

# Chapter 16

## The Impact of Big Data on the Physician

**Elizabeth Le, Sowmya Iyer, Teja Patil, Ron Li, Jonathan H. Chen, Michael Wang, and Erica Sobel**

The practice of medicine has historically been a very unilateral process. Physicians go through a great deal of training to possess near exclusive knowledge of human health and illness. They are trained to collect information from patients on a subjective level, combine this with objective data in the form of the physical exam, laboratory values, and imaging studies and then independently process this information into a diagnosis with a recommended course of treatment for the patient. Patients have been traditionally expected to accept and follow this recommendation, trusting that their physician has been adequately trained and is giving the best recommendation for their health without having any particular knowledge of the condition or treatment themselves. As the age of technology and now big data has evolved, this process and relationship is undergoing dramatic changes. While the advances are remarkable, the amount of information generated can be overwhelming and ultimately stifling to individual physicians. As such, a whole other realm of technology is being created to assist in analyzing this data, ensuring that physicians are utilizing it to the highest potential, and adhering to the most proven treatment regimens. As technology has evolved to support a deeper understanding of illness,

---

E. Le • S. Iyer (✉) • T. Patil  
Veterans Affairs Palo Alto Healthcare System, Palo Alto, CA 94550, USA  
e-mail: [elizabeth.le@va.gov](mailto:elizabeth.le@va.gov); [sowmya.iyer@va.gov](mailto:sowmya.iyer@va.gov); [teja.patil@va.gov](mailto:teja.patil@va.gov)

R. Li • J.H. Chen  
Stanford University, Palo Alto, CA 94305, USA  
e-mail: [ronl@stanford.edu](mailto:ronl@stanford.edu); [ronc101@stanford.edu](mailto:ronc101@stanford.edu)

M. Wang  
University of California, San Francisco, CA 94720, USA  
e-mail: [Michael.wang4@ucsf.edu](mailto:Michael.wang4@ucsf.edu)

E. Sobel  
Kaiser Permanente Santa Clara Medical Center, Santa Clara, CA 95051, USA  
e-mail: [erica.m.sobel@kp.org](mailto:erica.m.sobel@kp.org)

so has a multitude of new types and ways to collect data. Data is being both passively and actively collected in every aspect of life, from specific biometric data including glucose readings and blood pressure to everyday data including tracking an individual's daily steps and calorie counts for each meal eaten. Increasingly, these once ordinary activities of daily life are being analyzed as components of health and living, and becoming a portion of the medical chart. Social engagement and interaction with the health system is also growing and changing in directions never before anticipated or experienced. Patients now have the opportunity to directly compare their doctors and hospitals. Information about medical conditions and healthcare is more readily available to consumers than ever before. Historically, one would have had to speak to a physician to learn about their condition and treatment. Now, anyone can simply "Google" their symptoms and be provided with a list of diagnoses and potential treatments. Patients seek to learn more about the therapies being recommended, as well as form communities of individuals with similar diagnoses to compare treatment plans and lend support. With increasing adoption of electronic health records (EHRs), and increasing innovation in areas of big data and healthcare, the ways that physicians interact with patients, approach diagnosis and treatment, and strive for improved performance at an individual and a health system level are evolving. This chapter will discuss many of the big data applications that practicing physicians and their patients encounter.

## **16.1 Part 1: The Patient-Physician Relationship**

### ***16.1.1 Defining Quality Care***

Consumers in the healthcare economy strive to be cared for by the "best" doctors and hospitals, in particular as society increasingly moves towards individualized, consumer-centered healthcare. However, physicians and hospitals are not chosen based solely on the quality of care they provide. Geography and health insurance often play a big role in pairing patients with physicians. Beyond this, patients largely depend on word of mouth recommendations to infer quality of care (Tu and Lauer 2008). The result is a system in which patients choose healthcare providers based on subjective and non-standardized metrics. The challenge lies not only in quantifying quality, but then making that information transparent to the consumer. A data science solution can lend rigor and clarity to answering the question: With whom and where can I get the best care?

A general guideline for what can be considered a data science solution to assessing healthcare quality contains two major criteria. First, the methods must process large amounts of data from multiple sources, and second, the relevant results of the analysis must be presented in an accessible, interpretable and thus usable manner. This work is done in the belief that it can lead to better health decisions. However, quantifying healthcare quality is incredibly challenging as so many vari-

ables affect physician and hospital success. Factors to be considered include facets of patient experience, availability of appointments/access, treatment outcomes, and complication rates. Some measures are more easily quantifiable than others, and determining how heavily each factor should be weighed is quite subjective. Several tools have been recently developed to allow healthcare consumers to compare and contrast physicians and hospitals on key objective measures.

### ***16.1.2 Choosing the Best Doctor***

The Internet has given consumers an immense amount of new information regarding products and services. One may be tempted to cite online review sites such as Yelp.com as a big data solution for connecting patients with the best physicians. These online review forums report patient satisfaction through individual anecdotal experience in an open text format and a self reported star rating of overall experience. It is interesting to note that multiple studies have shown a strong positive correlation between a hospital's patient experience scores and strong adherence to clinical guidelines and better outcomes (Chaterjee 2015). Still, while patient satisfaction is an important measure of quality care and correlates well with other measures of success, patient satisfaction is only one metric. Additionally, not only is the data sourced through a single method, online review forums do not analyze nor present the data in an optimized manner.

Health Grades, Inc. (Denver, CO, USA), a data science company, compiles quality metrics on individual physicians as well as hospitals. The physician rating is comprised of demographic information about the individual (i.e. education, specialty, board certification, malpractice claims) and a separate patient satisfaction survey. While the survey portion has a free text Yelp-like review segment, it improves upon this model by standardizing the reviews through a series of questions. Patients rank the physician on a scale of 1–5 stars. Some questions include: level of trust in a physician's decisions, how well a physician listens and answers questions, ease of scheduling urgent appointments, and how likely the patient is to recommend this physician to others. Unlike Yelp, this scale provides more quantitative information about the patient experience in a standardized and thus comparable format. Despite this comparative data, a weakness of the Health Grades system is that it does not recommend any doctor over another and has no outcomes data. Therefore, two physicians with similar education and patient satisfaction scores will appear equal even if one physician has much higher complication rates than another. In addition, there is no single algorithm or composite score to compare physicians to one another, and the results are not customized to the patient. Instead, it is a platform through which demographic information about the physician and patient satisfaction review results are easily accessed.

A more personalized data science solution has been put forth by Grand Rounds, Inc. (San Francisco, CA, USA) which uses a multivariate algorithm to identify top physicians. The proprietary "Grand Rounds Quality Algorithm" uses clinical data

points from 150 million patients in order to identify “quality verified” physicians (Freese 2016). 96% of practicing physicians in the United States (~770,000) have been evaluated by the Grand Rounds Quality Algorithm (Grand Rounds n.d.). The algorithm scores variables such as physician’s training, publications, affiliated institution, procedural volumes, complication rates and treatment outcomes. These and other non disclosed variables are combined to create a composite quality score. The algorithm takes into account the quality score as well as location, availability, insurance networks and expertise in specialty topics. Finally, patient characteristics derived from the patient medical record (i.e. languages spoken) are also included in order to then match the individual patient to a “quality verified” physician. As a result, the recommended cardiologist for one patient will not necessarily be the best for their family member or friend. The great benefit of this model is that it synthesizes a massive amount of information on a vast number of physicians, and then works to individualize the results for each patient.

The major criticism of the Grand Rounds model is that the proprietary algorithm is not transparent. It is not known what data sources the company uses nor how it extracts individual physician level outcomes data. The exclusion criteria used and how different variables are weighted in the algorithm is also unknown. As a result of this lack of transparency, the model has not been externally validated.

### ***16.1.3 Choosing the Best Hospital***

Compared to physician quality metrics, patients have more options when examining hospital quality using tools such as Health Grades, U.S. News World Report, and Leap Frog. As an example, Health Grades evaluates hospitals in three segments: clinical outcomes, patient experience and safety ratings. The clinical outcomes results are derived from inpatient data from the Medicare Provider Analysis and Review (MedPAR) Database, as well as all payer state registries (Healthgrades 2016a, n.d.). From these databases it presents either mortality rate or complication rate for 33 conditions and procedures including heart failure, sepsis, heart valve surgery, and knee replacement. Each of the 33 conditions has a prediction model which takes into account confounding variables such as patient age, gender and relevant comorbidities. Using logistic regression, Health Grades then compares the health care systems actual mortality/complication rate against their own predicted mortality/complication rate for that condition in that hospital. Finally, hospitals are presented to the patient as “below expected”, “as expected” or “better than expected”. The second segment reports patient satisfaction based on the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) (Healthgrades 2016b, n.d.). HCAHPS data is available from the Centers of Medicare and Medicaid Services (CMS). It is a 32 question survey administered to recently discharged patients with questions focusing on factors including doctor and nurse communication, pain control, hospital cleanliness and noise levels, post discharge instructions, and whether they would recommend this hospital to others. The final

segment, patient safety, is constructed from Medicare claims (Healthgrades 2016c, n.d.). Health Grades has chosen fourteen patient safety indicators (PSI) such as catheter related bloodstream infections, development of pressures ulcers and mortality rates in post surgical patients. Like the clinical outcomes segment, Health Grades calculates predicted complication rates for each hospital. The predicted complication rates are adjusted for complexity of cases using the Medicare Case Mix Index. So hospitals with more complicated cases are given higher predicted rates of PSI. The actual rates are then compared to the predicted and the hospital is ranked either “below expected”, “as expected” or “better than expected” for each PSI. Health Grades presents its final results in a user-friendly format with simple graphics using star ratings and pie charts.

The major benefits of the Health Grades model are that the methodology and data sources are transparent and results are displayed in a user-friendly manner. However, there are also many limitations, both with the statistical analysis and with the final product. For instance, comparing a hospital’s actual complication rate to a predicted complication rate is fraught with statistical challenges. Even though the predicted complication/mortality rates are corrected for age and comorbidities, it is very hard for a logistic regression to take into account all confounding variables. This is especially true because patient’s comorbidities are not always accurately reflected in the billing data. Billing data is also significantly delayed, thus present day analysis is often carried out on data from several years ago. In addition, the results for each hospital are not generalizable to every condition or patient. Since patients can only see data regarding 33 conditions, this leaves out information on less common diseases or rare surgeries. In addition, data from some patients, such as those who are discharged to hospice or have metastatic cancer, are not included in the Health Grades outcomes analysis. Aside from issues of generalizability, there is a notable limitation of the final product application. While the Health Grades platform provides a report card for individual hospitals, it does not offer an easy way to compare multiple institutions.

Today patients have many more tools for making a data driven decisions in choosing their healthcare provider. However many challenges still exist both for determining physician and hospital quality, and for matching patients and healthcare providers in an optimal manner. In an ideal future state, treatment outcomes data at a physician, department and institutional level would be available and sourced directly from the electronic medical record as opposed to deduced from billing data. Additionally, treatment outcomes, patient satisfaction and safety metrics would be available at faster turnaround times. If available, the applications could extend beyond patient’s using data to select physicians and allow providers themselves to have more real time feedback on their individual performance, which could ultimately influence practice patterns.

