Photo credit Glenn Asakawa

Finding the Needle in the Haystack

Colorado leads the nation in developing Big Data for the advancement of health

WRITTEN BY JEANNE McADARA, PhD

IT IS ESTIMATED THAT A SINGLE HUMAN BEING HAS THE potential to generate more than 12 terabytes of personal data relevant to his or her health over the course of a lifetime. Multiplied by the billions of individuals living on the planet today, those data represent an incredibly vast, rich and virtually untapped resource with which to understand and address human health problems and to improve the human condition. The ability to combine, analyze and share structured and unstructured patient data could improve the speed and accuracy of clinical diagnoses, match prevention and treatment strategies to outcomes, identify patients or populations at risk for specific diseases, and suggest both individual and public interventions and policies to mitigate those risks for the greater good.

While efforts to harness this resource—often referred to as "Big Data"—are arguably still in their infancy, the ever-accelerating development of advanced computational and analytical methods is beginning to provide tools capable of bridging the transition from theory to real-world practice. In the U.S., national legislative and policy initiatives aimed at improving healthcare

Left: Somalogic molecular biologist, Kan Chantranuvatana, runs a quality check on SomaLogic reagents in one of the company's Boulder labs.

quality, shifting to evidence-based delivery of care, and controlling healthcare costs are further driving efforts to capture and utilize Big Data.

Through a flurry of public and private activity and initiatives, Colorado's governing bodies, along with its public and private bioscience and healthcare sectors, have signaled their collective intention to assume a forward-looking leadership position as the nation explores this new frontier.

Jonathan Mathieu, Ph.D., is vice president of research & compliance and chief economist at the non-profit, non-partisan Center for Improved Value in Health Care (CIVHC).

"I joined CIVHC in 2012, and have seen a lot of productive change in Colorado, with various public organizations and initiatives coming together to find uses for Big Data that help to lower costs and improve population health," says Mathieu. "There is a lot of collaboration in Colorado, and an independent, can-do spirit that fits well with the Governor's goal to make this both the healthiest state in the nation and an attractive place to locate business."

Per Mathieu, CIVHC's origins can be traced to recommendations made in 2008, by Colorado's Blue Ribbon Commission for Health Care Reform. Among its many recommendations, the panel urged the creation of an entity to coordinate healthcare reform efforts in the state, and establish Colorado's all-payer claims database (APCD) to aggregate, analyze and report on claims data collected from health insurance plans covering the majority of insured Colorado residents.

"Health reform leaders in Colorado saw what was happening and realized that, without an entity collecting claims data, we were never going to be able to understand variation in patterns of spending for medical services, utilization or prevalence of chronic disease," says Mathieu. "Everyone knew there was variation, but unless you understand where and what is driving that, you don't know to dig deeper and can't really do much from a policy perspective."

The blue-ribbon recommendations, supported by legislation passed in 2010, and regulations implemented in 2011, led to the creation of the Colorado APCD, which began collecting data in 2012. Separately, CIVHC was created within the Department of Health Care Policy and Financing and was subsequently spun out as an independent, non-profit organization to administer the APCD.

> We can turn clinical data into information that informs action, allowing us to design or change community-based intervention strategies, policy-based system changes, and environmental changes, and then monitor the progress...
>
> **- ARTHUR DAVIDSON, PROJECT DIRECTOR, CHORDS**

The Colorado APCD now contains claims from the 21 largest commercial payers, Medicaid and Medicare, and is actively supporting consumer decision-making, research, healthcare operations, payment, treatment and public health policy.

Mathieu acknowledges that, while powerful and useful, analysis of claims data alone has its limitations, especially when it comes to generating measures of quality and outcomes. "We can use claims data to understand the rate at which A1C and cholesterol levels are checked, for example" he says, "but that doesn't tell us what proportion of the population has good glucose control. To get to quality, we need to combine claims analysis with clinical data."

Enter organizations such as the Colorado Health Observation Regional Data Service (CHORDS), and the non-profit health information exchange (HIE), CORHIO. According to Arthur Davidson M.D., project director, CHORDS represents a collaborative effort between Denver-metro providers and convened by the Colorado Health Initiative (CHI), to standardize and normalize clinical data from multiple Electronic Health Record (EHR) formats into so-called virtual data warehouses (VDWs). Those data can be queried by health departments and other public entities to answer health questions about specific Colorado populations. CHORDS data are distinguished from and complementary to data generated by initiatives like the Colorado APCD in that they contain clinical information rather than claims.

"We are, for example, able to look at obesity as a condition rather than a diagnosis," says Davidson. "We can look in a timely way at tobacco use rates, or conditions like asthma, depression, cardiovascular disease and diabetes, all through the virtual data warehouse, and categorize them by region, sociodemographics, age, gender and race. We can visualize them at a granularity that goes down to the community level. We can thus turn clinical data into information that informs action, allowing us to design or change community-based intervention strategies, policy-based system changes, and environmental changes,

Developmental Chemist Brian Ream focuses on nucleic acid chemistry at SomaLogic in Boulder.

and then monitor the progress and outcomes of those interventions."

