Distributed computations for collaborative medical projects

Personalized medical care in the fields of statistics relies on comparing new patients profiles to existing medical records, in order to predict patients response to treatments or risk of disease based on their individual characteristics. The chances of finding past profiles similar to new coming patients, and therefore of more accurately treating them, naturally increase with the number of individuals in the database. For this reason, gathering patients information from different hospitals promises better health care. However, there are technical and social barriers to sharing medical data. Indeed, as the size of the database increases, it becomes more and more difficult to handle and store it. Simultaneously, institutions are usually reluctant to share their data due to privacy concerns. During my stay at Stanford, I will be working with Dr. Narasimhan on a technique called distributed computation, which will allow us to overcome both obstacles. It consists in leaving the data on sites and doing calculations separately, so that hospitals only share some intermediate results instead of the raw data. We will implement statistical methods in a software which will be available to hospital practitioners and will help them carry out collaborative medical projects.