Truveta and Boston Scientific Strive for the Ultimate in Data Sharing: Speed, Accuracy, and Privacy

The data aggregator Truveta and Boston Scientific have embarked on a partnership that tests how advances in healthcare system data sharing can lead to more nuanced insights into patient outcomes, as well as unprecedented information on the comparative effectiveness of specific medical devices and procedures.

▶ **Wendy Diller**

While medical device companies increasingly incorporate artificial intelligence and machine-learning-driven innovations into their clinical and commercial strategies, the data they rely on to power their algorithms is often incomplete and flawed, leaving them vulnerable to challenges and competitive disadvantages.

The problem isn't unique to the medical device industry or one of its own makings. Among the most obvious and prominent sources of patient data are healthcare systems. Yet, accessing the vast troves of patient data that are collected and stored within healthcare systems has been traditionally challenging, and as the device industry's reliance on data grows, these difficulties have become glaring. Healthcare systems are reluctant to share data, beyond select increments, and because the US healthcare system is so fragmented, the information they provide cannot capture the full longitudinal journey of a particular patient over time—a goal that has become invaluable, especially for subsectors of the industry adapting to value-based care and focusing on chronic conditions. (*See "Optimizing the Value of Healthcare System Data to Transform Care," this issue.*)

The pandemic intensified these flaws and highlighted the gaps between the promise of access to healthcare system data and the reality. But it also led to fast-tracking of creative strategies about the best ways to overcome long-standing barriers to data sharing, including the frustrating reluctance of healthcare systems—often but not always for valid reasons—to share their data with with external organizations.

An example of this new way of thinking and its potential to help medical device companies is embedded in a partnership between the two-year-old healthcare data aggregator **Truveta** and **Boston Scientific**.

The foundation for the collaboration began informally in 2020, shortly after Truveta was founded, when its chief medical officer, Ryan Ahern, MD, reached out to his former supervisor, Michael Jaff, DO, a vascular medicine specialist, who had been chair of the Department of Vascular Medicine and medical director of the Fireman Vascular Center at Massachusetts General Hospital (MGH), where Ahern completed a residency in primary care in 2015. The data world was moving quickly both in terms of technological advances and regulatory updates, driven by the urgency of the pandemic, as well as longer-term underlying trends, and Ahern was excited about Truveta's approach to helping life sciences companies. In particular, he saw it as a way for medical device manufacturers to solve obstacles to obtaining accurate, high-quality data about patients, and their medical journeys and outcomes.

Jaff, who had joined Boston Scientific in January 2020 as the company's chief medical officer and VP, medical affairs, innovation and technology, peripheral interventions, reacted immediately to his former student's overture. Truveta's potential, conceptually at least, seemed "almost too good to be true," he recalls.

Truveta enlists healthcare systems as "members," which agree to upload their complete, albeit de-identified, patient records (including electronic health records [EHRs], clinical notes, images, and

genomics) daily to Truveta servers. The frequency, speed, and careful curating of this data at scale differentiates Truveta from other aggregators. A key attraction is that the healthcare system members are the majority owners of Truveta and govern its processes, says Jaff, who was president of the Newton-Wellesley Hospital, an affiliate of MGH, for four years from 2016 to 2020, where he says patient privacy and data security was uppermost in the leadership team's minds. Furthermore, he adds, Truveta's data feeds are refreshed daily by its members and continually curated. "In all the population health research I've ever done in my career, we have relied on data from Medicare or commercial payors, and that data is often 12, 24, or even 36 months old. That's just the way the system works, particularly if it is federally funded."

The two doctors discussed opportunities to collaborate over the next year and a half and agreed on a pilot program to test the waters.

