Big Tech and COVID-19: Risks Associated with Engaging Our Digital First Responders

Cassandra Skapek


The U.S. healthcare landscape is primed for disruption. National health care spending climbed to 17.7% in 2019 and CMS projections estimate that they will reach an unprecedented high of 19.7% or $6.2 trillion by 2028 (CMS NHE Fact Sheet, 2020). Moreover, Americans are growing increasingly dissatisfied with the value of their care, which has sparked the beginnings of an industry-wide transformation of care delivery with a focus on quality improvement, cost containment, and increased patient centrivity. Given these trends, it comes as no surprise that four of the five Big Tech companies—Google, Microsoft, Amazon, and Apple—have spent the last two decades quietly laying the groundwork for their role in the disruption of the health care sector. In this moment of national crisis as a result of the COVID-19 pandemic, these companies seem perfectly poised to fulfill what Satya Madella, the CEO of Microsoft, described as their role of “digital first responders” (Foer, 2020). However, without sufficient regulatory protections in place to protect patient data from being monetized, providers and public health officials should proceed cautiously when deciding whether to outsource their COVID-19 response operations to Big Tech companies.

Technology developments over the last decade played an essential role in mediating the disruption caused by COVID-19. It is difficult to imagine life nowadays without the ability to stay connected with coworkers and loved ones alike on Zoom or to deliver groceries and essential supplies to our homes in a matter of hours via Amazon Prime. As the COVID-19 pandemic recently passed the one-year anniversary of the first confirmed case, there is now unanimous recognition of the value of digital health solutions in relieving the strain on health systems and expanding access to care for patients. While the pandemic accelerated the adoption of telehealth for care delivery, it also exposed major vulnerabilities and limitations in health systems’ technology capabilities; many provider organizations had to work quickly to upgrade and enhance their IT systems, which were woefully unprepared to support a full operational pivot to a virtual format seemingly overnight.

While partnering with Tech Giants is an undeniably attractive option for provider organizations lacking the computational capacity and operations expertise for efficient contact tracing and vaccine distribution, the level of risk associated with granting these companies access to sensitive patient data cannot be overstated. An analysis of Google’s prior investments in the healthcare space and its current role in the COVID-19 response can help shed light on the risks associated with engaging Big Tech in the analysis of patient clinical data.

An Overview of Google’s Healthcare Strategy: “Help everybody live their healthiest life”

Google is host to both the largely uncontested leading search engine in the world (used for 87% of search engine traffic) and the third largest cloud computing platform (with $2.78B in revenue in Q1 2020) (Dang, 2020). In recent years, the company has made significant moves in healthcare that have given traditional players cause for concern, leveraging its brand reputation as well as its technology expertise and unmatched access to enormous amounts of data in order to fulfill its mission statement, “to help everyone live more life every day through products and services that connect and bring meaning to health information” (“Google Health”, 2019). The company’s cloud computing platform, Google Cloud, is used by a number of healthcare organizations and in 2018, their health subsidiary launched a new Google Cloud Healthcare API to tackle data interoperability challenges in healthcare (Moore, 2018).

One of the critical components to success in launching a digital healthcare strategy for Big Tech companies is access to patient data, which has historically been the exclusive domain of the provider. While health systems have begun to recognize the untapped potential of this asset, with limited resources for innovation, they have struggled to reap the benefits.

Google has begun to align itself with provider organizations, leveraging its own technology expertise to deploy new clinical decision support tools and extract valuable insights for clinicians from electronic health record data. For example, Google has partnered with providers to deploy its artificial intelligence technology to mammography screening initiatives used in the early detection and treatment
of breast cancer (Pisano, 2020). In a 2020 article in *Nature*, this technology has been shown to outperform radiologists in accurately interpreting mammograms and was able to significantly reduce the incidence of false negatives (McKinney et al., 2020).

Despite this potential for clinical benefit, more recently Google has been expanding both the scope and scale of their provider partnerships at an alarming rate, examples of which include a 10-year collaboration with Mayo Clinic and a large-scale agreement with Ascension, codenamed Project Nightingale. In granting Google access to medical records of millions of patients across multiple states, both partnerships seek to apply AI and machine learning to the health system’s datasets in order to solve complex use cases and provide new clinical insights (Anastasijevic, 2019; Copeland, 2019). Despite the federal legality of these partnerships, the comprehensiveness of the patient data shared raised ethical questions and it is clear that Google intends to use Project Nightingale as a model for future work with other health systems.