### ***16.1.4 Sharing Information: Using Big Data to Expand the Patient History***

The practice of medicine is undergoing a cultural shift in which the paternalism ingrained in the patient-physician relationship is no longer expected, and patients are increasingly and rightfully becoming partners in their medical care. At the same time, technology has advanced to create near limitless access to information and data generation, further supporting the patient bid for greater autonomy. From creating communities and promoting patient advocacy to increasing use of shared decision-making (Shay and Lafata 2015), patients are playing larger and larger roles in their health.

Despite this shift, one area in which patients still have surprisingly little input is with regards to their health data. Currently a patient's history, whether obtained directly from the patient or their care-taker, in verbal or written form, resides within an electronic health record (EHR) controlled by the physician or medical system. In the most common EHRs, patients may at most be able to view their health data but do not have the ability to add to or edit their record. As patients become more engaged and involved in their care, it follows that they should have more direct input and ownership over their health data. This increasing responsibility and management over personal health data has the potential to lead to increasing patient activation, a concept in which patients have the knowledge, skills, and confidence to become more effectual managers of their own health (Hibbard et al. 2004). Studies of patient activation, a cornerstone in the management of chronic disease (Clark 2003), have demonstrated improvement in healthy behaviors (Rask et al. 2009) and overall health outcomes (Dentzer 2013; Greene and Hibbard 2012). Giving patients ownership over their health data should therefore help promote activation for improved health outcomes.

### ***16.1.5 What is mHealth?***

What will this expansion and shared ownership in patient health data look like? Where will this data come from? How will it enter the electronic record? How will it be used? Alongside this change in the patient-physician partnership has been a rapid expansion in healthcare oriented devices and applications, or "apps," giving rise to the field of mobile health (mHealth). mHealth describes mobile technologies used in healthcare diagnostics, monitoring, and systems support, and currently encompasses monitoring for the collection of individual biometric and environmental data, personal emergency response systems (PERS), telemedicine, mobile medical equipment, Radio Frequency Identification (RFID) tracking, health and fitness software, mobile messaging, and electronic medical records. mHealth has developed and flourished under the jurisdiction of the medical establishment, with remote patient monitoring (RPM) of chronic conditions, most prominently

using telehealth, capturing the largest portion of the mHealth market. As it stands, telehealth services are projected to cover 3.2 million patients in 2018, up from 250,000 in 2013 (Japsen 2013). The next frontier in mHealth is shifting from a focus on physician-collected data via RPM to patient-collected data using consumer targeted devices and applications.

### ***16.1.6 mHealth from the Provider Side***

Remote patient monitoring (RPM) has largely developed within integrated health-care systems and institutions with the goal of better managing chronic conditions. Two leading examples of RPM are with the Veterans Health Administration (VHA) Care Coordination/Home Telehealth (CCHT) program and Partners Healthcare in Boston, both of which have demonstrated substantial improvements in patient health care and utilization (Agboola et al. 2015 and Darkins et al. 2008).

Since 2003 the VHA, within the Office of Telehealth Services, has used the CCHT program to help veterans with chronic medical issues including diabetes, congestive heart failure, hypertension, posttraumatic stress disorder, chronic obstructive pulmonary disease, and other chronic conditions coordinate their care as well as receive remote patient monitoring. The home telehealth devices are configured to communicate health status including symptoms as well as capture biometric data. This data is transmitted and remotely monitored by a care coordinator working in conjunction with the patient's primary care provider. For the VHA, one of the primary goals of monitoring has been to reduce use of institutional medical care, particularly for veterans who may live remotely, and instead promote patient activation, self-management, and the earlier detection of complications. For patients with chronic illnesses enrolled in CCHT, monitoring has resulted in a 25% reduction in bed days of care, 19% reduction in hospital admissions, and high overall rates of patient satisfaction. By 2010, over 70,000 veterans have been enrolled in CCHT with plans for expansion to cover more.

Partners Healthcare (Partners Healthcare, Boston, MA, USA), a private integrated health system, has been using its own RPM system to improve outcomes in patients with heart failure. The Connected Cardiac Care Program (CCCCP) provides home telemonitoring combined with nursing intervention, care coordination and education for patients with heart failure. Telemonitoring consists of the daily transmission of weight, heart rate, pulse and blood pressure by patients with the program demonstrating a significantly lower rate of hospitalization for up to 90 days, and decreased mortality within the first four months after enrollment at discharge. This model has thrived on strong institutional leadership as well as the overarching goal of activating and engaging patients in greater self-care through technology.

These examples demonstrate how data gathered from patients remotely can be used to drive real time clinical decisions and improve healthcare outcomes. Still, they represent a model of provider driven data collection. This data is reliant on large teams of care coordinators, often nursing based, for interpretation and use, and does

not have the benefit of further analytics to distill and detect trends. The emerging frontier in health related data generation will come from the hands of patients and their caregivers, from data captured on consumer marketed devices including smartphones and wearable devices, and processed with sophisticated analytics to better support provider workflow.

### ***16.1.7 mHealth from the Patient Side***

In the U.S., smartphone ownership has increased from 35% of adults in 2011, to 64% of adults in 2014, with 62% of smartphone owners using their smartphone to look up information about a health condition within the past year (Smith 2015). 34% of adults have downloaded an app meant to support healthy living, and one in five have downloaded and regularly use an mHealth app (Witters and Agrawal 2014). Widespread ownership and use of smartphones is now evolving into ownership and use of wearable devices, including fitness trackers and smartwatches. A recent report by PriceWaterhouseCoopers has shown that over 20% of American adults now own a wearable device with the market share of these devices increasing each year (PwC 2014).

The rapid innovation in mobile consumer devices, including smartphones and wearables enabled to capture biometric and health-related data, and the increasing accessibility of these devices and apps has allowed consumers to capture their own health related data, or patient-generated health data (PGHD). PGHD encompasses health-related data including history, symptoms, biometric, and environmental data that is created, recorded, gathered, or inferred by or from patients or their care partners to help address a health concern (Shapiro et al. 2012). It is unique because patients determine with whom it will be shared. Use of consumer targeted mHealth devices and applications has allowed for greater capture of patient selected and controlled data, in contrast to provider-selected variables traditionally employed in remote patient monitoring.

With the generation of such large volumes of raw data, a new industry has emerged to determine how to best integrate and translate PGHD in a clinically useful manner into EHRs. One of the main pillars in emerging technologies is in the design and use of ecosystem-enabling platforms intended to bridge the divide between humans and technology (Gartner 2016). One particular platform-enabling technology as it relates to healthcare is seen in the Internet of Things (IoT). The IoT is an emerging concept that at its simplest is considered solutions generated from interconnected devices, and is used to describe “the range of new capabilities brought about by pervasive connectivity . . . involving situations where network connectivity and computing capability expand to objects, sensors, and everyday items that exchange data with little to no human involvement” (Metcalfe et al. 2016). As it pertains to health, the IoT promotes a user-focused perspective in how data is managed and information exchanged, particularly between consumers and their devices, allowing consumers to engage in greater self-monitoring and management.



The IoT is an opportunity to passively transmit data from consumer devices, which can then be transformed into big data using sophisticated analytics, and ultimately translated into information and insights that can aid in improving individual health.

Further spurring innovation in PGHD is Stage 3 of the meaningful use electronic health record incentive program by the Office of the National Coordinator for Health Information Technology, which has placed a high priority on bringing PGHD into the EHR (U. S. Department of Health and Human Services 2013). Stage 3 specifies that EHRs should “provide 10% of patients with the ability to submit patient-generated health information to improve performance on high priority health conditions, and/or to improve patient engagement in care.” With this mandate, several companies including eClinicalWorks (Westborough, MA, USA) and Apple, Inc. (Cupertino, CA, USA) have begun to build and deploy solutions, using platforms like the Internet of Things, to seamlessly integrate and translate PGHD captured by consumer devices into EHRs and create analytics to support clinical decision-making.

eClinicalWorks, a leading ambulatory EHR vendor, is integrating data from wearable devices into its Health & Online Wellness personal health record healow<sup>®</sup> with the goal of enhancing patient engagement in their health (Caouette 2015a). To date 45 million patients have access to their health records through healow<sup>®</sup>. According to a survey conducted online by Harris Poll and commissioned by eClinicalWorks, 78% of patients with wearable devices using them more than once a month feel that their physicians would benefit from access to the information collected (Caouette 2015b). Using its cloud platform, the Internet of Things, third party hardware platforms can collect, store, and analyze PGHD from home monitoring and wearable devices. These devices can include activity trackers, weight scales, glucometers, and blood pressure monitors. The healow<sup>®</sup> IoT contains analytics and dashboards designed to provide patients and physicians with high yield data culled from PGHD for more informed clinical decision-making. Though still early, this partnership has the promise to use big data to improve patient care using patient owned devices.

Although not a traditional corporation in the field of healthcare, Apple has been leveraging its existing technologies to collect, capture, and integrate PGHD. Apple has been rapidly evolving its HealthKit app, initially launched in 2014, into a platform to allow health data interoperability. It has partnered with numerous EHRs and over 900 health devices and apps and continues to expand its partnerships and potential applications. Working with large centers like Duke, Stanford, and most recently, in its largest patient integration of more than 800,000 patients, Cedars-Sinai Medical Center, Apple’s HealthKit allows for the centralization and transmission of data from consumer devices to healthcare operating platforms (Higgins 2015). Given the potential for large amounts of unfiltered data to be transmitted, HealthKit incorporates flow sheets to keep PGHD separate from other data and allows providers to specify the frequency of data transmission. As there are no clear recommendations for how PGHD is to be incorporated or handled once within EHRs, healthcare systems like Duke have created modified consents so that patients understand that the PGHD data transmitted has no guarantee to be viewed

or acted upon in a real time manner so as to avoid creating a false sense of security for patients (Leventhal 2015). Ultimately though, in contrast to the current state of health data in the EHR, patients retain control their PGHD, deciding to whom and what data is transmitted.

In one of the most promising published applications of mHealth and big data in clinical care, Apple's HealthKit has recently been used in a pilot to improve the management of pediatric diabetics in Stanford's outpatient clinics. Stanford partnered with Apple, a major glucose monitoring company Dexcom (San Diego, CA, USA), and its EHR vendor Epic (Verona, WI, USA) to automatically integrate patient glucose readings into its EHR and provide analytics to support provider workflow and clinical decision making (Kumar et al. 2016). Prior processes to input data from continuous glucose monitoring devices into the EHR have required either manual entry or custom interfaces, both of which limit widespread applicability and also impose a time delay variable, as data is only available to providers at clinic visits. Using the prior mentioned technologies, the study investigators were able to establish a passive data communication bridge to link the patient's glucometer with the EHR via the patient/parents smartphone. In this setting the patient wears an interstitial glucose sensor connected to a transmitter which sends blood glucose readings by Bluetooth connection to the Dexcom Share2app on the patient's Apple mobile device. The Dexcom Share2 app then passively shares glucose values, date, and time with the HealthKit app, which transmits the data passively to the Epic MyChart app. The MyChart patient portal is a part of the Epic EHR and uses the same database so that the PGHD glucose values are populated into a standard glucose flow sheet in the patient's chart. Importantly, this communication bridge results in the passive integration of the patient's data into the chart but can only occur when a provider places an order in the patient's electronic chart and the patient accepts the request within their MyChart app, placing ultimate control of the data with the patient/parent. Hesitations on the part of providers to accept integration of PGHD into the EHR have often centered on concerns over receiving large and potentially unmanageable volumes of data, which may lead to increasing liability and unrealistic patient expectations, as well as to what degree data is actionable (National eHealth Collaborative 2013). Providers have also expressed concern about the financial impact of using staff and physician time to review the data. The study investigators addressed these concerns by creating an analytic report to triage patients and identify actionable trends between office visits based on home glucose readings, supporting rather than hindering provider workflow. This report was not meant to replace real-time patient/parent monitoring and thus verbal and written notification was used to establish patient/parent expectations regarding only intermittent provider monitoring. The report was generated every two weeks to identify trends, such as episodic nocturnal hypoglycemia, rather than provide real time glucose monitoring. When viewed by the provider, these reports could also be shared and discussed with the patient/parent via MyChart to create an ongoing dialogue for care between visits. In this study, several actionable trends including nocturnal hypoglycemia in a toddler and incorrect carbohydrate counting in a teenager were identified leading to improvements in overall glycemic control

between office visits. Additionally, participants were noted to express gratitude that actionable trends were brought to their attention and there was no report of frustration regarding lack of contact for specific hypo or hyperglycemic episodes, further highlighting the importance of early expectation management.