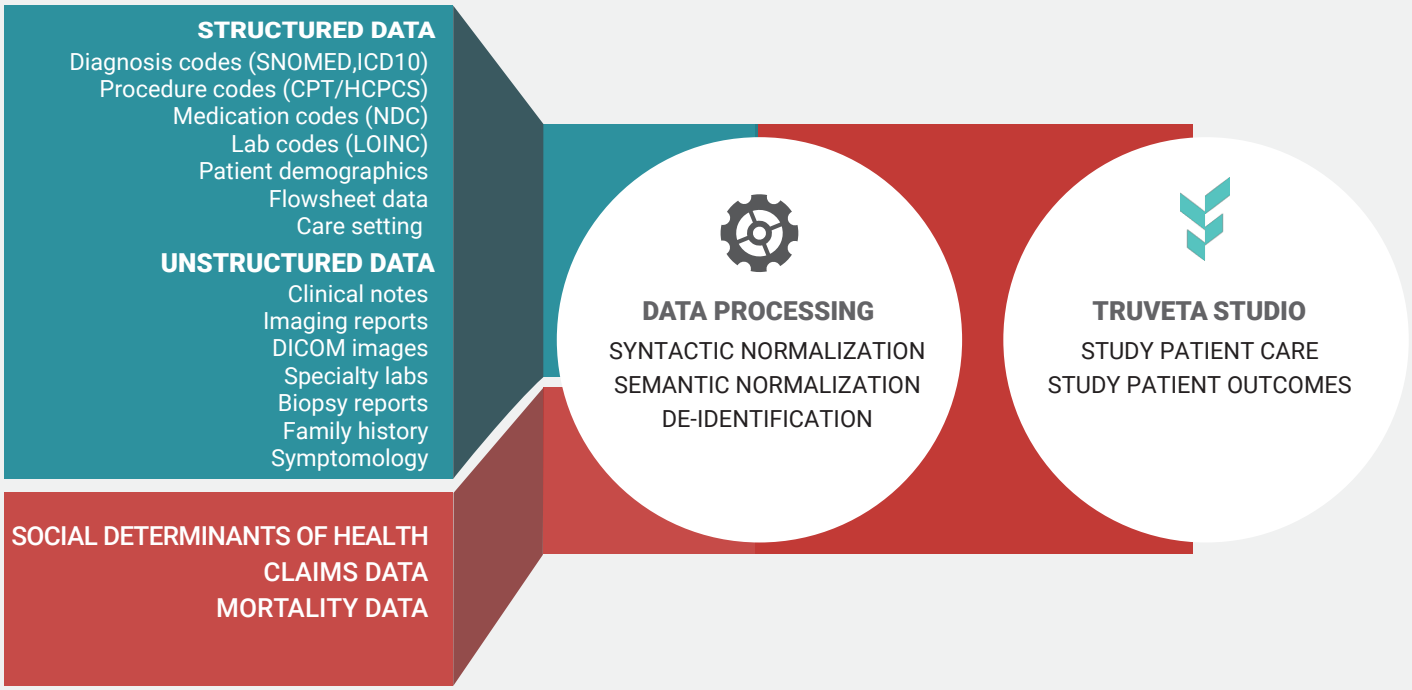
Healthcare Systems Take Ownership

Healthcare systems have been on an evolutionary journey—slow to figure out how to optimize and share the data they collect and store, and wary of enormous regulatory, technical, and security hurdles. Surveys show that providers and healthcare systems are overwhelmed by the onslaught of data on their patients, and that less than half of healthcare system executives have a formal digital health strategy, resulting in a reactive and slow procurement process. As a result, their ability to identify problems and find solutions is compromised.

Truveta was formed in the early days of the pandemic, by four healthcare systems looking for innovative pathways to "bring together talents that build on collective intelligence that can ask and answer questions about patient care and outcomes quickly," while balancing security and privacy demands, says Terry Myerson, a co-founder and CEO. Based on a concept outlined in a 2018 white paper by Providence, one of the largest health systems in the Northwest, the founders saw the collaboration as an innovative approach to tackling urgent challenges of caring for desperately ill patients facing a disease, COVID-19, that no one knew anything about. In addition to Providence, the founders included Advocate Aurora Health, Trinity Health, and Tenet Healthcare.

Myerson, a former Microsoft EVP, who led development of some of that company's key products, including Windows, had not previously worked in healthcare. But he saw the value of a healthcare-system-owned data

Figure 1
Truveta's Integrated Platform



Source: Truveta CMO Ryan Ahern's presentation to 3D Dartmouth Device Symposium, October 2022

aggregator and led the new company's work to design and implement sustainable organizational processes and a business model that would meet the exacting needs of the healthcare community and that also could underwrite the enormous costs of building an ambitious repository of longitudinal patient data that multiple external stakeholders could access.

