Health IT Policy Committee, Privacy and Security Working Group
Public Hearing on Health and Big Data
Panel on Health Big Data Opportunities
December 5, 2014

Testimony of Stephen J. Downs
Chief Technology and Information Officer
Robert Wood Johnson Foundation

Good afternoon. My name is Steve Downs and I'm the chief technology and information officer at the Robert Wood Johnson Foundation (RWJF), based in Princeton, New Jersey. In that role, and in other positions at RWJF over the last 12 years, I have been a funder, supporter and observer of a variety of health IT and now big data efforts. The subject of this panel -- "Health Big Data Opportunities" -- is quite broad and I will focus my remarks on the areas of potential benefit with which I am most familiar.

Robert Wood Johnson Foundation, the nation's largest philanthropy dedicated solely to improving the public's health, has a vision of building a Culture of Health[1] in the United States where everyone has the opportunity to make healthy choices every day, where our health care system consistently delivers high quality care -- and our economy is no longer burdened by excessive health care spending, where people work across sectors to improve the health of communities and where they hold public officials accountable for making policy changes that improve health. In short, it is a vision where health is a fundamental American value.

Achieving a Culture of Health involves recognizing that much of what drives health -- in individuals and in populations -- occurs outside of the formal health care system. While high value medical care is important, health happens where people live, work, learn and play -- in the communities where they walk and buy their food, in their workplaces, and in the schools they attend. The experience of health plays out in our daily routines -- what we eat for breakfast, how we get to work (if we are able to work), whether we have opportunities for physical activity, the environmental exposures we receive, and the stress upon our minds and bodies. Research has shown that these day-to-day experiences and health behaviors are strongly affected by a set of non-medical factors -- factors such as income, education, housing and access to transportation. RWJF's Commission to Build a Healthier America showed, for example, that poor Americans are more than three times as likely as Americans with upper-middle-class incomes to suffer physical limitations from

[1] For an overview of RWJF's vision, see "Building a Culture of Health," http://www.rwjf.org/en/about-rwjf/annual-reports/presidents-message-2014.html

a chronic illness.[2] These factors are known collectively as the social determinants of health.

As I look to the potential health benefits of big data, there is a broad variety of possibilities. Within clinical medicine and health care delivery, there is a longstanding opportunity to convert the data from day-to-day care, now increasingly captured electronically, to fuel a learning health care system in which each data point feeds an ever refining collective understanding of which treatments work for which conditions suffered by which people. And the data captured as part of the health care delivery process can be used for many different questions. For example, RWJF has commissioned cloud EHR services provider athenahealth to examine the impact of the Affordable Care Act on insurance coverage, using its aggregate patient data.[3]

One area where I see great opportunity for benefit is in the utilization of data that provides insight on the impact of the social determinants on health. As payment models evolve and as health care providers grow increasingly accountable for the health outcomes of their patients, they will need to pay increasing heed to the role that social determinants play in their patients' outcomes. One can imagine health care systems leveraging data on housing stock, on community walkability, safety and violence, availability of early childhood services, food accessibility, transportation infrastructure and more to understand the barriers faced by individual patients and by their population as a whole. For example, RWJF Young Leader Ruben Amarasingham, at Parkland Health and Hospital System in Dallas, has been using data on insurance status and the number of home addresses one has had in the past year, to develop predictive algorithms to understand which heart failure patients are at risk for readmission. By providing additional services to those at determined to be at greatest risk, Parklands was able to reduce readmission rates by 30%.[4]

One can imagine applying these data at a population level as well. For example, a health system, using a disease registry, could identify patients with diabetes and, tapping geographic information, could analyze both neighborhood walkability and food access metrics for the neighborhoods where high percentages of their patients reside. These

---

[2] See "Beyond Health Care: New Directions for a Healthier America," p. 10. http://www.rwjf.org/content/dam/farm/reports/reports/2009/rwjf40483

[3] See "ACAview: Tracking the Impact of Health Care Reform," http://www.rwjf.org/en/research-publications/find-rwjf-research/2014/07/acaview--tracking-the-impact-of-health-care-reform.html

[4] Moukheiber, Z., "Can Algorithms Save Parkland Hospital?" Forbes, December 28, 2012. http://www.forbes.com/sites/zinamoukheiber/2012/12/28/can-algorithms-save-parkland-hospital/

analyses could lead the health system to conclude that in order to improve the health outcomes of its patients, it would need to work on environmental changes to improve walkability and/or food access in those communities. This kind of "hot spotting," pioneered by Dr. Jeffrey Brenner in Camden, New Jersey,[5] could lead to more community-based interventions above and beyond providing care to individuals.

A second area of potential benefit is the utilization of personally generated health data. Since the advent of the smartphone and, more recently, wearable sensors, increasing numbers of people are generating data about an ever widening array of health-related behaviors and experiences. Apps and wearables track diet, steps, workouts, sleep, mood, pain, menstrual cycles, heart rate, as well as -- critically -- time and location. Recent products also include hydration, breathing rates and patterns, stress and whether one is sitting or standing. And still others are able to infer health experiences by analyzing data, such as location, movement and social activity -- not typically considered "health data."[6] These apps and wearable devices -- particularly those that collect data passively through devices are that are worn or carried around the clock -- are starting to provide a window onto people's day-to-day experience with health. And one can imagine tapping other sources such as supermarket loyalty card data or credit card purchase data to fill out the picture even further. Given that the most popular services count subscribers in the tens of millions, these data can lead to new insights -- at a population level -- on behavior, environment and health.

