Cal INDEX – A New Model for Collaboration
An Evolving Healthcare Landscape

The U.S. healthcare system has arrived at a tipping point. In the face of an aging population and alarming rates of chronic disease that threaten to overwhelm existing capacity and propel costs higher, stakeholders across the industry are united in calls for reform. Although the specific nature of reform remains a contentious issue, many agree that adopting health information technologies is an important first step. Health IT offers a new set of tools that can enable patients and providers to acquire, aggregate, access, analyze, and act on healthcare data to improve operational efficiency, assess the effectiveness of medical practices, drive down costs, and deliver better outcomes for patients.

However, health IT alone is not the solution, as demonstrated by recent challenges with electronic health record (EHR) adoption. Through the Health Information Technology for Economic and Clinical Health Act (HITECH), Congress appropriated as much as $36.5 billion to incent hospitals, practices, and clinicians to adopt and use EHRs. While the program has dramatically increased the number of providers with EHR systems, the industry now faces a significant barrier to interoperability: most EHR systems are proprietary and were not designed to communicate with other health information systems or a competitor’s products. Though the technology is now largely in place, users are still frequently unable to coordinate care outside the four walls of their institution.

In fact, the issue of competition has significantly hampered healthcare reform efforts for years. Lately, technology vendors have shouldered much of the blame, but other stakeholders are also at fault. Providers and payers have both historically spurned the open exchange of data for their member and patient populations, fearing that access to their data could contribute to consumers opting to use a competitor’s services instead.

Fortunately, many organizations are beginning to recognize that effective reform cannot occur within the competitive siloes that have long dominated healthcare. Instead, these innovative organizations have created unique partnerships that seek to leverage the combined power of data from different sources to improve patient care. One such organization is Cal INDEX, the largest statewide health data exchange in the U.S.

Collaborating for Change

Cal INDEX was founded in August 2014, but its conception can be traced back to conversations between executives at Blue Shield of California and Anthem Blue Cross twelve months prior. Though normally competitors in the health insurance market in California, executives at both organizations realized that optimizing revenue, reducing costs, and improving the quality of care using only their own internal data was becoming increasingly difficult. Because patients often traverse the care continuum while moving in and out of established provider and payer networks, healthcare organizations are frequently unable to maintain a comprehensive view of a patient’s care from the data they possess. Competitors
often hold much of the missing information. According to Dr. Kenneth Park, Vice President of Research Environments and Business Solutions at HealthCore (a wholly owned subsidiary of Anthem), “We’ve been trying to do all these things just using our own data, but our ability to get incremental gains and improvements is pretty limited at this point.”

With an increased focus on quality and cost brought about by the development of accountable care organizations (ACO’s) and bolstered by the growing trend of healthcare consumerism, Anthem and Blue Shield of California faced challenges about how they used their data to support care. David Watson, President and CEO of Cal INDEX noted, “As providers and payers participate in different models of care like ACOs, they’re recognizing that collectively managing risk requires this notion of interoperability and data sharing.” To truly improve care, organizations will have to change their mindset about the strategic value of data.

Consequently, leaders at Blue Shield of California and Anthem decided it was in their best interest to collaborate around data instead of competing to acquire it. They envisioned the creation of a public utility to serve all of California and facilitate the collection of and access to healthcare data. Importantly, the new entity would not only feature data from payer organizations, but would also incorporate data from participating healthcare providers and their clinical systems. With such an infrastructure in place, participants would instead compete around their use of the data. “We put a lot of money into trying to acquire data, but the data itself doesn’t actually generate direct value. The way we should all be competitively differentiating is on how we analyze data, interpret it, and then apply those results to change our business decisions and operations,” said Dr. Kenneth Park. In little more than a year, Cal INDEX was born.

As of early 2015, Cal INDEX has loaded data for 9 million members, including “medical information from claims records, prescription, and lab data,” explained David Watson. Clinical data, however, is not currently accessible through the platform. “We don’t have general access open yet because we’re still testing the platform and working through a few things we wanted to have in place first,” he added. Over the next few months, providers will be able to participate in Cal INDEX and contribute data such as lab results, electronic health records, and admission/discharge/transfer (ADT) data to the system. Providers will have access to the data through a clinical portal to view information in a structured manner; information will be presented on summary screens with the ability to display more detailed views of specific data. Moreover, to entice more providers to participate, Cal INDEX will enhance the clinical usability of the system by offering the capability for participants to download and import data directly into their analytics and data warehousing solutions.

