ABSTRACT:
The novel virus SARS-CoV-2 has produced a global pandemic, forcing doctors and policymakers to “fly blind,” trying to deal with a virus and disease they knew virtually nothing about. Sorting through the information in real time has been a daunting process—processing data, media reports, commentaries, and research articles. In the USA this is exacerbated by an ideologically divided society that has difficulty with mutual trust, or even agreement on common facts. The skills underlying statistical data scientists are central to this knowledge discovery process, filtering out biases, aggregating disparate data sources together, dealing with measurement error and missing data, identifying key insights while quantifying the uncertainty in these insights, and then communicating the results in an accessible balanced way. As a result, we have had a central role to play in society to bring our perspective and expertise to bear on the pandemic to help ensure knowledge is efficiently discovered and put into practice. I have authored a website and blog (https://covid-datascience.com) that represents my own personal efforts to disseminate information I have found reliable and insightful regarding the pandemic, accounting for subtle scientific and data analytical issues and uncertainties about our current knowledge, and seeking to filter out political and other subjective biases.

Using experiences with the covid-datascience blog as a backdrop, I will highlight how statistical and data scientific issues have been central in understanding the emerging knowledge in the pandemic, and will make a general case for the importance of our skill set in solving societal problems, and arguing we should position ourselves to have a seat at the table with policymakers and have visibility with the media in scientific communication.

BIO: Jeffrey S. Morris received his PhD in Statistics from Texas A&M University under the supervision of Raymond J. Carroll and Naisyin Wang. He is also a recipient of the H.O. Hartley Award from the Department of Statistics. His research interests focus on developing quantitative methods to extract knowledge from biomedical big data, including work to relate complex biomedical object data—including functions, images and manifolds—to patient outcomes and characteristics using flexible, automated regression methods, and to integrate information across multiple types of multi-platform genomic, proteomic, imaging, and wearable device data to uncover biomedical insights contained in these complex data.