Research interests and experience

My best research efforts have been collaborative. I want to highlight one that illustrates cutting-edge methodology and computational tools for solving high-dimensional problem. I also will describe my efforts to help others be successful in their research endeavors.

Mining the electronic health record

In January 2016, I was offered the opportunity to work on a research grant funded by the Patient Centered Outcomes Research Institute. My assignment was to develop a develop a phenotype of breast cancer from information in the electronic health record (EHR) and validate it against information in the breast cancer registry. The advances in high throughput genome sequencing and the linkage of that information with the EHR allows for exploration of novel precision medicine options. Developing a phenotype from the EHR, however, is fraught with peril because information in the EHR on basic issues like diagnoses and treatments is often coded inconsistently.

I used a big data model, LASSO regression, to predict whether a patient was in the breast cancer tumor registry and set up sparse matrices as input to better manage the size of the data sets. The breast cancer cases were compared against three separate control groups and in spite of the massive size of the independent variable matrix (more than 45,000 columns), this model ran in under ten minutes. The resulting sensitivity and specificity were very high, putting to rest concerns that the EHR data might be too incomplete or inconsistent to produce an accurate phenotype. The LASSO regression model could easily be run for other tumor types, and just as quickly validated. I have presented these results at a local research conference and plan to submit a peer-reviewed publication soon.

My work on the PCORI grant has been transferred a different grant and I plan to work with partners at Truman Health Center and Children's Mercy Hospitals to develop more research utilization of EHR at their locations. I also have been asked to help develop an analytics platform simplifies data mining of the EHR through a standardized library of functions interfacing SQL databases and R. This library would pull appropriate meta data descriptors as well, expanding the types of analyses available to the end user.

Helping others with research

I am a great believer in the Harry Truman quote "Anything is possible if you don't care who gets the credit." A major portion of my career has been helping others become successful researchers. This work is quiet, behind the scenes, and often leads to very little recognition for me. But I enjoy watching someone developed from a scared and timid person starting out with their very first research study to someone who has learned enough that now he/she is mentoring others.

A large part of my work is helping people who are struggling in their graduate studies. It might be some extra tutoring for that seemingly impossible statistics class. More often, though, it is guiding students through the difficult process of writing and defending their dissertation. I did this for free for doctors, nurses, and other health care professionals that I worked with at Children's Mercy Hospital. After I left that job, I started my own consulting business, and I got lots of referrals to graduate students. They typically are struggling with their disserations and with a dissertation committee that was not giving adequate direction on the data analysis. For a dissertation, you can't do the data analysis for them because they have to be able to explain their work during the dissertation defense. You have to teach them those things that they didn't pick up in their earlier statistics class and teach them so thoroughly that they can offer a clear and convincing presentation of the analysis that they ran. You have to help them anticipate the types of questions that they might get and how best to answer those questions. Perhaps the most important thing is to get them a sense of self-confidence that they are working on an important problem and that they have a solid and defensible plan for solving that problem.