delicious meals. She really enjoys interior design and loves to decorate and redecorate. Check out her purple-themed office when you get a chance!

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

Have the courage to drive – I know how but the fear that settles within when I think about getting behind the wheel and traveling on my own physically scares me.

When you have 30 minutes of free-time, how do you pass the time?

When I have 30 minutes of free time I like to redecorate. I have a passion for interior design and have always dabbled a little into it. I also love to read so that helps me to relax.

What was the last experience that made you a stronger person?

The last experience that me a stronger person was when I lost my oldest daughter in a fatal car accident. You think you are strong and can handle anything that is thrown at you, but that was one of the most difficult times of my life. I had to think about others and not grieve like I should have. My strength comes from my kids and my grandkids.

What’s your favorite indoor/outdoor activity?

My favorite indoor activity is sewing; my favorite outdoor activity is dancing/walking.

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

I guess I would say my grandmother who was a second mom to me.

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

Yes, leaving my job after 12 years and wondering what I would do next.

What three traits define you?

My ability to see good in others, Kindness, Giving.
Ten years ago, who did you think you would be now?

An interior designer. That is my passion but I love to do it without pressure from others.

What would you do (for a career) if you weren’t doing this?

I would have been a dental hygienist. My background of study was science.

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

My favorite thing about working in the bio department is the many different people who I have encountered. It’s always a pleasure when you’re able to connect with people. That has always been something I strive to do on a daily basis.

RICHARD BUCHSBAUM

Richard Buchsbaum has been creating databases, and the software to interact with them, for more than 20 years. His first experiences in this field, in business, especially in advertising, provided him with skills in both databases and user interfaces (it is unusual to be proficient in both), and a focus on the user experience.

Since coming to Columbia University in 1999, Richard has created data systems for patient registries, laboratory management systems, epidemiological studies, administrative units and clinical trials. He has extensive experience in collaborative research networks, including the North American Mitochondrial Disease Network and the Breast Cancer Cooperative Family Registries. He collaborates with clinicians and biostatisticians, and supervises the work of programmers, coordinators, and data quality staff. Richard is conversant with bioinformatics, relational database theory, user interface design and a variety of information technologies. He has developed an approach to research data management that integrates multiple technologies (websites, desktop applications, telephone systems, wireless connectivity, centralized data storage, etc.) to address the full spectrum of research data needs.

Richard is a unique resource. His impact has been pervasive throughout the Mailman School, the Columbia University Medical Center, and the greater Columbia community. More than fifty projects currently capture data using systems he and his team have created, touching faculty, staff, students and patients alike. In the era of data science, Richard is a major asset to the Columbia community.
JUSTINE HERRERA

The Biostatistics Enrichment Summer Training program celebrated their 10-year anniversary. BEST has accomplished a lot in introducing minority students to the field of Biostatistics and Public Health. Justine Herrera, the Program Director, plays a pivotal role at BEST, supporting a diverse public health workforce.

What is your role in the BEST Diversity Program? What experiences in your life led you to this work?

Before 2009, I was at the Morningside campus coordinating a very similar summer research pipeline program for undergraduate students. Back in my hometown of Virginia Beach I helped run Learning Bridge, a non-profit program that provided educational support for underserved populations. One of the major activities of Learning Bridge was providing a summer enrichment program where students took English, Math and Sciences courses.

My first job at Columbia was coordinating the educational outreach programs for an interdisciplinary center at the Morningside campus. One of my responsibilities was running Research Experience for Undergraduates (REU) programs in Materials Science and Nanoscience. Undergraduates from across the country applied to spend eight weeks at Columbia doing research with a faculty member in the Physics, Engineering or Chemistry departments.

In 2008, I came to the Biostatistics Department, which just finished its pilot summer of the BEST Diversity program. Emma Benn and Gary Yu, the two DrPH students who co-founded the program, were handling not only the teaching but also the administrative aspect of running the program. Since I had experience running summer programs, I asked if I could help them with the administrative side of BEST and they were happy to have me take over that piece. I then took on the role of overseeing BEST after Emma and Gary graduated from the department.

I am now responsible for recruiting students to teach the summer courses. Every year I play a role in matching the undergraduate students with a faculty mentor who will expose them to a research project with real-world implications. I think we are achieving the BEST goal of creating a pipeline program to introduce students to the field of biostatistics and public health. Out of the 114 students who have gone through the BEST program, we have had 13 of them pursue a degree at Mailman.

