Welcome to our 2022 Summer Newsletter

Meet Our Team: Jim Wan

Interviewed by Trish Goedecke

Background and Research

I earned a Bachelor's degree in Mathematics at the Chinese University of Hong Kong, and a doctoral degree in Statistics from Yale University. My research interests include the analysis of failure time data and the use of Poisson

regression and logistic regression in epidemiologic studies. I collaborate with faculty across the whole campus on statistical methods for clinical and epidemiologic data and health services research. These collaborations have resulted in more than 130 peered-reviewed articles and close to 200 abstracts presented in major national and international scientific conferences. The collaboration I have enjoyed most was with Dr. Jim Bailey around the turn of the millennium. Other collaborators have included Dr. Ithaar Derweesh, Dr. Joe Santoso and Dr. Jay Pershad. They have the common characteristics that they respect my contribution, and are able to explain their research objectives in a clear and precise way. My motto is *simple is beautiful*. If you can explain in ten words, don't use eleven.

Pastimes

I enjoy hiking. Before the pandemic, I went back to my hometown of Hong Kong 17 years in a row. Each and every time, I took a third of the days to hit the trails. This opportunity was thanks to our British governors, especially Murray MacLehose during the time I grew up in the '70s. A trail system named after him was my favorite. In fact, many Japanese tourists came to Hong Kong mainly for the trail system. Since Hong Kong is hilly and small, one can access a trail in no time. It is not difficult to find a trail that leads to a cliff, overlooking the outlying islands on one side and the high-rise buildings on the other. The night scenery there is unforgettable.

Back in the States, I have lived in four places. The one I miss most is Oak Mountain State Park in Shelby County, Alabama. It also has a trail system. I remember one New Year's Day, I came across a couple of students of mine there. After they learned that I hiked there, they were inspired and followed my steps. In the Memphis area, we do have a few trails. The closest is in Shelby Farms Park. Unfortunately, it does not give me a rural feel. Meeman-Shelby Forest State Park is rural, but it has too many mosquitos and flies. I enjoy hiking in Chickasaw State Park, which is about an hour away.

Besides hiking, especially in recent years, I pay more attention to commentary. I have learned many famous, sometimes funny quotes.

Here are a few of my favorite quotes:
The only thing that we learn from history is that we learn nothing from history — Georg Wilhelm Friedrich Hegel.

Success is not final, failure is not fatal: it is the courage to continue that counts

— Winston Churchill.

We know they are lying, they know they are lying, they know we know they are lying, we know they know we know they are lying, but they are still lying
— Aleksandr Isayevich Solzhenitsyn.

The butt directs the brain
— Cao Xueqin in The Dream of Red Chamber

Bad money drives out good
— Gresham's Law.

MacLehose Trail

GeneNetwork Funding Continues for Saunak Sen, PhD

Written by Hyo Young Choi

Funding continues for Saunak Sen's GeneNetwork grant through a four-year renewal of $453,683 from the National Institute of General Medical Sciences. The GeneNetwork grant, "A unified high performance web service for systems genetics and precision medicine," was initiated from 2018 and is being extended through 2021 to 2025 after an initial 4-year funding. The grant will fund building the next generation of the GeneNetwork platform for systems genetics and precision medicine, which will provide fast interfaces to genomic data and interactive tools for massive multivariate datasets.

Dr. Saunak Sen, professor and chief of the Department of Preventive Medicine at the University of Tennessee Health Science Center (UTHSC), is the contact PI on the grant and oversees the statistical and computational aspects of the research project. Dr. Rob Williams, department chair and professor of the Department of Genetics, Genomics and Informatics (GGI), and Dr. Pjotr Prins, assistant professor of GGI are other PIs on the grant.

For more than 20 years, Dr. Sen has dedicated his career to the development of statistical methods for understanding biological systems. This includes developing statistical approaches for matrix-valued high-throughput data, computational methods for large-scale linear mixed models, and statistical computing using the Julia programming language. His commitment to advances in statistical methods for complex and large-scale biological data has resulted in publication of a number of peer-reviewed manuscripts and statistical software packages including six Julia packages published just in the past year.

"The goal of the GeneNetwork project is to provide a platform of innovation to make the best use of the data and computational tools for systems genetics," Dr. Sen said. High-throughput technologies have transformed biology by enabling deep phenotyping of living organisms. Deep phenotyping in genetically randomized populations provides a powerful resource for understanding cause and effect. However according to Dr. Sen, despite the advances in data collection and the availability of genetic and genomic resources, the generation of actionable knowledge from these data presents computational and statistical challenges.

