

TECHNOLOGY INTEGRATION OBSERVATIONAL DATA: GENERATING NEW KNOWLEDGE FASTER

■ By John S. Toussaint, MD

In this article...

Observational databases such as Cosmos can increase the pace with which we gain knowledge about certain aspects of patient populations.

THE TIME HAS COME TO UNLEASH THE TRUE

power of the electronic health record. The potential to use the massive database that is accruing from the aggregation of patient data has long been a matter of speculation, but now it is here and it's called Cosmos.

Epic Systems' Cosmos is a clinical database that combines millions of patient records maintained to improve the health and lives of people everywhere. Users can query population health trends, compare like patients for trends in appropriate treatment, and potentially perform patient-level research.

Observational data are important, particularly in studies of living organisms (both functional and behavioral), our planet, climate, and the universe at large. The medical data we input from an electronic health record is based largely on observations. Such data result from processes that cannot be repeated but can reveal many important details about patient conditions.

Cosmos is an observational database. Pre-pandemic, it included 10 million patient records; however, in 2020, health-care leaders realized how important these data were to understanding overall patient trends, so they approved the transfer of their patient data into Cosmos.

By the end of January 2021, more than 90 million records were added to Cosmos, making it the largest comprehensive *clinical* database in the world. Within the next few years, the number of de-identified patient records in Cosmos may grow to more than 200 million.

RESEARCHING TRENDS FOR POPULATION HEALTH

Observational databases can increase the pace with which we gain knowledge about certain aspects of patient populations. Before Cosmos, researchers spent a significant portion of their time gathering data. Now they can start with data that are already mapped and normalized across patient populations and organizations.

For example, David Kaelber, MD, a pediatric and internal medicine physician and researcher at MetroHealth, a multi-hospital system in Cleveland, Ohio, was concerned that in March and April 2020, fewer patients were coming to the emergency room for life-threatening illnesses like stroke. He postulated patients were afraid of being infected with COVID-19 by seeing caregivers.

Kaelber queried the MetroHealth electronic health data, but the numbers were too small for him to make any conclusions. When he logged into Cosmos and queried the larger database, he discovered the overall incidence of embolic stroke across the country had dropped by more than 20% from the fourth quarter of 2019 to the second quarter of 2020 — from 14,801 cases to 11,281 cases. The same trend was true for heart attacks. Other studies revealed similar trends in cancer screening. Observational data also helped determine that anticoagulation led to better outcomes in COVID-19 patients early in the pandemic.¹

Many of these findings have been published in the *Epic Health Research Network* journal,² an open-access online journal where researchers can publish their findings using Cosmos data. Cosmos allows practicing clinicians and researchers to ask questions about entire populations, leading them to design better processes and support planning and capacity decisions at the point of care.

COMPARING PATIENTS LIKE MINE

The Cosmos database can be especially powerful for clinicians who see patients with rare diseases. They can research the treatment of rare diseases in a couple of hours instead of days or weeks. Using a tool currently in development called “look-alikes,” clinicians can enter clinical information about patients with unusual presenting signs and symptoms and determine if other clinicians have seen a similar case. They can then connect to discuss these difficult cases. Understanding treatments that clinicians have tried previously leads to better-informed decisions about how to treat.

Another tool in development is “best care for my patient.” Clinicians can compare a patient’s symptoms with those of tens of thousands of other patients with similar signs and symptoms. For example, a provider can compare a patient with rheumatoid arthritis with wrist and knee inflammation and Felty’s syndrome with an enlarged spleen and low white count to thousands of other patients with the same presentation. This decision tool allows the provider to choose a specific outcome such as joint preservation or white blood cell count level and determine which treatment modality achieves that outcome most effectively. The information leads to a more informed decision about the best treatment plan for individual patients.

Although the user interfaces for these two functionalities are still in development, during the next few years, these tools will allow powerful new ways for clinicians to determine how to manage not only complex care decisions but also common conditions. With just a computer, providers can perform a query in real-time while in the exam room with a patient or during a telehealth session.

POTENTIAL FOR INDIVIDUAL ANALYSIS AND RESEARCH

Although Cosmos is a secure database with patient identifiers removed, researchers need to run a regression analysis, study multiple variables, and look longitudinally at outcomes. They need the entire individual Patient Health Information file (PHI). Of course, providing this information is challenging due to many issues, not the least of which is HIPAA, but researchers can band together and study their own patient data much as they do today, if done so in appropriately designed studies.

Human subjects research must be approved by ethics committees/Institutional Review Boards (IRBs) and follow strictly enforced standards. Those same standards could apply if organizations wanted to pool their PHI data in Cosmos. One way to do this would be to “bring people to the data.” In other words, rather than sending the data files to the researchers, the researcher would access the data in Cosmos with a secure

access code. The data would be a read-only file. Examples of this type of data use for research includes access to the CMS Virtual Research Data Center.

ACCESSING THE DATA TODAY

While the Cosmos database has been in the testing phase for the past few months, superusers like Kaelber have been accessing it to answer questions like those discussed above. Queries such as “Has the encounter been closed or not?” have led to a cleaner data set. So far, the data seems to be fairly clean (at least as clean as clinical data can be), meaning clinical comparisons are accurate. An embolic stroke is an embolic stroke, not something else.

The goal is for every organization that submits data to have one superuser and a core group of people who will have access to the data. The superuser will be the lead in interacting with Cosmos staff; however, the others in the organization will have full access to the de-identified data.

Database use is a complex affair. Keeping the number of users small will allow everyone to understand how to use the database and what it can and cannot be used for. Access will expand over time.

CONCLUSION

Aggregating patient information to improve the health of individuals is a noble goal. Despite the discussion around big data, so far there has been little to show for all the hype. We now can imagine the day when we voice-activate the EHR; input the patient signs, symptoms, and test results; and then compare hundreds of millions of other patients to the person in front of us to determine the best treatment — all within a few seconds.

The future of medicine is bright, and Cosmos is helping to illuminate the way.



John Toussaint, MD, is executive chairman of Catalysis, Inc., a nonprofit education institute focused on training physicians and healthcare executives on organizational excellence. He is also a non-executive director on the Epic Systems Inc., board, the Shingo Institute board, and the Lean Enterprise Institute board.

jtoussaint@createvalue.org

@JohnToussaintMD

REFERENCES

1. Mast C and del Rio AM. Delayed Cancer Screenings—A Second Look. *Epic Health Research Network Journal*. July 17, 2020. <https://ehrn.org/articles/delayed-cancer-screenings-a-second-look>
2. Faulkner J and Toussaint J. Industry Voices—Epic’s Faulkner and Toussaint: Observational HER Data Are Helping Combat COVID-19. *Fierce Healthcare*. www.fiercehealthcare.com/tech/industry-voices-epic-s-judy-faulkner-and-john-toussaint-observational-ehr-data-powering-fight