This pilot study, thus far the only study to demonstrate automatic integration of PGHD into the EHR, demonstrated that integration using widely available consumer technology is not only possible, but when combined with smart and intuitive analytics can improve provider workflow for reviewing data and communicating with patients leading to better care for patients. Notably, this workflow did not require any institution-level customization or a specific EHR vendor. Additionally, other companies like Microsoft Health and Google Fit are both in the midst of developing similar patient-generated data platforms making it likely that any mobile device, not specifically Apple devices, may be able to be configured to perform similarly.

### ***16.1.8 Logistical Concerns***

The enormous promise of consumer devices and integrated health data platforms to revolutionize health care is two-fold in the goal of the establishment of the patient as an owner and active user of their health care data, and in the implementation of big data to support physician workflow and clinical decision-making. Inherent in the development and adoption of any new technologies, particularly within health care, are the logistics of addressing accessibility, privacy, security, and liability.

### ***16.1.9 Accessibility***

Despite rapidly expanding ownership and usage of mobile devices, economic disparities and technological literacy both stand as potential barriers to accessibility. Use of new technologies is most often associated with younger and more financially stable demographic categories. This is particularly seen in the wearable devices market, where approximately half of consumers are between the ages of 18–34 and one-in-three have a household income of greater than \$100,000 (Nielsen 2014). However, approximately two thirds of Americans now own a smartphone of some sort with 10% owning a smartphone without any other form of high-speed internet access at home, making them heavily dependent on their smartphone (Smith 2015). Those who tend to be most heavily dependent on a smartphone for online access include younger adults, those with low household incomes and levels of educational attainment, and non-whites. This implies that shifting technology toward smartphones may still be a reasonable strategy to cover diverse demographic groups. Interestingly, as they relate to health, consumers have indicated that they are not willing to pay much for wearable devices but would be willing to be

paid to use them (PwC 2014). This may signal a more broad interest in wearable technologies for health independent of financial status, and may offer an opportunity for insurers, providers, and employers to step in and potentially level the playing field for patients. In fact, consumers have noted that they are more willing to try a wearable technology provided by their primary care doctor's office than they are for any other brand or category (PwC 2014).

### ***16.1.10 Privacy and Security***

With regards to privacy and security, health care providers must address HIPAA requirements and malpractice issues while developers must pay attention to standards for product liability including Federal Drug Administration (FDA) and Federal Trade Commission (FTC) rules to protect patients (Yang and Silverman 2014). In the United States, it is generally held that an individual's medical record is owned by the provider or institution that retains the record, not the individual patient the record describes. Patients are however still covered by privacy provisions included in the Health Insurance Portability and Accountability Act (HIPAA) of 1996, that ensures the confidential handling and security of protected health information (PHI) (U.S. Department of Health and Human Services 2013). With increasing use of mHealth technologies, Congress has expanded the use of HIPAA through the Health Information Technology for Economic and Clinical Health (HITECH) Act (U.S. Department of Health and Human Services 2009), which sets forth requirements for mandatory breach notifications. Despite this, HIPAA coverage as it pertains to mHealth technologies remains complex. When a patient's health data is in the possession of health providers, health plans, or other "covered entities", it is protected under HIPAA. When it is transmitted among individuals or organizations that are not covered entities under HIPAA, it is not protected (Fisher 2014). For example, if a patient checks their heart rate and the data is recorded on their mobile device, it is not covered by HIPAA. However, if those readings are sent to their physician for use in clinical care, the data becomes HIPAA protected. This and many other scenarios will need to be identified and clarified to ensure data from mHealth apps and devices incorporated into health care are appropriately protected.

Patients themselves are becoming more and more aware of privacy concerns and even within the context of traditional methods for health information transfer (fax, electronic transfer), more than 12% of patients noted withholding health information from a health care professional due to privacy concerns (Agaku et al. 2014). Inspection of 600 of the most commonly used and rated English-language mHealth apps showed that only 30.5% had a privacy policy, and that bulk of these policies required college-level literacy to understand and were ultimately irrelevant to the app itself instead focusing on the developer (Sunyaev et al. 2015). More stringent attention from developers with regards to mHealth data privacy will be needed to both protect data and ensure consumer confidence in their devices.

Given the nature of mHealth technologies, data from mobile devices (that are themselves easily lost or stolen), is transmitted to cloud-based platforms over wireless networks that themselves are prone to hacking and corruption. Even prior to the expansion of mHealth devices, examples of lost or stolen patient data have already populated new cycles. In 2012, Alaska's Medicaid program was fined \$1.7 million by HHS for possible HIPPA violations after a flash drive with patient health information was stolen from the vehicle of an employee (U.S. Department of Health and Human Services 2012). Interestingly, health hacking has also become increasingly prevalent given the relative ease of hacking medical systems and devices, and the increasing worth of health care data. It is estimated that health care data is currently more lucrative than credit card information for fraudulent purposes (Humer and Finkle 2014). As more patients, providers, and healthcare organizations use mobile health technologies to augment and conduct patient care, more attention to security features and protocols will be needed to ensure privacy.

### ***16.1.11 Regulation and Liability***

The regulation of mHealth as it relates to medical licensure and liability is a complex issue without clear guidelines or answers. In the current state, mHealth devices and apps may be regulated in a piecemeal fashion by several agencies. Potential regulatory agencies may include the Federal Communications Commission, Food and Drug Administration (FDA), Federal Trade Commission (FTC), Office for the Civil Rights of the Department of Health and Human Services (HHS), and National Institute of Standards and Technology, each with a unique role. These agencies can help create standards for mHealth technology, authorize carriers for access and transition of information from devices connected to networks, ensure appropriate use of information by health providers, and regulate advertising related to the app use (Yang and Silverman 2014). The FDA has expressed it will focus on the subset of mobile apps that are intended to be used as an accessory to a regulated medical device or transform a mobile platform into a medical device using attachments, display screens, sensors, and other methods (U. S. Department of Health and Human Services 2015). In essence, as part of its risk-based approach to cover apps with the potential to cause the most harm, it will focus its attention to mHealth apps that transform consumer devices into medical devices but leave the large majority of other health apps unregulated.

Aside from unclear regulatory domains, jurisdiction, and liability also poses an issue for mHealth technology. When providers use health apps to communicate with other providers in different locations, issues regarding the cross-jurisdictional practice of medicine may arise. As most medical silencing requirements are state specific, there exist over 50 different sets of requirements. Telemedicine has led the way in helping to clarify cross-jurisdictional practice (State Telehealth Laws and Medicaid Program Policies: A Comprehensive Scan of the 50 States and District of Columbia 2016), but more clarity will be needed as data from mHealth devices

is potentially transmitted and acted upon across state lines. Additionally, without a clear standard of care with respect to the use of mHealth technologies, coverage of malpractice laws come into question. Traditional malpractice liability is based on a physician-patient relationship with direct contact and care. Given mHealth can be used to capture and monitor health data using a patient's own device or application, it is unclear what a physician's liability would be if patient injury resulted from faulty or inaccurate information from the patient's device. As such malpractice in the setting of use of mobile health technologies remains an open question.

Advancements in mobile health technologies promise to make healthcare more accessible and to more effectively engage patients in their medical care, strengthening the patient-provider relationship. Applying a big data approach with thoughtful analytics to the massive amounts of data generated, captured, and transmitted on mobile devices can not only supply providers with additional information, but also present that information in such a way to enhance provider workflow. Given the rapid growth of mobile health technologies and their increasing integration into health records and use in clinical decision-making, more attention will be needed to clarify privacy, security, and regulatory concerns.

### ***16.1.12 Patient Education and Partnering with Patients***

The meeting of a physician and patient is the start of a therapeutic relationship, but much of the work that happens to improve health occurs beyond the appointment time. Whether a physician meets a patient in the clinic, hospital, nursing home, or any other care setting, for the physician's recommendations to be successfully implemented by the patient there has to be a trusting partnership. Physicians are the experts in medicine and patients are the experts in their own lives, so both parties must be engaged for meaningful changes to occur.

Much research has been done in the areas of patient engagement, activation, and the patient-physician partnership. Patient engagement includes both the patient activation to participate in care, and the resulting positive health behaviors from this motivation (Hibbard and Greene 2013). Patient engagement has been linked to participating in more preventative activities including health screenings and immunizations, adherence to a healthy diet, and regular physical exercise. Patients who are very activated are two or more times more likely than their less activated counterparts to prepare for doctor visits, look for health information, and know about the treatment of their conditions (Hibbard and Greene 2013; Fowles et al. 2009; Hibbard 2008). Patient engagement in the therapeutic process has been studied and encouraged in specific areas such as chronic condition management and adverse event reporting, but the degree to which patients are able to be engaged in their healthcare in the current age of social media and Google is unprecedented.

As electronic health records are creating accessible mobile platforms, all patients, not just those with chronic conditions, are able to engage in their healthcare on a basic level to schedule appointments, ask questions to medical staff, and initiate

medication refills online. Despite the fact that not all patients have the technical knowledge, health awareness, and engagement to interact with their physicians in this manner between appointments, this is still becoming an increasingly common way that patients and their caregivers are getting engaged with their medical care.

Understanding the role that patient engagement plays in the physician-patient interaction and disease treatment and management sets the stage for understanding the new ways patients can now engage in their healthcare through online access to medical information. A 2013 study in the journal *Pediatrics* sums up what many physicians perceive in the current age of medical care, that “Dr Google is, for many Americans, a de facto second opinion” (Fox 2013). The Pew Research Center has studied health information online since 2000. The latest national survey in 2014 showed that 72% of adults search online for health issues, mostly for specific conditions and their treatments. Meanwhile, 26% say they have used online resources to learn about other people’s health experiences in the past 12 months, and 16% have found others with the same health concerns online in the past year (Fox 2014). Caregivers and those with chronic health conditions are most likely to use the internet for health information. Health professionals are still the main source of health information in the US but online information, especially shared by peers, significantly supplements the information that clinicians provide (Fox 2014). Many survey respondents report that clinicians are their source of information for technical questions about disease management, but nonprofessionals are better sources of emotional support (Fox 2013).

### ***16.1.13 Developing Online Health Communities Through Social Media: Creating Data that Fuels Research***

PatientsLikeMe Inc. (Cambridge, MA, USA) and the rare disease community exemplify the use of social media to both connect people with similar health conditions, and provide real-time data feedback to healthcare providers, healthcare systems, pharmaceutical companies and insurance companies. This feedback is unique in that it collects large amounts of patient-level data to help in the development of new care plans for specific patient populations.

