The health care industry has undergone deep transformations in recent decades, from the consolidation of smaller hospitals and medical practices into large systems, to the digitization of patient records and nearly every other aspect of health care. Despite these significant changes, North Carolina technology entrepreneurs see the potential—and need—for even deeper disruption of the industry through more effective use of data. By applying insights and analytical techniques from other industries (such as finance and retail), we can help health care providers better understand their own work and improve health outcomes for individual patients and broader communities.

Large health care systems have made tremendous strides in building digitized knowledge about all aspects of their operations. Digital records now exist of many aspects of doctor-patient interactions and medical outcomes. From business transactions to operational details around building management and employee support, each operational division maintains detailed records in numerous spreadsheets and large databases.

We argue that these traditional ways of collecting, organizing, and analyzing data are ripe for disruption in three ways: data management (collection, integration, curation), population scale analytics, and automation of human processes.

We now explain each of these in depth.

Data Management

When collecting different types of data—patient, procedure, payer, etc.—most health care companies and systems place the data in distinct and often unconnected silos [1]. Different divisions often record the same data in different ways. These data silos make it extremely difficult to integrate data across an entire organization to provide a unified, clean, and curated dataset. These same challenges multiply exponentially when trying to integrate data across health care systems and providers. Yet, this integration is crucial when a doctor tries to get a complete picture of an individual patient’s experience in a health care system, or when a patient has an accident far from home and needs health care records from a home hospital.

Our company, Tresata, has completely automated this entire data management challenge by using complex machine-learning disciplines to produce usable, quality data at massive scale. Our data management software enables providers and payers alike to have access to all of their data at the level of the individual entity, whether it be a patient, doctor, or health care problem. This integrated data, which we refer to as a “data asset,” allows doctors, hospitals, or health care systems to access all of a single patient’s records and transactions, or all of a single doctor’s patients displayed geographically on a map, or the prevalence of different health care problems in a health care system over time. The data asset enables personalized health care outcomes for each and every patient, doctor, or problem, resulting in improved health, understanding, and service.

Population Scale Analytics

Data assets also create the second opportunity for disruption: population scale analytics, which utilizes entire populations instead of samples of those populations. Many organizations, including those in health care, use predictive analytics to explore which medical diagnoses lead to better health outcomes, or which organizational structures or incentive systems lead to greater employee engagement. These models often use samples (of diagnoses, employees, etc.) to draw generalizations about entire populations, even when comparing “like” populations. Thanks to advances in distributed computing (defined as “large-scale data processing on clusters of commodity computers” [2]), Tresata’s software can now create models that incorporate data from the entire population—of patients, employees, health care outcomes, etc.—
regardless of the amount of data or size of the population. Instead of having to use a sample to generate a summary statistic about an entire population with an associated level of confidence about the prediction’s accuracy, we can now calculate that statistic using the entire population. If the analysis uses an entire population of patients, it can become easier to identify “like” sub-populations within the overall population.

Using entire populations of data significantly enhances our ability to make better, more accurate predictions. Sometimes a sample includes only a small number of cases of a significant subset (such as men over 50 who exercise frequently, or Asian Americans with a history of heart disease); with such small numbers, it is not possible to make confident statistical predictions about the subsets. But with a population, we are much more likely to find large numbers of most significant subsets. We also know the actual health care outcomes for the subsets, instead of having to predict them. As another example, health care organizations often collect samples of data on how often medical devices break down. These sample-based predictions are limited to the particular devices and contexts from which the sample data are collected. In contrast, if we collect the population of data, ie, all the data over time from every device made by the manufacturer that is currently in use, our data come from a much wider range of use contexts, and we can use the past history of each device to predict its future performance.

Automation of Human Processes

The third opportunity for disruption is automation. Software developed at Tresata automates the heavy lifting and tedium of data collection, integration, linkage, and enrichment. As a result, domain experts (such as doctors, researchers, and administrators) can focus on drawing insights from the data to address the critical challenges they face. (To our knowledge, no other company combines all the components of Tresata’s software at the same scale.) In automating all of the complex processes that go into curating data and generating predictions, our software quickly arms and augments health care professionals to make better-informed medical decisions. Practitioners have more data at their disposal than ever before and can access the information they need with the touch of a button.

In summary, we believe that data management, population-scale analytics, and automation of human processes can significantly help the health care industry improve health outcomes for individuals and communities. These changes are certainly disruptive and difficult to implement. But at Tresata we have helped numerous companies in other industries make the leap to these new data strategies, and we have seen their positive impact. In light of our country’s tremendous and continuing health challenges, we must make the same leap in the health care industry by supporting health care technology innovation and entrepreneurship.

Abhishek Mehta
founder & CEO, Tresata, Charlotte, North Carolina.

Acknowledgments

Potential conflicts of interest. A.M. is founder and CEO of Tresata, a Charlotte-based automation company that develops predictive data management and analytics software for health care organizations and companies in other sectors.

References

1. Clarke JL, Bourn S, Skoufalos A, Beck EH, Castillo DJ. An innovative approach to health care delivery for patients with chronic conditions. Popul Health Manag. 2017;20(1):23-30.
2. Lin J, Dyer C. Data-Intensive Text Processing with MapReduce (Synthesis Lectures on Human Language Technologies). Wilieston, VT: Morgan & Claypool Publishers; 2010.