According to Davidson, there are already 12 large providers in the region that are part of the project, including Children's Hospital Colorado, Kaiser Permanente and Denver Health, and CHORDS is working on bringing additional non-profit and university hospitals into the mix as well.

EHRs came into widespread adoption and pervasive use only within the last decade, so sharing EHR data between providers and using it for multiple purposes presents special challenges. Daily, HIEs facilitate patient-information exchanges from system to system, dealing with massive amounts of data to fulfill that role.

"We move 375,000 clinical messages around the state every day, and our web portal manages 250,000 queries per week. It's a lot of data," Morgan Honea, CEO at CORHIO, says. "As of today, we have 61 hospitals connected to our HIE, and clinicians, their key staff and other stakeholders at more than 1,100 organizations—including outpatient centers, public health organizations, diagnostic facilities and health plans—access and use the HIE in some capacity."

While public health is proving to be a key area in the movement to develop Big Data in Colorado, a great deal of activity is ongoing in private companies, academic centers and advocacy organizations. Companies like SomaLogic and Biodesix (both in Boulder) continue to develop and commercialize product platforms that use advanced analytics and Big Data collected from proteomics and genomics studies to shape precision medicine in the areas of wellness and disease. Matrix Analytics has developed software-learning algorithms to analyze Big Data for real-time clinical decision support, starting with identification and management of incidentally-found pulmonary nodules. Clinical Trial Site Solutions and its subsidiary, TOMR2 Development Designs, has developed a customizable platform that works with existing EHR platforms to improve clinical workflows and track outcomes-based data at clinical research centers and sponsor-driven clinical trial sites in virtually any medical sub-specialty.

Based in Denver, RxRevu aims to use Big Data and artificial intelligence to improve the consistency, safety and efficiency of the drug prescribing process. Their prescription decision-support tool, RxCheck, pulls together data from a patient's clinical and claims histories, and benefits plans and formularies, along with drug costs and clinical protocols. The tool then uses those "smart data" to inform the provider about appropriate prescribing options in a digestible, useable, hierarchical format in real time, at the point of prescribing.

"We want to help improve patient care by improving prescribing consistency so that patients receive the most efficacious, safest and most cost-effective medications based on the best evidence," RxRevu's chief executive officer, Carm Huntress says. "An important, data-driven feature of RxCheck is that it has a performance-measurement tool that uses data to drive downstream learning."

Huntress says that RxRevu is currently in partnerships with private, public and academic Colorado-based organizations to implement RxRevu into real-world practice.

Meanwhile, a wide variety of departments within the major university systems in Colorado have research programs dedicated to developing new uses for and tools to harness Big Data. At CU Boulder's Computational Bioscience program, faculty and students are focused on computational approaches to health problems from the molecular level to the population level. Elsewhere on campus, programs at the BioFrontiers Institute and the Department of Computer Science are developing novel computational and statistical techniques for understanding data from complex biological and social systems, and for characterizing the multitude of noncoding RNAs in the transcriptome.

In 2013, CU Denver established a new Division of Biomedical Informatics and Personalized Medicine (BIPM) in the Department of Medicine on the Anschutz Campus. Now headed by Kathleen Barnes, Ph.D., the division brings together medical informaticists, bioinformaticists, biologists generating large "omics" datasets and other academicians whose goal is to translate research into personalized and precision medicine to the clinic. Current areas of focus include application of computational biology and bioinformatics to cancer genomics and epigenomics, gut microbiology, hypoxia-related health conditions and pulmonary diseases.

> The lessons of Big Data are showing that inborn molecular traits, socioeconomic circumstances, environmental conditions, behavioral choices and medical care quality and utilization...all play a critical role in shaping health at both the individual and population level.

The Department of Computer Science at Colorado State University (CSU) has $15 million invested in active research projects applying Big Data, artificial intelligence and bioinformatics to computational problems in protein and RNA function and interactions, alternative splicing, genomics and transcriptomics, genome sequencing and resequencing, and detection of transcription regulatory elements.

The enormous level of statewide activity dedicated to collecting, sharing, analyzing, governing and responding to health data is the power of our collaborative and entrepreneurial environment.

Ultimately, the lessons of Big Data are showing that inborn molecular traits, socioeconomic circumstances, environmental conditions, behavioral choices, and medical care quality and utilization—and the interplay between these domains—all play a critical role in shaping health at both the individual and population level.

"So much about what happens in life is dependent on social experiences and economic environment," Davidson says.

Colorado is poised to bridge these divides to solve public health problems and amplify the impact of precision medicine for the health and well-being of future generations.