**Google’s COVID-19 Response**

From the early days of the COVID-19 pandemic, Google has opportunistically begun to amass user data in order to provide epidemiological insights to public health officials. Anonymized location history data is being used in COVID-19 Community Mobility Reports, which track travel to and from locations such as retail and grocery stores, recreation facilities, and workplaces (“Google Health: COVID-19,” 2020). The company has also leveraged its dominance as a search engine to create a self-assessment tool to reduce strain on limited healthcare resources and a COVID-19 Search Trends dataset, which is being utilized by health departments to identify emergent outbreaks. Officials can then use these datasets to adapt stay-at-home orders to local contexts based on current need. Finally, Google collaborated with Apple to create an exposure notification system for individuals whose location data indicated close proximity to a confirmed COVID-19 case. These initiatives demonstrate the enormous potential to do good that exists in utilizing this data to enhance clinical and public health decision-making.

**Privacy and Equity Concerns**

Google’s recent initiatives in response to the COVID-19 pandemic have caused questions regarding privacy and security to be brought to the forefront once more. Through the Health Insurance Portability and Accountability Act (HIPAA) and the U.S. Department of Health and Human Services Office of Civil Rights (OCR), the federal government established strict guidelines and requirements for organizations governing the use of individuals’ protected health information (PHI). Amongst partnerships with Big Tech companies, the focus thus far has been on ensuring HIPAA compliance, however, this begs the question of whether regulations dating back to 1996 are sufficient to defend the privacy of patient data today, in an era where many of the practices and expectations surrounding handling of data have shifted. One example of an area of potential risk deals with PHI that has been de-identified and thus no longer protected by HIPAA rules, opening the door for Big Tech companies to monetize this data to be used for secondary purposes (U.S. Department of Health and Human Services Office for Civil Rights, p.79, 2013). Prematurely entering into agreements with these companies without careful consideration of the limitations of current legislation jeopardizes the sacristy of the patient-provider relationship and creates the potential for patient data to be used outside the original terms of partnership agreements.

In addition to privacy concerns, reliance on machine learning and AI for clinical insights raises questions regarding the impact of bias on patient care and outcomes. As data scientist, Cathy O’Neil stated in the popular documentary, *The Social Dilemma*, “algorithms are opinions embedded in code...[they] are not objective. Algorithms are optimized to some definition of success.” The potential for machine learning algorithms to perpetuate biases in diagnosis and treatment is well documented, and identifying and addressing bias remains a challenge (Gianfrancesco, Tamang, Yazdany, & Schmajuk, 2018). Additionally, bias is common in source data; examples of this include missing EHR data, limited sample sizes, samples skewed towards the most severe cases, and misclassification by clinicians, all of which detrimentally affect the ability of machine learning algorithms to make accurate predictions regarding patient care and outcomes. Reliance on bias-ridden algorithms and data sets raises major ethical and health equity questions on how their use will contribute to the reinforcement of existing health disparities. Given the recent firing of a leading AI ethics researcher at Google for her criticism of the company’s approach to handling biases in AI (and other team members soon after), there is serious doubt regarding Big Tech’s willingness to address these issues (Metz & Wakabayashi, 2020; Metz, 2021).

**Future Directions**

The companies who experience success moving forward will be those who find the best avenue to access and leverage this data in the future. At this point, Big Tech companies are still dependent on partnerships with providers for their greatest asset: patient clinical data. Provider organizations have begun to see the risks associated with granting Big Tech unfettered access to their patient data. The company Truveta was formed as a partnership amongst 14 provider organizations to enable health systems to pool their data into larger datasets and perform similar analysis as Big Tech, but with fewer concerns over misuse of patient data (Vaidya, 2021). The overarching takeaway is
that the current conflict in America surrounding user data privacy is in part due to the role of technology companies expanding faster than the regulatory environment can adapt to support them, resulting in the mass monetization of user data and subsequent security and privacy challenges. The current regulatory environment for clinical data is not equipped to prevent PHI from following the same trajectory as user data. Therefore, it would be a grave mistake to delegate the responsibilities of the health care system from our actual first responders to our digital first responders prior to establishing strict guardrails to prevent similar abuse of confidential patient data.

References