PatientsLikeMe allows users to input and aggregate details on symptoms, treatments, medications, and side effects of various illness, connect to support groups dedicated to challenging health conditions, and creates a platform that can potentially set the stage for clinical trials and medical product development (Sarasoehn-Kahn 2008; Wicks et al. 2010). Patients who choose to enter their personal health data on the site are actively choosing to use their personal health information in a way that is different than the traditional method of encapsulating health information in the electronic health record, to be seen and used privately between the physician and patient. For example, those with Amyotrophic Lateral Sclerosis (ALS) may choose to share personal information about their symptoms, treatments, and outcomes with the ALS community within PatientsLikeMe. These community-level symptoms and treatments are aggregated and displayed so that users can discuss the data within the site’s forums, messages, and comments



sections. A study of PatientsLikeMe's data sampled 123 comments (2% of the total commentary posted) and noted that group members sought out answers to particular questions guided by this data, offered personal advice to those who could benefit, and made relationships based on similar concerns and issues. This study illustrated that individual patients who shared their personal health data benefitted by participating in conversations that may help with self-management of their disease. (Frost and Massagli 2008). In addition to individual patients benefitting from their usage of the site, PatientsLikeMe allows pharmaceutical companies to partner with patients to design clinical trials and research studies (PatientsLikeMe Services n.d.). Since 2007, PatientsLikeMe has achieved many milestones in patient-centered and patient-directed research. For example, in 2007 a study on excessive yawning in ALS patients was conducted in which the symptom was listed on the ALS page and each user had to rate the severity of the symptom experienced. This quickly created a method to evaluate the symptom in the context of the person's medications and disease course and ultimately data from the PatientsLikeMe users helped to identify that the excessive yawning was more likely a symptom of emotional lability associated with the disease state rather than a drug side effect, or a side-effect of respiratory issues with ALS. Soon after, it was found that the impact of the research extended beyond its clinical scope. While the potential physical pain of yawning in ALS patients may have been the impetus to study this issue, based on discussions on the site it came to light that people with ALS had lost friends due to the misinterpretation that the yawning represented lack of interest and was a sign of rudeness. With this study, patients/families and healthcare professionals are now better able to understand and be more sympathetic to this symptom, physicians can warn patients of this symptom, and researchers have a greater impetus to find a treatment for the emotional lability likely causing yawning (Wicks 2007). The example highlights the invaluable nature of user input in helping to guide research in a more patient-centered manner.

The rare disease community has also been a remarkable testament to the power of online communities for sharing healthcare data and furthering medical practice. When new findings are published as case reports in academic journals, the process relies on clinicians in various parts of the world to see those articles and recall the specific symptomatology at the time needed to make a diagnoses - a process that can take many years to diagnose individuals with the same rare diseases. With common online platforms like Facebook, Twitter, and blogs, parents of kids with rare diseases and people who have rare conditions are taking matters into their own hands to find and share more information. Patients and families are turning to social platforms to promote greater collaboration between patients, caregivers, and healthcare industry professionals.

Bertrand Might is a child with the first known case of NGLY-1 deficiency, a very rare illness that was not diagnosed until he was nearly 4 years old (Might and Wilsey 2014). His journey to diagnosis is extraordinary because of the hard work of his parents, a couple who epitomize a new perspective and the shift occurring in the work of clinical diagnosis. After years of moving from one genetic specialist to the next looking for an explanation for their son's condition, researchers at Duke



finally gave the family an uncertain diagnosis about a rare new enzyme deficiency. Bertrand's parents blogged and documented his journey including evaluations, symptoms, visits to specialists, wrong diagnoses, and ultimately the gene mutations leading to his condition. A second patient with Bertrand's condition was discovered after a clinician came across this blog and realized that the two children likely shared the same enzyme deficiency. Yet another patient was identified when parents on another continent came across the blog after searching for similar symptoms and were motivated to have sequencing of the NGLY-1 mutation performed on their child. Now, 14 children from around the world have been diagnosed with this deficiency (Might and Wilsey 2014). Compared to how diseases have traditionally been discovered and disseminated through medical literature, this innovative new way of using social media to fuel recognition of symptoms and prompt genetic data analysis is quite rapid, often leading to a more timely diagnosis for those with extremely rare conditions. The implications of this phenomenon extend beyond the rapid diagnosis of rare diseases. Often rare-disease communities prompt pharmaceutical companies to consider researching and developing treatment options, support creating patient registries and push for clinical trials, and attend FDA meetings to advocate for the approval of new therapeutic options (Robinson 2016).

### ***16.1.14 Translating Complex Medical Data into Patient-Friendly Formats***

The value of data from online sources comes from the aggregation of opinions on a specific topic, such that the sum of different user input is more powerful than a single person's comment. A current example of a site making big data useable for patients and physicians alike is iodine.com.

### ***16.1.15 Beyond the Package Insert: Iodine.com***

Since 1968, the US Food and Drug Administration (FDA) has required that certain prescription medications contain package inserts that consist of usable consumer medication information (CMI). The suggested CMI includes the name and brand name of the medication, use, contraindications, how to take the medication, side effects, and gene years, the FDA's recommendations have supported changes to the package insert (Food and Drug Administration 2006).

Expanding upon this concept, Iodine.com (San Francisco, CA, USA) was founded in 2013 and currently offers free medication information for more than 1000+ drugs. This information is compiled from FDA drug side effect data and augmented with user input on drug side effects from Google Consumer Surveys (GCS) (Iodine 2014), creating a experience that has been called the "yelp of medicine" (Sifferlin 2014). Ultimately, these medication-related experiences become part of a growing database that can guide new insights into how drugs work

in the real world. There are many other sources of drug information on the internet for consumers including [drugs.com](http://drugs.com), [medlineplus.gov](http://medlineplus.gov), [rxlist.com](http://rxlist.com), [webmd.com](http://webmd.com), and [mayoclinic.org](http://mayoclinic.org), but these sites lack the peer-to-peer recommendations that [Iodine.com](http://Iodine.com) provides. Medication side effect profiles provided for physicians and that are listed in the drug package inserts are notoriously long, listing all side effects regardless of incidence frequency from less than a 1% chance of occurring to frequent side effects. This type of information is hard for consumers to interpret, and impossible for physicians to memorize, so often patients are prescribed medications with minimal guidance: only important contraindications such as “don’t take with alcohol” or “take on an empty stomach” are communicated. Most physicians also do not have first-hand experience with these drugs, and so are unfamiliar with which side effects truly occur most often.

[Iodine.com](http://Iodine.com) helps to fill the void that is common in traditional prescribing practices by adding personalization. The data for [iodine.com](http://iodine.com) is collected from a variety of sources including traditional medical research literature and pharmaceutical product labels, center for Medicare and Medicaid Services (CMS) data, non-research sources of data like insurance claims formularies, pharmacist reports, patient reported data from more than 100,000 Americans on Google Consumer Surveys, and social data sharing health experiences extracted from many other sites ([Iodine Data n.d.](#)).

[Iodine](http://Iodine.com) users also complete online reviews of their medications, and this community-generated content contributes to the data on the site. The data from these sources covers 1000+ medications, and offers data subclassified by age, gender, and medical condition. This subcategorization allows users to see how similar populations experience and feel about different medications. This unique aspect, the user experience, is a valuable addition to the information provided by medical research literature primarily because medical research often has limited generalizability based on the study design and exclusion criteria. Results of medical studies can often only be extrapolated to populations similar to the study populations, which often exclude older adults and those with complex medical conditions limiting generalizability. These exclusions exist to make study results easier to interpret, but a consequence of such stringent exclusion criteria is that a large percentage of complex patients (precisely the ones for which clinicians most need guidance) do not have a lot of evidence-based management guidelines. If large amounts of data on patient experience can be captured from this complex population of older adults, this can help guide clinical decision making ([Bushey 2015](#)).

Despite the potential for significantly expanding the database of medication use and effects, there is a one major limitation of [Iodine.com](http://Iodine.com): validity. There is no mechanism for confirming the self-reported data within the system. The integrity of the reviews is potentially quite variable. In an attempt to combat this, the company manually reviews all of the user input. [Iodine.com](http://Iodine.com) reports that the site rarely receives reviews that are not legitimate, but if detected, these reviews are removed. They also check data trends to see if the patterns of reviews are consistent with known population side effects ([Bushey 2015](#)).

### ***16.1.16 Data Inspires the Advent of New Models of Medical Care Delivery***

The life of a physician is typically a delicate balance between the time demands of seeing patients, extensive medical documentation of encounters, and billing and coding. Through all of this, certain interventions such as behavioral health discussions rarely occur due to low provider knowledge and confidence, insufficient support services, and little feedback from patients that behavior interventions are needed and are effective (Mann et al. 2014). There are several companies that have recognized the need to identify and engage high risk patients through technological solutions with the goal to yield behavioral change. Two such companies are Ginger.i.o and Omada health.

Ginger.io (San Francisco, CA, USA) offers a mobile application for patients with various mental health conditions to have a personalized health coach. It allows users to communicate through text or live video sessions with a licensed therapist specializing in anxiety and depression, and has 24/7 access to self care tools through its app. The company combines mobile technology with health coaches, licensed therapists, consulting psychiatrists, and medical providers to create an interface that is technology based, but with a human element. The app collects both active data through regular mood surveys, and passive data through mobile phone data on calling, texting, location, and movement. This passive data can generate a norm for the user's regular patterns. Once regular patterns are known, changes in communication and movement can help predict depression or changes in mental health (Feldman 2014). The active and passive data in addition to in-app activities to build skills around managing mood changes are synthesized into personalized reports that can be shared with the person's doctor and the Ginger.i.o care team. Accessible mental health care is challenging in the US for a variety of reasons-availability of providers, cost and cost sharing limitations by insurance companies, distance to care, and availability of appointments (Cunningham 2009). This type of mobile platform can help narrow this gap and provide timely and convenient behavioral care when needed through mobile devices (Ginger.io Evidence N.d.). Ginger.io combines its data with behavioral health research data from the National Institutes of Health and other sources to help provide insights from the aggregate data. One particularly interesting insight is that a lack of movement from a user could signal that a patient feels physically ill and irregular sleep patterns may precede an anxiety attack. (Kayyali et al. 2013)

Omada Health (San Francisco, CA, USA) is an innovative company that offers "digital therapeutics," which are evidence-based behavioral treatment programs that are cost-effective and potentially more accessible than traditional programs. Their 16-week program, *Prevent*, offers a targeted intervention to individuals who are at high risk for chronic illnesses such as diabetes and heart disease. Each participant is paired with a personal health coach and online peer group for regular feedback and support. This program includes weekly lessons about nutrition, exercise, and psychological barriers. In 2016, a study was published examining long term clinical

outcomes of the Prevent Pilot and the effects of the program on body weight and Hemoglobin A1C, a marker of blood sugar control in diabetics. The 187 pre-diabetic participants who completed the four core lessons achieved an average of 5.0% and 4.8% weight loss at 16 weeks and 12 months, respectively and had some reduction in their A1C level at final measurement (Sepah 2015).

The type of behavioral change program developed by Omada health would require extensive staff and resources if replicated in person, rather than administered online. The benefits of these online interventions include increased access to care, convenience to patients by using mobile health delivery systems and avoiding travel time, and increased patient engagement in their healthcare. Potential limitations of these types of online behavioral intervention technology (BIT) programs are technological barriers, lack of engagement due to the design of the program, and issues with translating the program into actionable behavior change. Ideally, human support should complement BIT use, so that together all potential barriers are addressed and individuals can gain the maximum potential benefits from these behavioral change programs (Schueller et al. 2016).