These data could be used for research purposes and also for the practice of public health. Three characteristics of personally generated data -- the breadth of variables about day-to-day experience that can now be captured, the near continuous nature of its collection and the sheer numbers of people generating the data -- make it extremely interesting for research. Returning to the notion that health happens where one lives, learns, works and plays, data are now being generated about people's experiences in those environments. The combinations of data could be quite interesting -- imagine research that explores the relationships between neighborhood walkability and steps taken, between food access and dietary patterns, or between stress and geographic location.

Some early examples of research on aggregated data are emerging.

---

[5] See "How 'Hot Spotting' Cut Health Care Costs by 50%," http://www.rwjf.org/en/about-rwjf/newsroom/features-and-articles/Brenner11.html

[6] For example, see http://www.ginger.io

- RunKeeper has used data it collects on people's workouts to post information on the frequency, average pace and average distance people run in different U.S. states, noting that runners in northern states tend to run a bit farther and faster than runners in most southern states.[7]

- Data from RunKeeper's Breeze app, which tracks overall movement, showed a comparison of the different times of day when people get their steps in 25 different countries.[8]

- Massive Health (now a part of Jawbone) used data from its Eatery app, which had people photograph and rate the healthiness of their meals, to profile how America eats. Using data from 500,000 meals, they generated interesting possible insights, including the suggestion that people who eat breakfast eat healthier food than those who don't eat breakfast throughout the day.[9]

These are early, and in many ways, very modest examples of how these data might be used. At Robert Wood Johnson Foundation, we recognized that there is much work to be done before the promise of using these data for research can be fully realized. We supported an initial exploration -- known as the Health Data Exploration project -- into this topic and our grantee, Calit2, released a report[10] outlining the challenges that needed to be overcome. Key issues include privacy, informed consent, access to the data and data quality. The fact that adoption of smartphone apps and especially wearable sensors is, while growing, still skewed demographically, also poses methodological challenges. We have continued our involvement in this area and are now supporting a research network[11] that is pulling together researchers, data scientists from companies that gather data, and other stakeholders to work on these issues and learn through the experience of a number of small research projects.

---

[7] "In the USA, Who Runs the Fastest, the Farthest, and the Most Often?" http://data.runkeeper.com/rk-usa-running-stats-by-state

[8] "Rise and Shine: When People Around the World Get Their Steps In, "http://data.runkeeper.com/breeze-habits

[9] "How We Eat: Analyzing ½ a Million Meals," http://blog.massivehealth.com/post/21377192742/how-we-eat-analyzing-half-a-million-meals

[10] Personal Data for the Public Good," http://hdexplore.calit2.net/wp/project/personal-data-for-the-public-good-report/

[11] http://hdexplore.calit2.net/wp/network/

In addition to general insights that might result from research on aggregated data, there are more immediate public health applications. Data on people's everyday patterns can help public health and other government agencies understand community needs, make interventions and monitor the responses to those interventions. For example:

- In Louisville, Kentucky, with RWJF support, the Community Foundation of Louisville and the City of Louisville are working with Propeller Health, which makes a GPS-connected asthma inhaler. Together, they are mapping hotspots of asthma inhaler use in an effort to understand the environmental and neighborhood drivers of asthma in their community.[12]

- The Oregon Department of Transportation is working with Strava, whose app helps cyclists track their rides, to analyze when and where people ride bicycles so that they can see where bike lanes are needed or where current traffic patterns might pose safety threats.[13]

- Jawbone, makers of the UP wristband that tracks activity and sleep, used data on when people woke up in different locations to show the time and distribution of the impact of the 2014 earthquake in Napa.[14]

- Scientists at Johns Hopkins University have used Twitter data to map the incidence of flu by time and geographic location.[15]

Again, there are issues associated with using the sort of data featured in these examples for public health purposes -- there are questions of representativeness, questions of access to the data, questions of privacy and user consent, and methodological issues. Nevertheless, these few examples suggest that there might be promise in applying certain kinds of personally generated health data to certain public health questions.

In summary, we are still very much at the dawn of these new possibilities. We have seen glimpses of exciting potential benefits, but there are cautions to be heard and challenges

---

[12] "Is Louisville, Kentucky, the new face of asthma healthography?" http://propellerhealth.com/2014/11/louisville-kentucky-new-face-asthma-healthography/

[13] Davies, A. "Strava's Cycling App Is Helping Cities Build Better Bike Lanes," WIRED, June 3, 2014, http://www.wired.com/2014/06/strava-sells-cycling-data/

[14] "How the Napa Earthquake Affected Bay Area Sleepers," https://jawbone.com/blog/napaearthquake-effect-on-sleep/

[15] "Johns Hopkins researchers go local with their Twitter flu tracking efforts," http://hub.jhu.edu/2014/03/18/twitter-data-flu-tracking-new-york

to be overcome. As noted in the just-released report[16] by the JASON group, there is much work to be done to develop an infrastructure that would facilitate the integration of all of these types of data. It is important to recognize that in every technology innovation and adoption cycle, many of the imagined benefits will not pan out and many of the challenges will turn out to be just transitional. Bill Gates has noted "we always overestimate the change that will occur in the next two years and underestimate the change that will occur in the next ten," and I think that applies to the benefits of big data in health, though I'd prefer to be less precise about the timeframes. Given the early, transitional stage we are in, full of possibility and also potential pitfalls, it is important to allow for experimentation, for the technology and the methods to get better, and most importantly, to allow our institutions to catch up so that they can learn how best to take advantage of the opportunities and realize the potential benefits most fully.

[16] Data for Individual Health. http://healthit.gov/sites/default/files/2014-JASON-data-for-individual-health.pdf