A key differentiator is Truveta's capital structure. Its healthcare system members, now numbering 25, are the majority owners, and in exchange for an equity stake, and percentage of profits, each makes a financial investment and agrees to regularly provide de-identified, complete patient medical records. Truveta raised $195 million in a Series A financing, to which its members contributed, completed in mid-2021.

The Truveta platform consists of billions of clinical data points, uploaded automatically and daily by each member into its own dedicated silo, which Truveta has branded as an "Embassy." That way, the members retain control over access to their own data, as well as ownership of that data. The fully integrated offering was recently branded by Truveta as Truveta Studio (see *Figure 1*). Truveta also obtains regularly updated, de-identified claims data and information on SDOH (social determinants of health) from a partnership it has with LexisNexis—adding to the comprehensiveness of the data mix.

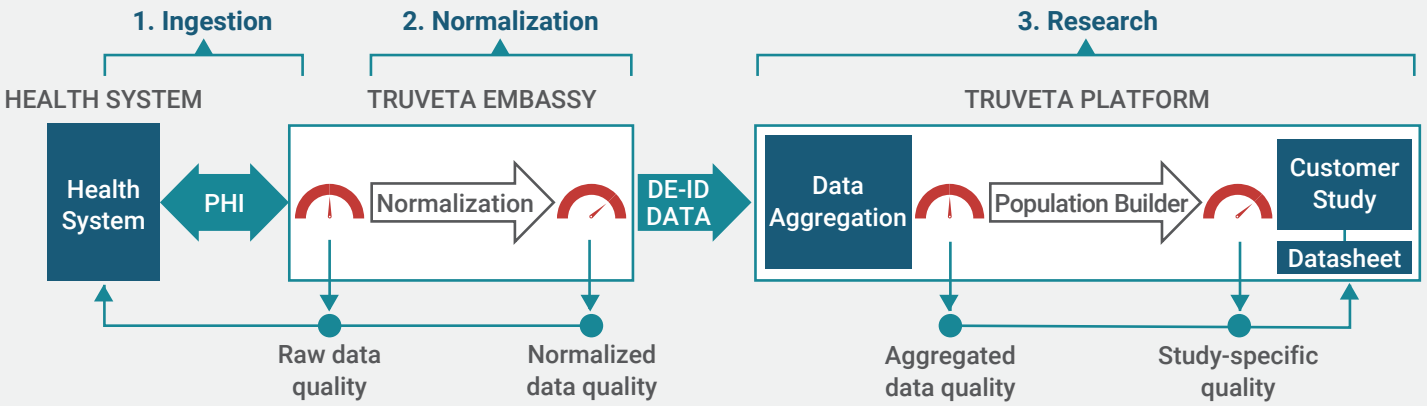
Truveta then curates and de-identifies each member's dataset in a highly structured, extensive process that it describes in white papers that it publishes online for transparency—obscurity about curation is a common complaint in the data field. The steps involve normalizing, de-identifying, and evaluating the quality of the data submitted by healthcare systems, then storing it in separate silos, and the devil is in the details (see *Figure 2*).

Normalization is a key component; Truveta's aggregator maps unstructured content to clinical ontology standards such as LOINC for lab tests and GUDID for medical devices. The de-identification process—another critical sticking point for healthcare systems that are considering sharing data—is elaborate and follows HIPAA Privacy Rule standards, which the company also has outlined in detail and published on its website, as part of its effort to bring transparency to the process. Truveta relies on a multi-step de-identification process that covers millions of medical records while also safely merging related data from LexisNexis and other partners.

As part of the process, it removes personal patient information (PPI) from data sets to minimize the risk of connecting that data back to individuals and applies tokenization from LexisNexis *Gravitas* to block the potential for users to re-identify already de-identified data—one of the most concerning hurdles that has hampered data sharing until recently. Tokenism is used across industries—not just in healthcare—to transform what

Figure 2

Truveta's Approach to Data Quality for Researchers



Points along Truveta's data pipeline that measure data quality. Metrics generated at the Embassy level are sent to health system members to support ongoing data quality improvements. Metrics on the platform are reported to researchers to inform studies.