## **16.2 Part II: Physician Uses for Big Data in Clinical Care**

### ***16.2.1 The Role of Big Data in Transforming How Clinicians Make Decisions***

#### **16.2.1.1 Imagine the Following Scenario**

An elderly man, Mr. Williams, develops a fever, back pain, and nausea, and is admitted to the local hospital where he is diagnosed with a kidney infection. He tells his physician that he has several other medical conditions, such as heart disease and diabetes, and was hospitalized several times in the past year for different infections, although they were at another hospital where medical records are not available. The physician chooses to start him on ciprofloxacin, which is an antibiotic that she typically uses to treat kidney infections. She also faxes a form to the other hospital to request records, although the request would not be processed until the following day.

In the middle of the night, the physician receives a page from the nurse saying that Mr. Williams' heart rate is elevated. She walks over to his room to examine him, and notices that he is a bit anxious about being in the hospital. This was actually the fifth time she was woken up that night; the last four times were all for patients who were anxious and needed something to help them calm down. She takes a minute to scan Mr. Williams' chart and does not notice anything else that looked worrisome, so instructs the nurse to give him an anti-anxiety medication and goes back to sleep.

The following morning, Mr. Williams develops a high fever and becomes confused. The physician, now more concerned, looks through his chart and notices that his heart rate had continued to increase throughout the night. She diagnoses him with sepsis and immediately starts intravenous fluids and switches his antibiotic to

piperacillin/tazobactam, which is the antibiotic that she usually uses if ciprofloxacin is not working. However, he continues to worsen throughout the day and is later transferred to the intensive care unit. That afternoon, the medical records from the other hospital are faxed over, which consist of 50 printed pages of progress notes, discharge summaries, medication lists, and lab reports. Although the physician's shift had ended two hours ago, she spends the time looking through the faxed charts because she is particularly worried about Mr. Williams (she had taken care of a patient with a kidney infection who ended up dying in the hospital just two weeks ago). After pouring through the pages of records, she finally notices a line of text in a lab report that is a critical piece of information: two months ago at the other hospital, Mr. Williams had developed another kidney infection and his urine at the time grew out a bacteria that was resistant to both ciprofloxacin and piperacillin/tazobactam. She remembers that meropenem would be the antibiotic of choice in this situation, but takes the extra time to look it up in her online reference since she has not used that medicine in over a year. At 5 pm, almost 24 h after Mr. Williams was admitted to the hospital, the physician starts him on the correct antibiotic. He eventually recovers, but was weakened because of the prolonged hospital stay and had to be discharged to a nursing home.

The above hypothetical scenario is not uncommon in modern hospitals. Although the resources needed to treat Mr. Williams' kidney infection were readily available and his physician did her best to care for him, the limitations in the process by which medical decisions are made in healthcare led to missed opportunities that may have improved the care that he received in the hospital. He could have received the correct antibiotic much earlier, and his sepsis could have been recognized and acted upon sooner. Clinical decisions are often made in situations of uncertainty with incomplete information and sometimes inconsistent levels of expertise. As demonstrated in the scenario with Mr. Williams, many decisions are heavily influenced by the experiences of the individual physician, which can lead to high variability in cost and quality of care (Institute of Medicine 2012). Existing clinical knowledge is often inconsistently applied, with compliance with evidence-based guidelines ranging from 20%–80%. Additionally, most clinical decisions including common diseases such as heart attacks, lack strong evidence from clinical research studies (Institute of Medicine 2012) with only ~11% of clinical practice guideline recommendations supported by high quality evidence (Tricoci et al. 2009). This section will describe the potential for big data to address these gaps and transform how decisions are made in healthcare.

The methods by which physicians are currently trained to make medical decisions are not very different from those utilized one hundred years ago. Physicians undergo a long and arduous training process that heavily relies upon developing the individual's knowledge base and experience. The amount of data available for today's physicians to process, however, far exceeds the capacity of any individual and is increasing at a rapid rate. Advances in medical research have not only given us many more diagnostic and treatment options for long-standing diseases, but have also defined new diagnoses that add to the already expansive medical vocabulary. As our understanding of disease grows more granular, the management of patients is

becoming increasingly complex. A physician in the intensive care unit will manage a range of 180 activities per patient per day (Institute of Medicine 2012). This rapidly accumulating knowledge base has allowed us to reimagine what medicine can accomplish, but also means that today's physicians are required to know and do more than ever before to deliver the standards of care that we now expect from modern healthcare systems. The cognitive tools that physicians are given, however, have not advanced at the same rate, thus creating a critical need for newer tools to help physicians make clinical decisions.

The promise of big data to address this need is driven by the increasing wealth of clinical data that is stored in the electronic medical record (EMR). Information such as patient histories, laboratory and imaging results, and medications are all stored electronically at the point of care. This critical mass of data allows for the development of clinical decision support computational tools that interact with the EMR to support physicians in making clinical decisions. The development of clinical decision support is an evolving field with different computational models that include earlier systems using probabilistic and rule-based models and newer data driven approaches that more effectively harness the power of big data.

### ***16.2.2 Probabilistic Systems***

The diagnosis of disease follows a Bayesian model. When approaching a new patient, a trained physician will come up with a list of potential diagnoses in order of likelihood, and adjust the probabilities of those diagnoses being present based on new information from the examination and diagnostic tests. A patient coming to the emergency room with chest pain could have anything from a heart attack to a muscle strain, but additional relevant information such as the character of the pain and the presence of elevated serum levels of cardiac enzymes will increase the posterior probability of the diagnosis being a heart attack and decrease that of the diagnosis being a muscle strain. Physicians use this mental model to diagnose diseases, but will often be affected by psychological biases. In the case of Mr. Williams, his physician mistakenly assigned a higher posterior probability to the diagnosis of anxiety as the cause of his elevated heart rate overnight, likely because she had just seen several other patients with anxiety. This phenomenon, known as the availability heuristic, is a common reason for misdiagnosis and physicians become increasingly susceptible to it when exhausted with data and tasks that exceed what can be processed by an individual person.

Probabilistic clinical decision support systems use computers to simulate the Bayesian model with which physicians are trained to think. For a presenting symptom, a computer can start by considering the prior probabilities of the differential diagnosis, which is typically based on the known prevalence of those conditions, and modify the posterior probabilities in a sequential approach based on new inputted data, such as symptoms and diagnostic test results (Shortliffe and Cimino 2014). These clinical decision support systems require a knowledge base of

estimated conditional probabilities of a set of diseases for the pieces of data that are entered. VisualDx (Rochester, NY, USA) is an example of a commercially available probabilistic clinical decision support system that is designed to help users create a differential diagnosis at the point of care that is not affected by human error such as recall bias. A user can enter combinations of clinical data such as symptoms and laboratory results to generate lists of likely diagnoses that are ordered by their posterior probabilities.

### ***16.2.3 Rule-Based Approaches***

Rule-based clinical decision support systems rely on encoded concepts in a knowledge base that are derived from content experts to simulate recommendations that experts might provide. The knowledge base can include probabilistic relationships such as those between symptoms and diseases, and medications and side effects. For example, a commonly used rule-based clinical support system is the drug interaction alert system that is built into many modern EMRs. The system alerts the physician if a new medication ordered has an adverse interaction with an existing medication that the patient is taking. When the medication is ordered in the EMR, the action is then analyzed against a previously encoded knowledge base of all known medication interactions, which then triggers an alert. Knowledge bases can be created for specific diseases to create tools such as early disease detection systems. For example, many hospitals now use EMR embedded clinical decision support tools to detect severe sepsis, which is a life-threatening physiologic state that patients can develop when they have an infection. These alerts are powered by algorithms designed based on probabilistic rules derived from sepsis treatment guidelines (Rolnick et al. 2016). Such a clinical decision support system may have been able to help Mr. Williams' physician recognize earlier that he had sepsis based on clinical data in the EMR that the physician did not otherwise notice, or could have initiated an alert when the physician ordered the first antibiotic that his previous urine culture grew a resistant antibiotic.

### ***16.2.4 Data Driven Approaches***

Rule-based and probabilistic models, which comprise the majority of existing clinical decision support systems that are in use, rely on manually curated knowledge bases that are applied to clinical data in a top down approach. They act on, rather than use the data in the EMR to generate insights into clinical decisions. The paradigm of clinical decision support is now evolving towards a data driven approach, which employ machine learning techniques to mine EMR data for new knowledge that can guide clinical decisions. Rather than relying on a pre-formed knowledge base, a data driven approach seeks to "let the data speak for itself"



by extracting patterns and insights from the data generated by everyday clinical practice. The most direct approach is to attempt discovery of new knowledge that can guide clinical decisions. This approach is particularly powerful for the majority of clinical decisions that do not have an existing evidence base in the form of clinical trials or guidelines, but for which many “natural experiments” occur in regular practice. Efforts are underway to develop systems that would allow physicians to generate real-time, personalized comparative effectiveness data for individual patients using aggregate EMR data (Longhurst et al. 2014). For example, there may have been thousands of other patients who share Mr. Williams’ specific clinical history who were treated at the same hospital. His physician could query the EMR to find out how this cohort of patients responded to certain treatments in order to choose the most effective treatment for him. Reliable conclusions can be challenging due to confounding by indication, but can somewhat be mitigated by causal inference methods that risk adjust for different clinical factors (e.g., propensity score matching) as has been demonstrated through established methods in retrospective observational research.

Another example of an emerging data driven approach is known as collaborative filtering. Traversing the hierarchy of medical evidence, we first look to randomized controlled trials to guide our medical decision making, followed by observational studies, before accepting consensus expert opinion, or finally our own local expert (consultant) opinions and individual anecdotal experience. With only ~11% of clinical practice guideline recommendations backed by high quality evidence and only about a quarter of real-world patients even fitting the profile of randomized controlled trial inclusion criteria, it should not be surprising that the majority of medical decisions we have to make on a daily basis require descending the entire hierarchy to individual experience and opinion. For a practicing clinician, the established norm is to consult with other individual local experts for advice. The advent of the EMR, however, enables a powerful new possibility where we can look to, not just the opinion, but the explicit actions of *thousands* of physicians taking care of similar patients. Right or wrong, these practice patterns in the collective community reflect the real world standard of care each individual is judged against. More so, these may reflect latent wisdom in the crowd, with clinical practices refined through years of hard won experience, but which never before had a fluid mechanism to disseminate through conventional social and publication channels. Such an approach can represent an entirely new way of generating and disseminating clinical practice knowledge and experience, owing heavily to methodology established in product recommender systems for companies such as Netflix and Amazon. Active research is underway to help discern how such approaches can separate the wisdom of the crowd from the tyranny of the mob, and the potential impacts of integrating such dynamic information into a physician’s point of care decision making process (Chen et al. 2016). When the computer systems are trained to recognize established standards of care through readily available clinical data, they will be able to seamlessly anticipate clinical needs even



without being asked. This will translate endpoint clinical big data into a reproducible and executable form of expertise and, deploying this right at the point-of-care, can close the loop of a continuously learning health system.