To overcome this, his research will focus on broadening the impact of the web service by encouraging community participation and strengthening engine

powering of the service. In this funding period, the team will enhance the database by providing automated tools for data entry and curation; improve tools for genetic analysis of high-throughput traits; enable bidirectional information sharing between animal model and human genetic studies; and provide a platform for toolset prototyping and testing by the computational community.

According to Dr. Sen, "A major piece of this project is to develop statistical models and computational tools for structured matrix-valued data obtained from high-throughput technologies and develop web 'apps' using the GeneNetwork database. This will serve as an interactive tool for biologists to interrogate hypotheses using diverse datasets in the database and translate findings across species."

COVID-19 Efforts

Chi-Yang Chiu, PhD:
Dr. Chiu and his collaborators studied a mathematical model for accurately describing the transmission and spread of pandemic. They proposed a susceptible-exposed-infected-recovered-decreased (SEIRD) system of differential equations that illustrates relationships among the number of susceptible individuals, the number of exposed individuals, the number of infected individuals, the number of recovered individuals, and the number of deaths in a town, state or country. While the existing models consider that only an infected individual is infectious, the proposed SEIRD system assumes that a susceptible individual may interact with an exposed person and contract the virus. Their empirical results show that a very rapid decline in contact significantly reduces the numbers of exposed, infected and deceased. These findings clearly support that wearing masks and reducing social activities can be helpful during the pandemic.

Fridtjof Thomas, PhD
Dr. Thomas supports the City of Memphis and Shelby County COVID-19 Joint Task Force: Emergency Support Function #8 – Public Health and Medical Services and has since April 2020 contributed COVID-19 related summaries and analyses from daily new cases to predicted bed occupancy in the hospital

system incl. required ICU resources in the Memphis Metropolitan Area. These updates were delivered daily starting April 2020 and are currently (Oct 2021) provided three to four times per week. Starting May 12, 2020 to the present, he is contributing figures and analyses to the UTHSC Local Data Projections web site - Updated COVID-19 information for the Memphis Metropolitan Area: confirmed cases in Shelby County with medium-term predictions, Harvard Global Health Institute Metrics for COVID suppression for counties in the Memphis Metropolitan Area, hospital bed occupancy with medium-term prediction, choropleth maps with SARS-CoV-2 PCR test results; https://uthsc.edu/coronavirus/local-data/index.php (1,218 figures archived as of 10/15/2021). His figures and analyses have been used by the Shelby County Health Department, as well as the office of the Mayor (City of Memphis Mayor Jim Strickland) to motivate COVID-19 related measures in the city of Memphis, and by local print and TV media (Commercial Appeal, Daily Memphian, WMC Action News 5, Fox13).

His work in April 2020 regarding the timing and intensity of the "epi-curve" in the Memphis Metropolitan Area and the U.S. in general is published in the American Journal of Epidemiology (Olney et al. 2021). Dr Thomas has been taking the lead for compiling the COVID-testing results for the UT community testing site ("Tiger lane"/University Clinical Health) as well as our community partners Christ Community Health Services and Church Health (project "COVID-19 Observational Registry" - PI Dr. Schwartz – UTHSC IRB 20-07287-XM).   Dr Thomas is the biostatistician on the NIH-funded Tennessee Community Engagement Alliance COVID-19 Disparities project with UTHSC site PIs Drs. Bailey and Stewart (PI Dr Juarez/Meharry Medical College; part of the nationwide NIH Community Engagement Alliance (CEAL) Against COVID-19 Disparities initiative).

Dr. Thomas has authored a white paper (together with Dr. Manoj Jain) in December 2020 that projected a doubling of the COVID-19 deaths in Shelby County from Dec 18, 2020 to Feb 18, 2021 if the then ongoing rampant community spread would not be brought under control. Unfortunately, that projection largely materialized despite a following Safer-at-Home directive by the Shelby County Health Department (in effect Dec 26, 2020 – Jan 23, 2021) and the deaths attributed to COVID-19 rose from 783 to 1,452 in Shelby County (+ 85% until Feb 18, 2021; total deaths as of 10/20/21 is 2,210).

