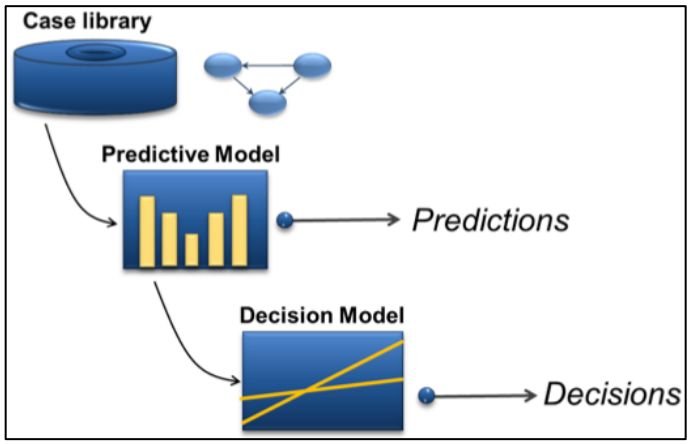
Data to Insights to Actions: Enabling Evidence-based Healthcare

Eric Horvitz
Microsoft Research



Nearly 2500 years ago, Hippocrates called for the recording of data about patients and their illnesses, establishing evidence-based healthcare as a long-term goal. Although 25 centuries have passed since Hippocrates' call, we have not yet attained the dream of true evidence-based healthcare. To date, large quantities of data about wellness and illness continue to be dropped on the floor, rather than collected and harnessed to optimize the provision of care. However, we now stand at the brink of a potential revolution in data-centric healthcare, made possible by advances in multiple areas of computer science. Critical funding of fundamental research in the computing and decision sciences over several decades by NSF, DARPA, NASA, the Office of Naval Research, various institutes of the NIH, and other Networking and Information Technology Research and Development (NITRD) Program agencies have provided methods for collecting, storing, and reasoning with health data.

Computational systems and methods stand ready to enhance the efficacy and quality of healthcare. Key computational ingredients for powering a revolution in evidence-based medicine include affordable large-scale computation and storage resources, connectivity among computing systems and devices, and the development of computational procedures for machine intelligence, including methods for learning and predicting, planning, and decision making. These core ingredients can be used to develop computational systems and services that provide assistance with such difficult problems as error-detection, triage and diagnosis, acute therapy, long-term disease management, prediction of outcomes, and the formulation of ideal healthcare policies that weigh the costs and benefits of different actions by providers and patients.

Machine learning procedures can transform biomedical and clinical data into *predictive models*. Such predictive models can be used to generate forecasts with well-characterized accuracies about the future, or diagnoses about states of a patient that we cannot know or inspect directly at low cost. Patient-specific forecasts or diagnoses can be harnessed within cost-benefit analyses to generate recommendations for actions in the world, and decisions about when it is best to collect more information about a situation before acting – considering the costs and time delays associated with gathering more information to enhance a decision. The pipeline of *data to predictions to actions* can be used to automate or provide decision support for such goals as minimizing errors, performing accurate triage and diagnosis, generating predictions about health outcomes, generating effective plans for chronic disease management, and formulating and evaluating healthcare policies on a large scale.

While numerous core computational problems have been solved through NITRD investments over the last two decades, a critical standing bottleneck is the lack of data, based both in inadequate capture and in difficult challenges with sharing clinical data for research and development. Efforts to promote the installation of electronic health record systems and the storage of healthcare data

promise to change the availability of data that can drive analyses, decision support, and policy. Successful collection and use of data will depend additionally on deeply embedding the use of such systems into the workflows of clinical medicine, and enhancing the efficiency with which data is captured. In addition to in-hospital data collection, sources of data with relevance to wellness and disease may become available via the growing ubiquity of capable smartphones, inexpensive sensors, and portable or embedded medical instruments that communicate with web services.

This talk presents research efforts that highlight the promise of machine learning and decision support for reducing the costs and enhancing the quality of healthcare – with examples spanning research projects focused on reducing hospital readmission, addressing hospital-based errors, and minimizing hospital-associated infection. The specific directions and results make the ideas concrete and underscore the opportunities ahead. A longer CCC whitepaper on this topic can be accessed here.