Note: PHI=Protected Health Information
Source: Truveta, "Our Approach to Data Quality," White Paper, Spring 2022

is known as "direct identifiers," which are metrics that can identify specific individuals, into random strings of numbers and letters, which cannot be reversed to reveal the identify of a previously de-identified patient.

Truveta is now opening access to its data to life sciences companies and researchers, including medical device companies, with goals of enabling them to conduct faster, more nuanced research at more affordable costs. "A lot of companies are working on healthcare data-driven solutions, but we are the only company that is trying to answer questions about the comparative efficacy of drugs and devices on a large-scale, AI-driven basis," says Myerson. "Revenue cycle management, risk assessment, and quality metrics are all being calculated by these companies, but the technology to structure unstructured data rapidly, in high volumes, and the kind of algorithms we are using are all new advances, made in the last two to three years."

For medical device companies, the opportunity to gain insights from large data sets that can link their devices with updated clinical longitudinal patient records to pinpoint 30-day infection rates, hospitalizations, and other metrics, has not previously existed, says Ahern, whose career has mixed business consulting with medicine. Prior to joining Truveta, Ahern was head of corporate development and data strategy at Clarify Health Solutions and also had held faculty positions at New York Presbyterian and Washington University School of Medicine.

The medical device industry has historically relied on purchasing market share data from GPOs, which did not provide outcomes or safety results of treatment, Ahern continues. Claims data can provide additional information, but it is not manufacturer specific and often lags in timeliness. The only way for device companies to get brand-level data linked to clinical outcomes has been to enter into one-off partnerships with large healthcare systems, but those have insufficient scale to meet today's needs and takes months or years to analyze.

A Chance to Gain Greater Insights into Specific Devices

For Boston Scientific, Truveta offers a chance to study patient care and outcomes at a more cost-effective, faster, and more accurate level, says Jaff. The breadth and quality of data supplied by Truveta—covering 72 million patients who interact with more than 20,000 clinics and 700 hospitals in 43 states to date—is unprecedented, he continues. "This is a much broader population than the one Boston Scientific usually studies, as it covers all patients treated at Truveta member hospitals, whether they are covered by commercial or government insurance, or are uninsured."

A small, multi-disciplinary team at Boston Scientific is piloting the Truveta initiative, which Jaff says could have "incredible" potential and very limited downside risk. Truveta, in turn will get early feedback from Boston Scientific on the accessibility, utility, and accuracy of its platforms. "Clearly, a lot of work was done before we moved forward on this project to ensure that patient records were exactly as Truveta promised and completely de-identified," Jaff adds.

Truveta's leadership team also shares a mutual interest in one of Boston Scientific's long-time priorities: exposing disparities in US healthcare and creating opportunities to ameliorate these differences. Initial work is focusing on peripheral artery disease (PAD), a field in which Boston Scientific is a leader.

PAD is a common condition that causes narrowing of the arteries in the lower extremities and affects approximately 6.5 million people aged 40 or older in the US, according to Truveta. While disparities in PAD outcomes following revascularization are well known, large national studies that provide a better understanding of the disease in different populations have not been done; studies that have been conducted were limited by sample size, geography, and timeliness of data.

An initial study, published by Truveta in a white paper in October, confirmed earlier findings on disparities and added more detailed information on the extent of the gaps and subsequent outcomes. It found, for example, that Black and Hispanic patients diagnosed with PAD had lower rates of revascularization procedures and receiving drug-eluting stents—considered to be the most effective care—in fewer cases, compared with their White counterparts.

There are guardrails for Boston Scientific and other Truveta partners: the data can provide manufacturer-specific information, but it can't be used to market products to specific doctors, although it can be used to identify regions of the US that are under- or overdiagnosed with PAD, and to document adverse consequences of those disparities. Findings can also be used to focus strategies and for publications.

Boston Scientific is still evaluating how useful the information is, whether it is better than current data sources, and whether the data can accurately answer the specific questions that are being asked, Jaff continues. And while Boston Scientific evaluated the data protection aspects of the relationship in great detail before entering into the relationship, he adds "Healthcare systems [and companies] can never actually relax about breach of confidentiality of their data." MTS

Posted on MyStrategist.com Nov. 17, 2022