**Powering New Opportunities with Open Data**

Aggregating data into an integrated database offers a slew of potential benefits. On the payer side, combining data affords economies of scale. Dr. Kenneth Park described how costly the process of building complete patient records can be in
the current environment. “The cost of doing a chart abstraction can be anywhere from $20-$30 per member. Right now, we can only do chart abstractions on a tiny handful of our members.” By investing in a technology infrastructure that enables data exchange, “there is a large upfront investment, but the cost of transmitting data afterwards is significantly lower per member. For approximately the same cost [as performing chart abstraction on a fraction of members] we can now obtain data on our entire membership.” Collaborating with other organizations will help to reduce the costs of the upfront technology investment.

Similarly, aggregating data across institutions can contribute to revenue optimization. Various quality measurement and reporting systems like HEDIS and Medicare’s Star Rating System evaluate the performance of health plans on dimensions of care and service such as customer service, receipt of preventative care, chronic disease management, and prescription medication benefits. Employers and consumers often use these scores to help them select the best health plan. According to Dr. Park, “these translate into meaningful revenue, but the ability to report on measures is largely dependent upon how accurate our patient data is. As a payer, we’re missing out on a lot of clinical data.”

Brian Russon, Director of Sales for National Accounts at Orion Health, highlighted the importance of data collection for payers in the wake of the Affordable Care Act’s disruption of the insurance marketplace. “You now have the opportunity to bring in new member populations through the health insurance exchanges with their medical histories available immediately, and intact, rather than waiting for claims data files to come through. New members can be rapidly enrolled in preventative and disease management programs to promptly address their needs.” Acting swiftly to manage new populations can improve the patient experience while reducing the long-term costs of allowing medical issues to go unaddressed.

Health plans also need clinical data on their member populations for internal risk calculations. Bill O’Connor, MD, Orion Health’s Vice President of Consulting Services suggested, “As we transition from a volume to a value type of model, providing value means managing and controlling risk. Having access to that entire dataset allows payers and providers to risk stratify their patient populations and proactively target patients they think will need intervention, often before they get sick.” Preventative measures can reduce downstream costs and alleviate health issues before they become more difficult to manage.

For providers, combining claims and clinical data is essential to establishing a comprehensive view of a patient’s care. Clinical data is often incomplete, particularly if a patient has visited with clinicians outside of a given network. Claims data can add depth to the clinical record by contributing additional information about the patient and their medical history. Dr. Chris Hobson, Chief Medical Officer at Orion Health, emphasized the importance of supplementing clinical data with claims data: “Patient demographics may be more up-to-date in the claims data. The encounters a patient had may be more accurate, and certainly more comprehensive. You can actually get perspective on how intensely a patient is using the system.”
With a more complete record, clinicians can better identify and address gaps in care. For example, a clinician knows which medications they have prescribed, but doesn’t have information about which ones were actually filled by the patient. Claims data includes this information. “Now I can see that I’ve written seven medications for this patient, but they only picked up five. Maybe the two that they skipped are the more important ones, and maybe they didn’t pick them up because of the price. Somebody can now work with the patient to find a solution,” said Dr. Hobson.

Data access can also improve care coordination and population health management. Cal INDEX can help close communication loops between members of a care team, different departments, and even different facilities using admission, discharge, and transfer (ADT) data. “The primary care provider or specialist on a care team can be notified when one of their patients is admitted or discharged from the hospital,” stated David Watson. “They otherwise might not know their patient was even in the hospital until the next patient visit.” With this information, clinicians can know to look for a discharge summary or care plan to promptly follow-up with their patient. “There are a lot of basic things around gaps in care and care coordination that we haven’t seen today that transparency is going to provide,” noted David Bennett, Executive Vice President for Healthier Populations at Orion Health.

Comprehensive datasets can improve operational efficiency for individual clinicians, hospitals, and health systems. A complete longitudinal patient record can save clinicians valuable time otherwise spent tracking down missing data. “Instead of having to look all over the place for information on a patient, Cal INDEX gives clinicians a whole dataset with which to work with to make care decisions,” explained Dr. O’Connor. Dr. Chris Hobson raised the question of value and financial efficiency. “We spend billions of dollars on medical services, but are we giving that amount of value back to our population,” he asked. Administrators can use more accurate and complete data on their patient populations to compare outcomes with healthcare spending to identify overutilization and promote cost-effective services.

A Data Platform for Today’s World...

Beyond the unprecedented cooperation between competitors, Cal INDEX is unique in its capabilities as a data platform. Integrating data from two large health insurers and a number of participating provider organizations presented a host of new challenges to traditional healthcare data management practices. For one, the sheer quantity of data required a different technical approach to the database.