Hyo Young Choi, PhD:
Dr. Choi examined complexities of COVID-19 comorbidities to understand their

contribution to illness severity or death. She and her collaborator, Dr. Shin, explored all COVID-19 confirmed and hospitalized patients in South Korea for the study patients. They conducted comorbidity network and multinomial logistic regression analyses to identify risk factors associated with illness severity or death. They showed that underlying diseases have different contributions to the severity of COVID-19 disease. In particular, chronic kidney disease, chronic obstructive pulmonary disease, heart failure, and malignancy were significantly associated with case severity or death. These findings show that underlying comorbidities should be closely monitored to efficiently allocate limited medical resources.

Mehmet Kocak, PhD

Associate Professor Mehmet Kocak, PhD, has served as an unofficial adviser to the Turkish Ministry of Health. His work focused on forecasting and benchmarking the impact of the pandemic, with a focus on peer countries.

Interactive Mapping - Tristan Hayes MSc

The staff of the Division of Biostatistics, led by Tristan Hayes, helped create novel web-based tools to map the virus' impact, both at the state and national level. These maps were developed using R Shiny and Javascript and allowed anyone, anywhere to quickly zoom in on a particular area in the US and move forward and backwards in time to track the development of the pandemic. This tool works on any device from a lightweight smartphone to a desktop computer and scales instantly. Since developing these tools, staff and faculty have given a well-attended virtual presentations, both on Covid-19 data options in general and also on R Shiny. See below for a TN example.

References:
- Olney AM, Smith J, Sen S, Thomas F, Unwin HJT. Estimating the Effect of Social Distancing Interventions on COVID-19 in the United States. Am J Epidemiol. 2021. Epub 2021/01/07. doi: 10.1093/aje/kwaa293. PubMed PMID: 33406533; PMCID: PMC7929448.
- Jain, M., & Thomas, F. (2020). Projected Covid-19 Deaths in Shelby County From Dec 18, 2020 to Feb 18, 2021. White Paper.
- Shin EK*, Choi HY*, Hayes N. The anatomy of COVID-19 comorbidity networks among hospitalized Korean patients. Epidemiol Health. 2021;43:e2021035. doi: 10.4178/epih.e2021035. Epub 2021 May 7. PMID: 33971700; PMCID: PMC8289479. (*co-first authors)
- Shuqi Wang, Wen Tang, Liyan Xiong, Mengyu Fang, Bingsong Zhang, Chi-

Yang Chiu, Ruzong Fan. Mathematical modeling of transmission dynamics of COVID-19[J]. Big Data and Information Analytics, 2021, 6: 12-25. doi: 10.3934/bdia.2021002

Summer 2021 Internships

Our division recruits summer interns every year. The selected interns work with faculty in our division for about 3 months. The program was started in 2016, and 9 interns have completed the internship so far. To date, all have gone on to pursue careers in data science. The class of 2021 with 4 interns was the largest to date. This article briefly introduces the projects of last year's interns. Congratulation on their great work!

Summaries provided by Mentors

Nadeesha Thewarapperuma (mentors Chi-Yang Chiu and Xiaoyu Liang) is a PhD student at the University of Kansas Medical Center. The aims of her project were to explore the DNA methylation difference between participants with and without oral and pharyngeal carcinoma and to investigate the role of DNA methylation in the mechanism of oral and pharyngeal carcinoma development. She performed an epigenome-wide association study for identifying oral and pharyngeal carcinoma-related CpGs by utilizing functional data analysis techniques.

Winston Miller (mentor Fridtjof Thomas) is a graduate student at the University of Memphis whose internship was mentored by Fridtjof Thomas. He developed an R package, "bigBERD," to generate a biostatistics report template. There are existing packages which will generate somewhat automated reports, but the advantage of this project is that it will generate a template which can be easily modified. We expect bigBERD to save investigators significant time.

Ye Eun Bae (mentor Saunak Sen) is a graduate student at Florida State University whose internship was co-mentored by Tom Juenger of University of Texas at Austin and Saunak Sen of UTHSC. Bae developed a statistical method for genetic analysis of allotetraploid species using data from switchgrass, a grass species native to North America used as biofuel. Her approach will shed light on the adaptation of switchgrass varieties to upland and lowland environments.

Zifan Yu (mentor Saunak Sen) is a graduate student with the University of Washington, whose internship was mentored by Gregory Farage. Yu developed a fast computational algorithm in the Julia programming language for modeling large data matrices common in high-throughput biological experiments. His approach models matrix data with covariates on rows and columns and allows for an elastic net penalty for sparse estimation (ie, when most effects are expected to be zero or null).

Find Out More