How is BEST different from your previous pipeline program on the Morningside campus? What keeps you here?

What’s different about BEST is that it really focuses on students from underrepresented populations in the STEM field (i.e. minority students, students from low socioeconomic background, first generation students with a disability) providing them with an opportunity to explore a field that they may not be familiar with. Also, I feel more of a connection with BEST students. Maybe it’s because I was a first-generation college student and I wish I had a BEST-like program when I was in undergrad.

The students keep me here. BEST is a small program, 12-14 students each summer, so it feels like a family. I get to know each student while they’re in the program and I stay connected with many of them after the program. Every summer I get to work with a new group of diverse students. We have had many BEST alumni pursue degrees in our department and elsewhere. I enjoy seeing students learning more about what graduate school is like and figuring out if this is a fit for them. Regardless of where BEST alumni end up, they will always be a member of the Biostats family.

How can our students learn more about BEST?

Biostats students are all invited to teach our courses to the undergraduates over the summer. I encourage all students to attend the BEST poster session and see the great work our BEST students are doing. I also recently noticed an increase in the number of students asking me to look at their resume and personal statement before they apply for Mailman or other graduate schools. It would be great for these students to have another set of eyes on their application materials. I think our BEST alumni would benefit from receiving feedback from current Mailman students of color.
Can you tell us about the various biostatistics roles that you’ve had since you graduated from the department?

Since graduating from the Columbia University’s Biostatistics program at the School of Public Health, I have had four major roles. First, I was immediately employed in June 1998 by Lenox Hill Hospital as Health Information Consultant working for one of the two Vice Presidents of Operations. In this role, I used my Business Administration background in conjunction with my Biostatistics training to prepare various analytic reports for Administration. In addition to supporting hospital administration, including the Chief Operating Officer and the President of the hospital, in my second year, I held the responsibility of providing biostatistical support to all the Gastroenterology fellows. In addition to this, I also supported the Cardiology Department reviewing the articles they selected for Journal Club to give input on any strengths and weaknesses of the research design. I had some personal matters to take care of so I then started to work for myself part-time and created my own corporation, named Spectrum Stats. My main client in this role was Wyckoff Heights Medical Center where I provided statistical support to the Preventive Medicine department. This role opened the door to a teaching opportunity to residents at the hospital, which was a great experience for me. Since I had so much exposure to working with physicians at Lenox Hill, I was able to design a course that addressed the statistical deficits clinicians had performing their own analysis (such as the assumptions of statistical tests). I went back to full-time employment at a public health department in Westchester County, New York. While there, I continued to consult for Spectrum Stats. At the County, my first major assignment was to go through about ten years of Early Intervention data including financial data. Another assignment later on was to determine the weight status of the school children of Westchester County which was based on CDC weight for age parameters needed to calculate Z-scores. The analytic results ended up on the front page of the Westchester paper, The Journal News. I worked on Community Health Assessments, Healthy People 2010 statistics, the New York State Preventive Agenda and even served as a Dose Assessor to quickly calculate during drills how many rems of exposure of radiation persons living around Indian Point Power Plant would be subject to given a radiation leak partly depending on wind speed and proximity to the plant to decide whether the residents should shelter in place or evacuate. One time, I received a call on a Sunday night which was not a drill. Thankfully, the event did not result in a leak. Finally, I decided to return to the hospital environment as I enjoyed my time at Lenox Hill. Northwell Health now owns Lenox Hill so I have come full circle and am analyzing their data and data on an additional 17 hospitals. Since starting in 2012, I have helped to create over 100 quality indicators and process most measures being monitored by the Affordable Care Act Programs (Value Based Purchasing, Readmissions Reduction Program, and Hospital Acquired Conditions) for the health system. In addition to the quality measures, I have done extensive research in sepsis and have presented at APHA Conferences on (1) analysis on mortality and length of stay as end points to the introduction of the use of a sepsis bundle and (2) lactate clearance for death prediction in severe sepsis and septic shock.

You were a recent participant in the department’s Master’s Alumni Panel, how did it feel reconnecting with the department?

I truly felt honored coming back to my alma mater to speak to the students. There is a certain energy in the Biostatistics department, with talented enrollees
and skilled professors. In a way it felt like coming back home. There were more students than I expected present in the auditorium eager to hear what we could share with them. When you are a student, about to graduate, there is so much expectation about the future but also many questions. This panel served to help answer those questions and share knowledge from our past experience. At the end of the discussion, I was truly glad I came back to help give insight into how my job search was exactly 20 years ago and what my experience after Columbia has been. I appreciated the time spent with the students and my fellow alumni and know that the Class of 2018, the Class of 2019 and many more to come will shine brightly in various careers.