### ***16.2.5 Challenges and Areas of Exploration***

Significant challenges remain in the development and adoption of clinical decision support. Although the EMR stores clinical data electronically, much of the data is not in a format that is easily readable by computers. Inherently structured data such as medication lists and laboratory values are often the sources for data used in existing rule-based clinical decision support systems. The promise of data driven machine learning approaches to clinical decision support, however, require the use of data in the unstructured, free text narratives that comprise the majority of valuable, actionable EMR data generated by clinicians. The question of how to structure large amounts of clinical data into reliable variables that can then be used by computational tools remains one of the “grand challenges” of clinical decision support (Sittig et al. 2008). Natural language processing, which is a technique that has been used in other applied fields of computer science, is being explored as a way to translate clinician generated text into encoded data (Liao et al. 2015). How this data can then be organized into meaningful groups, or phenotypes, that can be analyzed is an area of active research. For example, EMR phenotypes can be used to create electronic cohorts of patients around specific disease states, either at the point of care to generate real time comparative effectiveness data for clinical decision support, or for the creation of electronic cohorts that can be used for clinical research (Xu et al. 2015).

The infrastructure for how clinical data is stored and shared is also in need of change to accommodate a data driven healthcare system. In our example with Mr. Williams, a critical piece of clinical data, the urine culture results, was actually located in another hospital and had to be faxed over as printed text to be read by the physician. Any clinical decision support system would be limited by the amount of clinical data available to analyze. Currently, electronic clinical data is stored throughout disparate EMR systems that are owned by individual healthcare delivery systems and often not shared electronically. Although the Health Information Technology and Clinical Health Act (HITECH), which was enacted as part of the American Recovery and Reinvestment Act of 2009, includes standards for EMR interoperability that envision a system where clinical data can be shared electronically among clinicians across the country in real time, issues such as privacy, misaligned financial incentives, and the lack of technological infrastructure continue to remain as barriers to adoption (Ball et al. 2016). Regional health information exchanges are beginning to have some success in allowing for data sharing among health systems, although the scope remains limited (Downing et al. 2016). Further, insights are needed to understand how to successfully scale implementation of clinical decision support systems into clinical enterprises. Issues such as physician

workflow integration, system usability, and alignment with financial incentives of healthcare delivery systems all need to be considered. Nevertheless, in spite of these challenges, the convergence of the need for improved healthcare quality, an unprecedented amount of available clinical data, and the rapid development of powerful analytical tools is pushing the healthcare system to a tipping point and into an era of big data.

### ***16.2.6 Using Big Data to Improve Treatment Options***

One of the greatest promises of big data as it relates to medical care is in precision or personalized medicine. These terms are often used interchangeably, with precision medicine emerging more recently and persisting as the preferred term to describe the concept of taking individual variability into account in creating prevention and treatment plans (National Academies Press 2011). Precision medicine has existed for at least a century with the most prominent example seen in blood typing to more safely guide blood transfusions (Collins and Varmus 2015). Additionally, complete sequencing of the human genome at the beginning of this century has led to a wealth of data towards the better understanding of disease states, developmental variability, and human interaction with pathogens (Lander et al. 2001). More recently, precision medicine has come to define a framework to combine huge databases of patient health data with OMICS, primarily genomics but also proteomics, metabolomics, and so on, to facilitate clinical decision-making that is “predictive, personalized, preventive and participatory” (P4 Medicine) (Hood and Flores 2012). The overarching hope for precision medicine is to be able to select therapies for predictable and optimal responses, and identify potential side effects to therapies based on a patient’s genetic makeup and individual characteristics.

In his 2015 State of the Union Address, President Obama announced details about the Precision Medicine Initiative (PMI), a \$215 million research effort intended to be at the vanguard of precision medicine (The White House 2015). Funds have been invested into the National Institutes of Health (NIH), National Cancer Institute (NCI), Food and Drug Administration (FDA) and Office of the National Coordinator for Health Information Technology (ONC) to pioneer a new model of patient data-powered research intended to accelerate the pace of biomedical discoveries. The PMI’s initial focus is on cancer treatment, but long-term goals emphasize preventive care and disease management.

The Precision Medicine Initiative encourages collaboration between the public and private sectors to accelerate biomedical discoveries using technology to analyze large health datasets alongside advances in genomics. This ambitious goal, while simple in framing, in practicality will require an immense amount of oversight and regulation alongside the actual research components to ensure the safety, privacy, and security of data used. As such, the NIH is tasked with building a voluntary national research cohort of over one million Americans to collect a broad collection of data including medical records, genetic profiles, microbes in and on the

body, environmental and lifestyle data, patient-generated information, and personal device and sensor data. The ONC is specifically tasked with developing standards for interoperability, privacy, and secure data exchange across systems. Additional provisions within the PMI have been included to protect privacy and address other legal and technical issues. In sum, the Precision Medicine Initiative will help support and also make practical the transition into the era of precision medicine.

Currently, the best examples of precision medicine can be seen in the field of oncology, where patients increasingly undergo extensive molecular and genetic testing so that physicians can choose treatments best suited to improve survival and reduce side effects (Kummar et al. 2015). One particularly encouraging advancement in targeted oncology has been with the treatment of metastatic melanoma, which prior to 2011 was thought of as a rapidly fatal condition with a prognosis usually under one year (Jang and Atkins 2013). Studies of melanoma biology and immunology revealed that almost 50% of melanomas harbor mutations in BRAF, mainly at codon 600. Ipilimumab and vemurafenib, two BRAF Val600 selective inhibitors, demonstrated significant tumor response with improved progression-free survival. Despite this promising initial response, patients often suffered disease progression at a median of 5–7 months due to multiple resistance mechanisms within the tumors. It was then discovered that some patients with BRAF Val600 mutations were able to obtain more durable responses with the addition of certain immunotherapies like high-dose interleukin 2. While further studies are still needed to ascertain the extent of downstream mutations and optimal combination therapies for greater survival (Robert et al. 2015), targeted therapies like the BRAF V600 selective inhibitors remain the goal in increasing precision medicine in oncologic care.

While targeted therapies are being designed to treat tumors, testing for somatic germ line mutations is also being employed to assess risk and stratify management decisions for certain types of cancer. The most well known mutations, BRCA1 and BRCA2, can be tested in breast cancer patients to identify optimal surgical, radiotherapeutic, and drug choices for patients (Trainer et al. 2010). When tested in patients without cancer, identification of these mutations can significantly alter an individual's knowledge of their risk profile and affect downstream management of cancer screening and prevention (U.S. Preventative Services Task Force 2014). Studies are currently being undertaken to examine the potential benefits versus harms of BRCA genetic testing at the population level (Gabai-Kapara et al. 2014).

With respect to the management of chronic diseases, precision medicine has already yielded some concrete improvements in patient health. In 2014, direct and indirect mental health expenditures exceeded those of any organic health condition including cardiovascular disease and diabetes (Agency for Healthcare Research and Quality 2014). Treatment of mental health conditions, particularly refractory conditions, can be exceedingly challenging and is often based on trial and error with various medications. The field of pharmacogenomics has developed to identify genetic differences in the pharmacokinetic and pharmacodynamic profiles of individuals, and stratify their likely responses to different drugs. This can not only lead to more effective drug use but also mitigate adverse effects and potentially deliver cost savings to the health care system. The GeneSight Psychotropic test

was developed to provide clinicians with a composite phenotype for each patient applied to the known pharmacology of certain psychiatric medications (Winner et al. 2013). A recent study showed that pharmacogenomic-guided treatment with GeneSight doubled the likelihood of response for patients with treatment resistant depression, and identified patients with severe gene-drug interactions enabling them to be switched to genetically preferable medication regimens. A later study showed that use of GeneSight in medication selection resulted in patient exposure to fewer medications with greater adherence and an overall decrease in annual prescription costs (Winner et al. 2015). These examples exemplify the improvements in individual clinical care that can come about as a result of the collaboration between informatics, research, and clinical medicine.

Harnessing big data can also transform the way physicians apply research to their daily practice. Clinical research has traditionally relied on time-consuming acquisition of data and human driven analysis to conduct studies. However, this exposes them to only a fraction of published data. Furthermore, clinically relevant research then is distributed through published journals that often take weeks to months at minimum to disseminate to clinicians. True changes in practice patterns take 17 years on average to then be fully implemented (Morris et al. 2011). Well designed EHR interfaces could one day intelligently match publications to clinical situations, significantly augmenting physicians' ability to apply new published evidence at the point of medical decision making. Meanwhile real-time aggregation of data at smaller levels such as a city or county can lead to timely public health interventions.

Despite the huge promise of big data in precision medicine, some of the largest issues to be addressed include variability and reliability of information (Panahiazar et al. 2014). Health records can potentially provide fragmentary information if the health record is not complete and there is not systematic quality control for data elements gathered to ensure data accuracy. Additionally, with such a large amount of information being gathered and inputted from various sources, issues with incongruent formatting and lack of interoperability exist. Several different strategies are being employed to address this, from platforms such as the Internet of Things to loop in data from devices to semantic web technologies to make information interpretable for search and query and integration. Data must additionally be standardized and processed to ensure high quality input.

The limits of what data may be captured, in particular, has potential for controversy. For example, healthcare organizations that wish to reduce healthcare costs may wish to see patients' credit card data to assess their risk for disease based on things such as travel history or food, alcohol, and tobacco purchases. However, further discussion is needed to determine where the boundaries of this big data "creep" may lie. Meanwhile, as the pure volume of data available exceeds what providers can process in a reasonable timeframe, physicians will increasingly rely on big data analytics to augment their clinical reasoning. Already, robust decision support systems are being developed with IBM's Watson to help clinicians match advanced molecular therapies with an individual's tumor (Parodi et al. 2016). Future providers and patients will need to determine the role of increasingly powerful clinical decision support without inappropriately making it a clinical decision maker.

Moving away from traditional fee-for-service payment models and historic medical disease management techniques is necessary to facilitate innovation and incorporation of big data into medical care. In many current health systems, payers and providers are incentivized differently and care is influenced by what will be covered by insurance and what will not, rather than all stakeholders focusing on high quality effective care (Kayyali et al. 2013). Political and financial pressures are already improving the landscape for standardized data. In some cases, this has taken the form of interoperability and led to immediate clinical benefit. Kaiser Permanente has fully integrated their electronic EHR HealthConnect, to allow information to cross over all health care settings, inpatient and outpatient, and across all facilities. The integrated system has shown improved outcomes in heart disease management and has created an estimated \$1 billion savings from reduced office visits and redundant diagnostic tests (Kayyali et al. 2013). Meanwhile, the Health Observational Healthcare Sciences and Informatics collaborative has aggregated over 680 million records across Australia and many countries in Asia and has been working on the early identification of adverse drug reactions (Duke 2015). In other cases, larger institutions are intentionally curating more robust databases to drive advanced analytics. Both the VA and the NIH are creating databases with one million or more patients with a specific emphasis on combining large genomic datasets with clinical datasets (VA 2016, National Institutes of Health 2016). These two movements will ultimately create richer, standardized datasets to drive higher quality big data analytics.

