Big Data, big influence

Why our collective data is a social determinant of health.

BY EDWARD P. EHLINGER, M.D., M.S.P.H.

Who owns the Earth? By definition, it is a public good. It is the “commons,” the natural resource available to all, that helps sustain every one of us. Yet, for centuries, the Earth has been arbitrarily partitioned, with groups of people claiming ownership over vast areas and resources. We have further divided and subdivided it so that some relatively small groups and individuals now claim pieces of it as their own.

The Earth’s resources are often extracted, harvested, harnessed and used for the benefit of all—for food, shelter, clothing, security and quality of life. But sometimes those resources are used for private gain, rather than the public good. When that happens, some people benefit disproportionately.

Power and money are generally the levers used to gain control over our natural resources. They are also the levers used to gain control over our socioeconomic resources—the social, economic, educational, physical and health care systems that form the environment in which we are born, grow, live, work, age and receive care.1 We refer to these complex, integrated and overlapping socioeconomic systems as the social determinants of health, as they have a huge impact on our overall well-being. Although our socioeconomic resources can be used for the benefit of many, often they are not, and thus we have unfair and avoidable differences in health status among and within countries.2

**Big Data and health**

Big Data has emerged as a new and powerful force that is rapidly and dramatically transforming our world. The term refers to the massive and complex datasets now being generated, along with our ever-expanding capacity to analyze their content and extract and use information.3 Big Data is expanding and modifying our understanding of every aspect of our world and giving us the ability to educate, inform, influence, monitor, track, assess and direct people in ways that were once only imagined.

Big Data has a particularly profound impact on health. In fact, it might be one of the most important social determinants of health because of its overwhelming influence on every aspect of our lives.

The advent of Big Data has allowed us to rapidly decode human DNA, track and prevent disease, predict human behavior, monitor physiological systems, find treatments for cancer, foil terrorists, improve airline safety, build self-driving vehicles and personalize marketing efforts.4 Yet there is a downside. As more data are collected in multiple ways from multiple sources, our privacy is eroding. As anyone who uses a credit card, a cell phone or the Internet knows, personal information is already widely shared, often without one’s knowledge or consent.

The impact, both positive and negative, of Big Data is only going to grow, as the amount of data we generate grows. Google’s executive chairman Eric Schmidt said at the 2010 Techonomy conference: “From the dawn of civilization until 2003, humankind generated five exabytes of data. Now we produce five exabytes every two days … and the pace is accelerating.” Big Data will affect everyone on the planet.

With improved health outcomes as a goal, Big Data can be a powerfully positive force. It can be useful in determining the effectiveness of current treatments for all sorts of injuries and diseases. It can provide clues to the causes of infections, cancers and noncommunicable diseases, which can lead to new cures.5 It also can improve health care by linking electronic health records with vital records, disease surveillance systems, social media, and housing, environment, transportation, employment, finance and education data. Such integration will allow treatment of individuals to be informed by a better
understanding of the physical and social environment in which they live.

In the field of public and population health, the potential benefits of Big Data are innumerable. One example is crowd-sourcing, which may one day enable us to identify infectious disease outbreaks at an early stage, when interventions might mitigate their severity. Linking public health datasets with health records also could help identify issues that affect the health of community members such as environmental triggers for asthma or the density and location of tobacco advertising. It also could help in the development of public policies that promote healthy behaviors. Combined, these efforts could lead to more than $300 billion in health care savings, according to McKinsey and Co.  

If profit, rather than health, is the goal, Big Data could be used to alter clinical practice in ways that benefit one health care entity or one device manufacturer or pharmaceutical company over others. Similarly, it could allow health plans or insurance companies to target outreach efforts and services to populations based on economic rather than health-improvement goals. We are already seeing Big Data being used to encourage and support unhealthy behaviors such as the consumption of sugar-sweetened beverages and to market products such as menthol cigarettes to minority populations.

Private good or public commodity?

Because of its power, Big Data can influence the direction of our health care system and the health of our communities. Beyond that, it could influence the course of our entire society, depending on how it is used.

This raises multiple questions: Is Big Data part of the “commons”? Is it a public good or a private commodity? Who owns it? Who controls it? Who decides what to collect and analyze? Who decides how to use it? Who is responsible for evaluating its impact?

Obviously, these are rhetorical questions designed to stimulate an overdue conversation. If Big Data is considered part of the commons, broad community input and involvement will be needed to determine how it could be harnessed to help society more effectively address disparities and guide us toward health equity. If it becomes a private commodity, Big Data most likely will be controlled by those with power and money and used to promote their vested interests. This could lead to even greater inequities.

As we learn more about Big Data, we probably will reach an understanding that it can be both a public good and a private commodity. Like all the social determinants of health, it embodies aspects of both. Our challenge will be to ensure a balance between broad community interests and the more narrow interests of private entities when it comes to influencing the health of our society. The stakes are high because whoever owns Big Data may not own the Earth, but they certainly will be a powerful force in determining the well-being of its inhabitants for the foreseeable future.

Edward Ehlinger is Minnesota’s Commissioner of Health.

REFERENCES