How has the department changed since you were a student?

The department has changed immensely. I actually came back to the campus for the first time in years just a few months before being invited on the Alumni Panel to attend the Levin Lecture Series as I was feeling nostalgic in 2017. Back in 1996-1998 when I was a student, the Biostatistics department was located on the 18th floor of the old Presbyterian Hospital building and shared that floor with the Epidemiology Department. I may surmise that it may have been more intimate then. I recall walking down the hall and knowing where each professor's office was, including Dr. Paul Meier, Dr. Bruce Levin or my Introduction to Biostatistics professor, Dr. Heitjan. To give a better idea of the contrast in size, I can recall that my graduating class including doctoral students was less than 10! Everyone seemed to be in the Epi program and no one dared to take biostats back in my day. Now the Biostatistics department is stronger with professors focusing in diverse areas, it is well structured, and has a broader course selection. The poster presentation requirement is a plus which we did not have when I was a student. All of these additions will only make the graduate more well-rounded, prepared and competitive with biostatistics graduates from other programs.

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

Communication skills are paramount in being a successful biostatistician. When you deal with clients, you need to ask appropriate questions to seek a complete understanding of the needs of the client. Part of being a good communicator is being a good listener, so in order for you to answer the question at hand, you must obtain the information with a complete understanding from the beginning of the project to develop a statistical plan—or to design your experiment. Also, it is necessary for you to be able to convey clearly, completely, and accurately statistical information to persons who may not have an understanding of statistics. Always summarize your meetings after you meet with your clients and follow up with an email to guarantee that both parties are on the same page and include a timeline for next steps. It may be that some requests are not possible and this has to be expressed to the client giving them the reasons why.

In my present position, presentation skills are necessary to relay the results of analysis to an audience of mainly non-statisticians. You should not speak too softly, too slowly or too quickly, and you do not want to speak in a monotone. You do want to be clear, articulate, make good eye contact while maintaining a good posture. I like to ask questions when I am presenting to get a feel of the knowledge base of the audience. Do not assume they know the statistical jargon you may be speaking about. Once you understand your audience, you will
be able to explain your data at the right level to them. It is best to be prepared for your presentation by practicing in advance. This will help you to deal with time constraints and organizing your thoughts.

Organizational skills in the world of data analysis is key. You often have to go back to prior analysis or you may be running 3-4 types of analysis at the same time. It is important to categorize your data in folders in a way that you can find your files when required. Also, when you write code it is really important for you to write comments in your code, because a year or two later, you may not remember why the data was analyzed a particular way. Writing down clear comments as you write your code will aid you and anyone who may want to review your analysis as many journal are now open and transparent for replication.

Finally, time management skills are essential to possess as you will have to meet deadlines and most likely will be juggling more than one project at a time. Being able to work within time constraints may mean putting in extra hours some weeks. I always try to get my smaller projects done first so I can spend more time on the larger ones without being distracted which is the way I tend to be more productive.

What advice would you give to current students who are looking to work after graduation?

The one thing I would advise all current students to do is to augment their theoretical studies with as much practical application as possible by learning how to perform analysis in both R and SAS as they do not want to limit themselves when they interview by only knowing one language or the other. I highly recommend attending a Joint Statistical Meeting before you graduate or the year you graduate. Publishers have booths set up with biostatistics books and you can browse through programming books and pick up ones that appear most helpful. You will also meet recruiters at these meetings so you can give them your resumes and they will connect you with opportunities. In addition to this, every student should work with the Office of Career Services. They offer invaluable timely information and connect you with various resources for the present job market. It goes without saying that everyone should develop a LinkedIn Profile now detailing your work experience along with your educational background. That is why I even would suggest doing an extra practicum before graduation. You will most likely perform different types of analysis in each internship which would broaden your practical experience. Last but not least, it is imperative to work on your interviewing skills. You should always do research on the company that is interviewing you, go prepared with questions about the position and the company, and even brush up on your programming skills as some companies may even request you demonstrate that you can indeed program in the languages you have listed on your resume.
Publications


Significant Moments


Riss, P. T., Goldsmith, J., Shang, H. L., and Ogden, R. T. (2017). Methods for scalar-on-


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

Support Biostatistics at the Mailman School

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