This chapter is intended to be an overview of the applications of big data as they interface with a physician's practice. Currently, healthcare data is derived from a multitude of sources including those internal to a health system like electronic health records, computerized order entry systems, and data from devices and sensors. External sources including billing records from insurance companies, pharmacies, social media, and mobile and wearable consumer devices are becoming more and more prominent. The acquisition and aggregation of data from these new sources, ranging from standardized biometric data to highly variable patient captured data, adds both a wealth of potentially actionable health information as well as a layer of complexity related to systems interoperability, privacy and security concerns, and legal and regulatory challenges. Despite these challenges, promising examples of big data applications have emerged in both the public and private sector to transform the patient-physician relationship, expand knowledge of patient health factors, activate and engage patients in their health care, strengthen evidence-based physician decision-making, and accelerate research for more individualized patient care. Currently, much work is needed to create mechanisms to translate data and data-informed insights into useable data that can directly affect patient care and drive quality improvement in healthcare (Neff 2013). While the technical aspects of big data applications may be out of the scope of knowledge for many physicians, every aspect of medicine and health care will soon be influenced by the transformative potential of big data to achieve high quality, efficient, and effective patient-centered care.

## References

- Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., & Connolly, G. N. (2014). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, 21(2), 374–378. doi:10.1136/amiajnl-2013-002079.
- Agboola, S., Jethwani, K., Khateeb, K., Moore, S., & Kvedar, J. (2015). Heart failure remote monitoring: Evidence from the retrospective evaluation of a real-world remote monitoring program. *Journal of Medical Internet Research*, 17(4), e101. doi:10.2196/jmir.4417.
- Agency for Healthcare Research and Quality. (2014). *Total expenses and percent distribution for selected conditions by type of service*. Retrieved from [https://meps.ahrq.gov/data\\_stats/tables\\_compendia\\_hh\\_interactive.jsp?\\_SERVICE=MEPSSocket0&\\_PROGRAM=MEPSPGM.TC.SAS&File=HCFY2014&Table=HCFY2014%5FCNDXP%5FC&\\_Debug=](https://meps.ahrq.gov/data_stats/tables_compendia_hh_interactive.jsp?_SERVICE=MEPSSocket0&_PROGRAM=MEPSPGM.TC.SAS&File=HCFY2014&Table=HCFY2014%5FCNDXP%5FC&_Debug=).
- Ball, M., Weaver, C., Kim, G., & Kiel, J. (2016). Healthcare information management systems. *Journal of Information Systems Education*, 14. doi:10.1007/978-3-319-20765-0.
- Bushey, R. (2015, March 5). *Meet thomas goetz, The co-founder of iodine*. Retrieved October 31, 2016, from <http://www.dddmag.com/article/2015/03/meet-thomas-goetz-co-founder-iodine>.
- Caouette, H. (2015a). *eClinicalWorks announces integration between wearable devices and healow platform*. Retrieved from <https://www.eclinicalworks.com/pr-eclinicalworks-announces-integration-between-wearable-devices-and-healow-platform/>.
- Caouette, H. (2015b). *healow Announces Internet of Things Integration*. Retrieved from <https://www.eclinicalworks.com/pr-healow-announces-internet-of-things-integration-2/>.
- Chatterjee, P. (2015, October 09). *Delivering value by focusing on patient experience—AJMC*. Retrieved October 25, 2016, from <http://www.ajmc.com/journals/issue/2015/2015-vol21-n910/Delivering-Value-by-Focusing-on-Patient-Experience>.
- Chen, J. H., Goldstein, M. K., Asch, S. M., Mackey, L., & Altman, R. B. (2016). Predicting inpatient clinical order patterns with probabilistic topic models vs. conventional order sets. *Journal of the American Medical Informatics Association*. doi:10.1093/jamia/ocw136 [Epub ahead of print].
- Clark, N. M. (2003). Management of chronic disease by patients. *Annual Review of Public Health*, 24, 289–313. doi:10.1146/annurev.publhealth.24.100901.141021.
- Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine*, 372(9), 793–795. doi:10.1056/NEJMp1500523.
- Cunningham, P. J. (2009). Beyond parity: Primary care physicians' perspectives on access to mental health care. *Health Affairs*, 28(3), w490–w501.
- Darkins, A., Ryan, P., Kobb, R., Foster, L., Edmonson, E., Wakefield, B., & Lancaster, A. E. (2008). Care coordination/home telehealth: The systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. *Telemedicine and E-Health*, 14(10), 1118–1126. doi:10.1089/tmj.2008.0021.
- Dentzer, S. (2013). Rx for the “Blockbuster Drug” of patient engagement. *Health Affairs*, 32(2), 202–202. doi:10.1377/hlthaff.2013.0037.
- Downing, N. L., Adler-Milstein, J., Palma, J. P., Lane, S., Eisenberg, M., Sharp, C., et al. (2016). Health information exchange policies of 11 diverse health systems and the associated impact on volume of exchange. *Journal of the American Medical Informatics Association*, 34(13), 150–160. doi:10.1093/jamia/ocw063.
- Duke, J. (2015, June 14). *An open science approach to medical evidence generation*. Retrieved October 31, 2016, from [http://www.ohdsi.org/wp-content/uploads/2014/07/ARM-OHDSI\\_Duke.pdf](http://www.ohdsi.org/wp-content/uploads/2014/07/ARM-OHDSI_Duke.pdf).
- Feldman, B. (2014, January 30). *Using big data and game play to improve mental fitness*. Retrieved September 20, 2016, from <http://radar.oreilly.com/2014/01/using-big-data-and-game-play-to-improve-mental-fitness.html>.
- Fisher, M. (2014, June 12). *Apple's health app and healthkit and HIPAA*. Retrieved from <http://www.hitechanswers.net/apples-health-app-healthkit-hipaa/>.
- Food and Drug Administration. (2006). *Guidance: Useful written consumer medication information (CMI)*. Rockville, MD: US Food and Drug Administration.



- Fowles, J. B., Terry, P., Xi, M., Hibbard, J., Bloom, C. T., & Harvey, L. (2009). Measuring self-management of patients' and employees' health: Further validation of the Patient Activation Measure (PAM) based on its relation to employee characteristics. *Patient Education and Counseling*, 77(1), 116–122.
- Fox, S. (2013). After Dr Google: Peer-to-peer health care. *Pediatrics*, 131(Suppl. 4), S224–S225.
- Fox S. (2014). *The social life of health information*. Washington, DC: Pew Research Center's Internet and American Life Project. January 15, 2014, Available at <http://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/>.
- Freese, N. (2016, May 11). *It's high time for change—And health care data is paving the way*. Retrieved 25 October, 2016, from <https://medium.com/@hdpalooza/its-high-time-for-change-and-health-care-data-is-paving-the-way-cac1e54ee088#.wtgembknu>.
- Frost, J., & Massagli, M. (2008). Social uses of personal health information within PatientsLikeMe, an online patient community: What can happen when patients have access to one another's data. *Journal of Medical Internet Research*, 10(3), e15.
- Gabai-Kapara, E., Lahad, A., Kaufman, B., Friedman, E., Segev, S., Renbaum, P., et al. (2014). Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2. *Proceedings of the National Academy of Sciences, USA*, 111(39), 14205–14210. doi:10.1073/pnas.1415979111.
- Gartner. (2016, August 16). *Gartner's 2016 hype cycle for emerging technologies identifies three key trends that organizations must track to gain competitive advantage*. Retrieved October 10, 2016 from <http://www.gartner.com/newsroom/id/3412017>.
- Ginger.io Evidence (n.d.) Retrieved from <https://ginger.io/evidence/>.
- Grand Rounds (n.d.). Retrieved October 25, 2016, from <https://www.grandrounds.com>.
- Greene, J., & Hibbard, J. H. (2012). Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. *Journal of General Internal Medicine*, 27(5), 520–526.
- Healthgrades. (2016a). *Methodology: Mortality and complications outcomes*. Retrieved October 16, 2016, from <https://www.healthgrades.com/quality/methodology-mortality-and-complications-outcomes>.
- Healthgrades. (2016b, April 05). *Patient experience measures data source*. Retrieved October 25, 2016, from <https://www.healthgrades.com/ratings-and-awards/data-source-patient-experience>.
- Healthgrades (2016c). *Patient safety ratings and patient safety excellence award 2016 methodology*. Retrieved October 25, 2016, from <https://www.healthgrades.com/quality/2016-patient-safety-methodology>.
- Healthgrades (n.d.). *Healthgrades find a doctor doctor reviews hospital ... (n.d.)*. Retrieved 25 October, 2016, from <http://www.healthgrades.com/>.
- Hibbard, J. H. (2008). Using systematic measurement to target consumer activation strategies. *Medical Care Research and Review*, 66, 9s–27s.
- Hibbard, J. H., & Greene, J. (2013). What the evidence shows about patient activation: Better health outcomes and care experiences; fewer data on costs. *Health Affairs*, 32(2), 207–214.
- Hibbard, J. H., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the patient activation measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Services Research*, 39(4 Pt. 1), 1005–1026. doi:10.1111/j.1475-6773.2004.00269.x.
- Higgins, T. (2015, April 26). Apple's healthkit linked to patients at big los angeles hospital. [Bloomberg.com](http://www.bloomberg.com/news/articles/2015-04-26/apples-healthkit-linked-to-patients-at-big-los-angeles-hospital). Retrieved from <http://www.bloomberg.com/news/articles/2015-04-26/apples-healthkit-linked-to-patients-at-big-los-angeles-hospital>.
- Hood, L., & Flores, M. (2012). A personal view on systems medicine and the emergence of proactive P4 medicine: Predictive, preventive, personalized and participatory. *New Biotechnology*, 29(6), 613–624. doi:10.1016/j.nbt.2012.03.004.
- Humer, C., & Finkle, J. (2014, September 24). Your medical record is worth more to hackers than your credit card. *Reuters*. Retrieved from <http://www.reuters.com/article/us-cybersecurity-hospitals-idUSKCN0HJ21I20140924>.
- Institute of Medicine. (2012). *Best care at lower cost*. Washington, DC: The National Academies Press. doi:10.17226/13444.

