TRANSCRIPT: From Traditional Statistics to Big Data

KERRY LEE: The work of the statisticians is not all just analyzing data. We also are very heavily involved in designing these studies.

ZHEN HUANG: The Biostats group, in particular the one I'm involved with is a Clinical Trials Statistics group, we play a multi-faceted role in all stages of our clinical research. And in particularly, our focus is to provide the statistical expertise in study execution. So no matter how great a study design is, at the end of the day are were able to execute them in such a way that we're collecting quality, reproducible results that conclusively answers the research questions at hand.

LAINE THOMAS: Most Health Services Research projects are heavily dependent on biostatistics, and biostatisticians are constantly learning novel methods in order to deal with the problems of more complex data. So one of the trends that's changing is the data sources we are using are larger data sources from more heterogeneous sources, mixing platforms that patients enter their own data on—or patient reported outcomes—so a complex array of data sources coming together. And our job is to pull that together, make it clean, and then answer questions about health and improving outcomes.

ZHEN HUANG: I think we have had multiple initiatives of our groups that are motivated by the problems that we have encountered. Either that problem is triggered by a new study, a more complex design, or triggered by some existing historical limitations.

Whenever there is a new study design, there is calls for a new way or an innovative way to run the study and for us to collect the data and also to analyze them.

KERRY LEE: I had the great privilege to be in on the ground floor and the pioneering development of the Duke—what is now the Duke Clinical Research Institute. And it has been so much fun over the years, and so exciting to work very closely with the clinicians.

I can honestly say that we've been able to make a mark, a useful contribution, not only to the clinical issues, but the methods, the approaches for analyzing and modeling this information has also had an impact in the statistical world.

We've published manuscripts both in the clinical literature, as well as manuscripts in the statistical literature, that have outlined some of the methods that have been developed over the course of the years. One of the papers that was published by Frank Harrell and myself and Dr. Dan Mark has been one of the most highly referenced articles with regard to statistical modeling.

LAINE THOMAS: The new data sources are allowing us to answer new types of questions that are more complex—questions that we've always had, but they weren't always feasible in smaller data sources or with less information. And so new problems arise that we have to address. So, a recent example would be trying to study the personalized or patient-specific medicine—personalized medicine—in women with uterine fibroids. There's a study that we're running in Health Services Research called COMPARE-UF that wants to understand what treatment works best for what type of woman.

And so it's much more specific in the question, but the statistical challenges turned out to be unique. So one of the new methods that we've developed—we have a recent paper coming out showing how to do

subgroup analysis in observational data to look at what treatment works best for what type of person. And that was really exciting, I think, reflects the beauty of biostatistics at DCRI. All of the different team members working together came up with an array of innovations. And I think that's what, you know, is unique at DCRI—is the way the team comes together to innovate.

ZHEN HUANG: We are required to be nimble and to learn. And by learning from the protocol, learning from the experts who designed the protocol and who designed the statistical analysis plan, I think it's a very good opportunity, I think, that is afforded to us at DCRI. So by having the exposure to those new methods and you, you just grow from them.

KERRY LEE: And that's what makes working at the DCRI so much fun and so rewarding because of the fact that, as we design these studies and conduct the studies and answer the questions, it makes a difference in the lives of real patients.