Most healthcare databases today use a standard relational model, which stores data in sortable tables with rows representing individual records (e.g. patient 1, 2, 3, etc.) and columns representing attributes about those records (e.g. name, date of birth, ID number, etc.). To display information, a relational database will call for the appropriate data from a table and return the result. Tables can also
reference one another, which speeds up the process of more complex queries; instead of requesting data from each table individually, the database will request information from a merged table that references the location of the data in other tables. However, relational databases are ill-equipped to handle massive quantities of data. As more elements are added to the database, locating and calling data from tables and then building new tables to support complex relationships is progressively more difficult.

Because Cal INDEX is intended as a utility service for all of California, it needs to handle a large volume of data and process that data quickly. Instead of using a relational model, Cal INDEX leverages an open-source NoSQL database called Cassandra. NoSQL databases were developed in response to the growing volume of data being collected in other industries and are used by pioneering technology firms like Twitter, Instagram and Netflix. They are both more scalable and flexible than a relational database. NoSQL databases, for instance, utilize distributed computing technology. Instead of relying on a single processing unit or server like a relational database, distributed databases use networks of computers to speed up processing of data transactions and queries. “If we want to add performance we can elastically scale our capabilities by throwing more hardware at the problem,” explained Drew Ivan, Director of Business Technology at Orion Health. While it’s very difficult to improve the processing capabilities of a single server without purchasing a more powerful version, distributed networks can bring the computing power of additional units to bear on a problem as needed.

NoSQL databases also store data in different formats. Again, in a relational database, records are stored in relation to attributes in a table. This means that those relationships must be defined prior to data importation. On the other hand, NoSQL databases don’t use the relational model, so data can be imported and then merged into the database’s data model without having to refactor the entire structure of the database. “NoSQL technology is more flexible in terms of storing whatever you get and then making sense out of it later,” noted Ivan.

In the case of Cal INDEX, this flexibility is especially important because of the different types of data being hosted in the database. In general, clinical data is fairly well structured. “It has a really refined code set and there are defined criteria for moving clinical data from point A to point B; most standards already exist in the clinical data space,” said Brian Russon. “Claims data is a different beast. There are also standards that exist, but few standards for integrating claims data with clinical data.” Merging the two types of data required new thinking about how to integrate the information in a cohesive fashion, and Cassandra offered the flexibility to address the issue.

Finally, Cal INDEX’s NoSQL technology creates new opportunities for using data. As organizations try to become more intelligent about how they can improve outcomes and lower costs, they’ve turned toward advanced applications of analytics. Dr. O’Connor described how analytics are performed on traditional relational databases: “Typically you have to do what’s called an extract, transform, and load (ETL). You have to take the data that’s in the database, move
it somewhere else, and then run your analytics on it. It’s very slow and very inefficient to do it that way.”

In contrast, NoSQL databases are capable of real-time streaming analytics. According to Ivan, “We can send a patient record or other data set out to an analytics tool and immediately get an answer back about whether there are gaps in care or other things we need to look into instead of waiting a day or week for the analytics engine to crunch all the data and send back a report.” Analytics tools can leverage the benefits of distributed computing by harnessing additional processing power as needed to return results more quickly, which “can benefit the user from an operational cost perspective as they only have to pay for that added elastic computing power,” added David Bennett. Currently, Cal INDEX is focused on importing and integrating data into its platform and does not yet offer advanced analytics capabilities. Until new capabilities are added, Cal INDEX offers other ways to access and use data.

The simplest method is to securely pull data from the platform to import into a user’s own analytics tools. Participants can run a file extract of the longitudinal patient record to import into a data warehouse or analytics system. Alternatively, Cal INDEX takes advantage of another prominent technology tool to enhance participants’ abilities to use the data in the database, application programming interfaces (APIs). APIs are standardized instructions and protocols for developers to build tools and applications that can make use of other proprietary software or services. APIs enable developers to interface directly with Cal INDEX’s database to use the data for a range of potential purposes. “The ability to have a common set of APIs will enable customers to build their own applications right on top of the database and easily extract data for those programs. Providers will be able to build tools like mobile applications to support disease management or connect their analytics solutions,” said Dr. O’Connor.

APIs hold great importance for Cal INDEX’s long-term sustainability. Dr. Kenneth Park believes that one of the issues that has plagued traditional health information exchange (HIE) organizations is the question of value. “They almost always start off with grant funding but know that it will run out. To create value to support charging usage fees to providers, they start building proprietary services and analytics tools.” While this may have been an effective strategy at one point in time, healthcare organizations have now committed to improving many of the same issues that these HIEs seek to address. As a result, “the largest providers, the ones with the most money to pay for HIE, have already invested in their own analytics tools and software. They see the HIE’s software as competitive with their own.” By leveraging APIs, Cal INDEX can reduce competitive concerns while also fostering a more collaborative environment where data supports innovation and accelerates overall improvements in healthcare.