- Iodine. (2014, September 23). Retrieved August 26, 2016, from <http://www.prnewswire.com/news-releases/iodine-launches-service-that-transforms-100000-americans-real-life-experience-into-unique-data-driven-tools-about-medications-for-consumers-276858401.html>.
- Iodine Data. n.d.. Retrieved from <http://www.iodine.com/data>.
- Jang, S., & Atkins, M. B. (2013). Which drug, and when, for patients with BRAF-mutant melanoma? *The Lancet Oncology*, *14*(2), e60–e69. doi:10.1016/S1470-2045(12)70539-9.
- Japsen, B. (2013, December 22). *ObamaCare, doctor shortage to spur \$2 billion telehealth market*. Retrieved October 10, 2016, from <http://www.forbes.com/sites/brucejapsen/2013/12/22/obamacare-doctor-shortage-to-spur-2-billion-telehealth-market/>.
- Kayyali, B., Knott, D., & Van Kuiken, S. (2013). *The big-data revolution in US Health Care: Accelerating value and innovation* (pp. 1–13). New York City, NY: Mc Kinsey & Company.
- Kumar, R. B., Goren, N. D., Stark, D. E., Wall, D. P., & Longhurst, C. A. (2016). Automated integration of continuous glucose monitor data in the electronic health record using consumer technology. *Journal of the American Medical Informatics Association*, *23*(3), 532–537. doi:10.1093/jamia/ocv206.
- Kummar, S., Williams, P. M., Lih, C.-J., Polley, E. C., Chen, A. P., Rubinstein, L. V., et al. (2015). Application of molecular profiling in clinical trials for advanced metastatic cancers. *Journal of the National Cancer Institute*, *107*(4), djv003. doi:10.1093/jnci/djv003.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. doi:10.1038/35057062.
- Leventhal, R. (2015, April 21). *How duke is using HealthKit to get patient-generated data into the EHR*. Retrieved October 5, 2016, from <http://www.healthcare-informatics.com/article/how-duke-using-healthkit-get-patient-data-ehr>.
- Liao, K. P., Cai, T., Savova, G. K., Murphy, S. N., Karlson, E. W., Ananthakrishnan, A. N., et al. (2015). Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ*, *350*(apr24\_11), h1885. doi:10.1136/bmj.h1885.
- Longhurst, C. A., Harrington, R. A., & Shah, N. H. (2014). A “green button” for using aggregate patient data at the point of care. *Health Affairs*, *33*(7), 1229–1235. doi:10.1377/hlthaff.2014.0099.
- Mann, D. M., Quintiliani, L. M., Reddy, S., Kitos, N. R., & Weng, M. (2014). Dietary approaches to stop hypertension: Lessons learned from a case study on the development of an mHealth behavior change system. *JMIR mHealth and uHealth*, *2*(4), e41.
- Metcalfe, D., Milliard, S., Gomez, M., & Schwartz, M. (2016, October 3). *Wearables and the internet of things for health*. Retrieved from <http://pulse.embs.org/september-2016/wearables-internet-of-things-iot-health/?trendmd-shared=1>.
- Might, M., & Wilsey, M. (2014). The shifting model in clinical diagnostics: How next-generation sequencing and families are altering the way rare diseases are discovered, studied, and treated. *Genetics in Medicine*, *16*(10), 736–737.
- Morris, Z., Wooding, S., & Grant, J. (2011, December). *The answer is 17 years, what is the question: Understanding time lags in translational research*. Retrieved October 30, 2016, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3241518/>.
- National Academies Press. (2011). *Toward precision medicine: Building a knowledge network for biomedical research and a new taxonomy of disease*. Washington, DC: National Academies Press. Retrieved from <http://www.nap.edu/catalog/13284>.
- National eHealth Collaborative. (2013, December). *Patient-generated health information technical expert panel final report*. Retrieved October 21, 2016, from [https://www.healthit.gov/sites/default/files/pghi\\_tep\\_finalreport121713.pdf](https://www.healthit.gov/sites/default/files/pghi_tep_finalreport121713.pdf).
- National Institutes of Health (2016). *About the Precision Medicine Initiative Cohort Program*. Retrieved October 30, 2016, from <https://www.nih.gov/precision-medicine-initiative-cohort-program>.
- Neff, G. (2013). Why big data won't cure us. *Big Data*, *1*(3), 117–123.
- Nielsen. (2014, April 16). *Hacking health: How consumers use smartphones and wearable tech to track their health*. Retrieved October 3, 2016, from <http://www.nielsen.com/us/en/insights/news/2014/hacking-health-how-consumers-use-smartphones-and-wearable-tech-to-track-their-health.html>.



- Panahiazar, M., Taslimitehrani, V., Jadhav, A., & Pathak, J. (2014). Empowering personalized medicine with big data and semantic web technology: Promises, challenges, and use cases. *Proceedings: IEEE International Conference on Big Data, 2014*, 790–795. doi:10.1109/Big-Data.2014.7004307.
- Parodi, S., Riccardi, G., Castagnino, N., Tortolina, L., Maffei, M., Zoppoli, G., et al. (2016). Systems medicine in oncology: Signaling network modeling and new-generation decision-support systems. In U. Schmitz & O. Wolkenhauer (Eds.), *Systems medicine* (pp. 181–219). New York: Springer.
- PatientsLikeMe Services(n.d.) Retrieved from <http://news.patientslikeme.com/services>.
- PwC. (2014, October 21). *Wearable technology future is ripe for growth—Most notably among millennials, says PwC US*. Retrieved October 5, 2016, from <http://www.pwc.com/us/en/press-releases/2014/wearable-technology-future.html>.
- Rask, K. J., Ziemer, D. C., Kohler, S. A., Hawley, J. N., Arinde, F. J., & Barnes, C. S. (2009). Patient activation is associated with healthy behaviors and ease in managing diabetes in an indigent population. *The Diabetes Educator*, 35(4), 622–630. doi:10.1177/0145721709335004.
- Robert, C., Karaszewska, B., Schachter, J., Rutkowski, P., Mackiewicz, A., Stroiakovski, D., et al. (2015). Improved overall survival in melanoma with combined dabrafenib and trametinib. *New England Journal of Medicine*, 372(1), 30–39. doi:10.1056/NEJMoa1412690.
- Robinson, R. (2016, February). *Patients and patient organizations power rare disease therapies—Pharmavoice*. Retrieved October 1, 2016, from <http://www.pharmavoice.com/article/2016-02-rare-disease-therapies/>.
- Rolnick, J., Downing, N. L., Shepard, J., Chu, W., Tam, J., Wessels, A., et al. (2016). Validation of test performance and clinical time zero for an electronic health record embedded severe sepsis alert. *Applied Clinical Informatics*, 7(2), 560–572. doi:10.4338/ACI-2015-11-RA-0159.
- Sarasohn-Kahn, J. (2008). *The Wisdom of Patients: Health Care Meets Online Social Media*. Retrieved from <http://www.chcf.org/topics/chronicdisease/index.cfm?itemID=133631>.
- Schueller, S. M., Tomasino, K. N., Lattie, E. G., & Mohr, D. C. (2016). Human support for behavioral intervention technologies for mental health: The efficiency model. *Management*, 21, 22.
- Sepah, S. C., Jiang, L., & Peters, A. L. (2015). Long-term outcomes of a web-based diabetes prevention program: 2-year results of a single-arm longitudinal study. *Journal of medical Internet research*, 17(4), e92.
- Shapiro, M., Johnston, D., Wald, J., & Mon, D. (2012). *Patient-generated health data. White Paper: Prepared for Office of Policy and Planning*. Retrieved from [http://healthitgov.ahrqdev.org/sites/default/files/rti\\_pghd\\_whitepaper\\_april\\_2012.pdf](http://healthitgov.ahrqdev.org/sites/default/files/rti_pghd_whitepaper_april_2012.pdf).
- Shay, L. A., & Lafata, J. E. (2015). Where is the evidence? A systematic review of shared decision making and patient outcomes. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 35(1), 114–131. doi:10.1177/0272989X14551638.
- Shortliffe, E., & Cimino, J. (2014). *Biomedical informatics: Computer applications in healthcare and medicine* (4th ed.). London: Springer.
- Sifferlin, A. (2014, September 24). *Look up your meds on this massive new drug database*. Retrieved November 1, 2016, from <http://time.com/3425527/iodine-prescription-drug-reviews/>.
- Sittig, D. F., Wright, A., Osheroff, J. A., Middleton, B., Teich, J. M., Ash, J. S., et al. (2008). Grand challenges in clinical decision support. *Journal of Biomedical Informatics*, 41(2), 387–392. doi:10.1016/j.jbi.2007.09.00.
- Smith, A. (2015, April 1). *U.S. smartphone use in 2015*. Retrieved from <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>.
- Sunyaev, A., Dehling, T., Taylor, P. L., & Mandl, K. D. (2015). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, 22(e1), e28–e33. doi:10.1136/amiajnl-2013-002605.
- The White House. (2015, January 30). *FACT SHEET: President Obama's Precision Medicine Initiative*. Retrieved October 20, 2016, from <https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative>.

- Trainer, A. H., Lewis, C. R., Tucker, K., Meiser, B., Friedlander, M., & Ward, R. L. (2010). The role of BRCA mutation testing in determining breast cancer therapy. *Nature Reviews Clinical Oncology*, 7(12), 708–717. doi:10.1038/nrclinonc.2010.175.
- Tricoci, P., Allen, J. M., Kramer, J. M., Califf, R. M., & Smith, Jr. S. C. (2009). Scientific evidence underlying the ACC/AHA clinical practice guidelines. *Journal of the American Medical Association*, 301(8), 831–841.
- Tu, H. T., & Lauer, J. (2008, December). *Word of mouth and physician referrals still drive health care provider choice*. Retrieved November 1, 2016, from <http://www.hschange.com/CONTENT/1028/>.
- U.S. Department of Health and Human Services. (2013, January). *Realizing the full potential of health information technology to improve healthcare for americans: The path forward*. Retrieved October 4, 2016, from [https://www.healthit.gov/sites/default/files/hitpc\\_stage3\\_rfc\\_final.pdf](https://www.healthit.gov/sites/default/files/hitpc_stage3_rfc_final.pdf).
- U.S. Department of Health and Human Services. (2015, February 9). *Mobile medical applications: Guidance for industry and food and drug administration staff*. Retrieved from <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf>.
- U.S. Department of Health and Human Services. (2009, October 28). *HITECH Act enforcement interim final rule*. Retrieved from <http://www.hhs.gov/hipaa/for-professionals/special-topics/HITECH-act-enforcement-interim-final-rule/index.html>.
- U.S. Department of Health and Human Services. (2012, June 26). *Alaska DHSS settles HIPAA security case for \$1,700,000*. Retrieved October 18, 2016, from [www.hhs.gov/hipaa/for-professionals/compliance-enforcement/examples/alaska-DHSS/index.html](http://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/examples/alaska-DHSS/index.html).
- U.S. Department of Health and Human Services. (2013, March 26). *HIPAA administrative simplification*. Retrieved from <https://www.hhs.gov/sites/default/files/oct/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf>.
- U.S. Preventative Services Task Force. (2014, December 24). *BRCA-related cancer: risk assessment, genetic counseling, and genetic testing*. Retrieved from <https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/brc-related-cancer-risk-assessment-genetic-counseling-and-genetic-testing#citation27>.
- US Department of Veterans Affairs Office of Research and Development. (n.d.) *Million veterans program*. Retrieved October 30, 2016, from <http://www.research.va.gov/mvp/veterans.cfm>.
- Wicks, P. (2007, July 7). *Excessive yawning in ALS/MND*. Retrieved October 1, 2016, from <http://blog.patientslikeme.com/2007/07/07/excessive-yawning-in-alsmnd/>.
- Winner, J. G., Carhart, J. M., Altar, A., Allen, J. D., & Dechairo, B. M. (2013). A prospective, randomized, double-blind study assessing the clinical impact of integrated pharmacogenomic testing for major depressive disorder. *Discovery Medicine*, 16(89), 219–227.
- Winner, J. G., Carhart, J. M., Altar, C. A., Goldfarb, S., Allen, J. D., Lavezzari, G., et al. (2015). Combinatorial pharmacogenomic guidance for psychiatric medications reduces overall pharmacy costs in a 1 year prospective evaluation. *Current Medical Research and Opinion*, 31(9), 1633–1643. doi:10.1185/03007995.2015.1063483.
- Witters, D., & Agrawal, S. (2014, November 3). *How mobile technology can improve employees' well-being*. Retrieved October 10, 2016, from <http://www.gallup.com/businessjournal/179111/mobile-technology-improve-employees.aspx>.
- Xu, J., Rasmussen, L. V., Shaw, P. L., Jiang, G., Kiefer, R. C., Mo, H., et al. (2015). Review and evaluation of electronic health records-driven phenotype algorithm authoring tools for clinical and translational. *Journal of the American Medical Informatics Association*, 22 (6), 1251–1260. doi:10.1093/jamia/ocv070
- Yang, Y. T., & Silverman, R. D. (2014). Mobile health applications: The patchwork of legal and liability issues suggests strategies to improve oversight. *Health Affairs*, 33(2), 222–227. doi:10.1377/hlthaff.2013.0958.