...And the Future

Opportunities for leveraging data are omnipresent in healthcare. A utility service with the capacity to onboard and integrate emerging data types is valuable for
many different stakeholders interested in impacting care. In addition to offering a robust longitudinal view of patient records to providers and payers, Cal INDEX plans to support the research community. “When you have longitudinal records for 10, 20, or 30 million people, there’s a lot that you can learn. The medical research community is very interested in having access to that data. We’re already in talks with research-focused medical centers who are interested in joining us as a result,” said David Watson.

Moreover, as Cal INDEX evolves, its progressive technical architecture is designed to support new capabilities. The scalability and flexibility afforded by Cassandra could be put to use for other types of healthcare data. According to Don Soucy, Vice President of Sales for National Accounts at Orion Health, “There’s a tsunami of data coming. You have some that’s already there, the provider data and claims data. But then you have all of the personal device data as folks start using those tools. Beyond that, you have to think about personalized medicine, genomics, and other data that will help provide a full snapshot of each patient’s life.” Dr. O’Connor described a powerful use case for incorporating new types of data: “Imagine we have a 40 year old male. We know from his biometric device that he has high blood pressure and isn’t moving around a lot. From genetic tests, we know he has a marker that predisposes him to heart disease. With something like Cal INDEX, we can use analytics to find him, bring him in, and engage him instead of waiting six years until he shows up in the ER with a heart attack because of the uncontrolled high blood pressure and obesity that maybe we could have prevented.”

**Big Data Brings Big Responsibility**

For all of the potential advantages brought about by creating a public data infrastructure, there are also new risks to consider. Health information has grown as a popular target for digital attacks. Consumerization has led to a population that is more aware of how companies are using their data and more demanding for protections on that data. A massive database with a suite of healthcare data is bound to draw some negative attention. Recognizing the risks involved, Cal INDEX combines carefully crafted governance policies with technology to guarantee the greatest degree of privacy and security.

To maintain data security, all information in the database is encrypted at rest and in transit. Standard security mechanisms like firewalls work to detect and prevent intruders. There are also systems monitoring user access to rapidly alert security staff to patterns that appear abnormal. The platform uses a roles-based delegated security model to govern who can view patient records, and how much data they can see.

Cal INDEX’s privacy controls begin with a standard business associate (BA) agreement and participation agreement establishing a trusted relationship between organizations and outlining fundamental expectations around data access, management, and use. BAs then designate a security administrator to oversee the authorization of users with access to Cal INDEX and determine their
roles. Physicians have access to all of the data for patients with whom they have an established care relationship, but administrative staff would see less information. Clinicians without an established care relationship can view patient data in emergency situations, so long as they attest to having a valid reason to access the system. All such cases are reviewed by Cal INDEX’s privacy officers.

David Watson emphasized the importance of protecting patient privacy: “We’ve thought carefully about how to best provide consumers with the opportunity to manage their privacy.” Cal INDEX utilizes an opt-out consumer model to bolster patient choice about whether their data is accessible through the platform. The participation agreement instructs covered entities to notify patients that their data is being sent to Cal INDEX and that at any time, the patient is free to opt-out. The participation agreement also states that patients have the right to see what’s in their record and make corrections, as well as to request a disclosure of who has viewed their record, as required under HIPAA. In the future, Cal INDEX plans to allow patients to access their records directly in order to further empower consumers to manage their care and protect privacy.

Additionally, Cal INDEX has established other governance controls to reduce the risk of inappropriate access and protect participants. For instance, sensitive data such as behavioral health information, substance-abuse status, and HIV status, medical records for persons of public interest, and financial information won’t be included in the database. “You have to protect the privacy of individuals, the security of the system, and you have to be trusted. Trust means protecting the interests of the stakeholders participating in the system,” explained Dr. Kenneth Park. Consequently, data from Cal INDEX cannot be sold for profit, and all research must go through a formal approval process.

**Concluding Thoughts**

Cal INDEX is an important next step in the healthcare system’s evolution toward high quality, value-based care. Unlike many other efforts before it, Cal INDEX represents a restructuring of the competitive culture that has contributed to the lack of data sharing currently stifling care coordination. Founded on cutting-edge database architecture, Cal INDEX is one of the few healthcare organizations today capable of leveraging the types of technologies that have contributed to innovation and growth in other industries. With thoughtful governance to protect the interests of patients and users, Cal INDEX may reach unprecedented levels of participation. In the words of David Watson, Cal INDEX is “an experiment at scale” of a privately funded data exchange platform to “create a consistent way for participants to consume data and make it valuable for their enterprise and consumers alike.” It is likely that many will be watching Cal INDEX as a potential model for the future of healthcare interoperability.