

Big Data Analytics v Clinical Data Registry: Which Is Best for Performance Improvement?

written by Theresa Hush | September 30, 2015



A few weeks ago, Epic announced that it would create an unidentified database of patient data for customer research, with a plan that the [“Cosmos Research Network”](#) will leverage data for better decisions.

There’s been a flurry of activity in recent years to create Big Data in health care. The Clinical Data Registry (CDR) is a variation of the concept, soon to be populated by data from Meaningful Use public reporting requirements.

The question: Can both these data engines help performance improvement? The answer depends on a key distinction between Big Data Analytics and the CDR, with important implications for how providers should use these tools.

Big Data Analytics: De-Identified for Analysis of Volume and Trends

The concept of Big Data is intriguing. Even the National Institutes of Health is sponsoring an initiative to provide better access to Big Data for research. Health systems and academic hospitals are also creating data repositories for research. Many believe that just aggregating and analyzing the data will be enough to find the answers to our cost and quality issues. Whether this belief is valid must be examined more closely to determine if Big Data can improve performance.

Although Big Data in health care can have several definitions, for our purposes here we are distinguishing between de-identified, aggregated databases built from various sources of data

(e.g. hospital, physician groups, lab and multiple sources of each type), which is true “Big Data,” and databases that maintain a form of identified patient information along with the clinical and transactional data, which is the basis for most Clinical Data Registries. Epic’s Cosmos Research Network will be Big Data. Big Data Analytics is the primary product that comes out of the database.

Big Data has good capabilities for evaluation and is best applied when examining volume, variables, associations and trends:

- Identifying associations that cannot be seen on a small scale, such as common characteristics of patients with longer hospital stays, readmissions or mortality;
- Dividing populations into cohorts that may be used for research;
- Calculating quantity or volumes of any event, population or disease incidence;
- Identifying characteristics of patients or outcomes;
- Evaluating sequences and timing of services (e.g. ER admit, tests, diagnosis, admission, inpatient events, discharge to home) and thereby creating a basis for some operational improvements;
- Seeing volume trends in services;
- Benchmarking.

But it’s important to understand what Big Data does *not* do well, without identified data:

- Reviewing long term outcomes;
- Creating a longitudinal patient record for cause-effect research;
- Clarifying variability among individual patients and/or providers;
- Adjusting risk beyond aggregate population;
- Creating a foundation for patient-specific improvement;
- Collecting or validating root cause information.

In short, Big Data works well for seeing the big picture. But it is notoriously difficult to use Big Data to engage physicians in their own data or results if they can’t identify patients and can’t explain or challenge the information. And while Big Data is effective for identifying systemic issues in order to change processes to become more efficient, it is almost impossible to use Big Data in performance improvement plans that are focused on moving patient outcomes in a positive direction.

Clinical Data Registry: Database Plus Solutions

CDRs usually collect identified patient data, because most CDR have other purposes that require that information. (We are distinguishing a CDR on this basis, even though there may be

some exceptions.) For example, a Qualified CDR like ICLOPS may perform PQRS Reporting, and this requires identified data. Population health or performance improvement, including risk stratification to set up performance improvement initiatives, is a key CDR function that also requires patient details. Perhaps most important, the validation or patient outreach function of a Registry requires identifying patients.

The CDR can leverage many of the Big Data Analytics assets: evaluating volume, trends, variables and processes. But in addition, it has the ability both to benchmark and then to initiate performance improvement through additional processes or applications. In other words, the CDR is a database plus a series of applications that focus on outcomes. These include outcomes research, population health, performance measurement and a variety of performance improvement initiatives.

What distinguishes a CDR from Big Data Analytics solutions? The CDR deploys actionable analytics and simultaneously offers the instruments to improve individual patient outcomes:

- Outcomes research with identified patients and all their data, taking into account a variety of variables such as utilization, disease history, comorbidities, lifestyle factors, socioeconomic factors and even belief systems;
- Longitudinal patient research facilitated by the use of all patient data, not a cross-section, and cause-effect research;
- Performance measurement with risk adjustment or stratified populations, then solutions to fill the gaps in care through easily deployed outreach technology;
- Validating and/or adding data from patient feedback into the database to improve the quality of outcomes information and determine why;
- Testing of interventions to improve outcomes over time among patient cohorts;
- Performance improvement projects to identify optimal post-acute care options.

Not all CDRs perform all of these functions, however, with variability among Registries. Here are some of the issues that cause limitations for CDRs in this newly emerging model:

- Availability of patient-identified data, which currently is obtained as a byproduct of the CDR's offering of specific solutions, such as PQRS Reporting;
- Variation in CDR focus and distribution, with very different business models and missions.

The CDR is only just developing as more data becomes available and business opportunities grow. It is vulnerable to business and revenue cycles, which could be problematic for long-term analytical functions. Time is needed to grow and mature the model. Experts and resources will be needed to develop it as a viable research tool with business applications and to increase

provider understanding about this powerful performance improvement tool.

CDR or Big Data for Your Future?

Clinical Data Registries and Big Data Analytics are different ends of the patient database spectrum. Big Data focuses on larger trends and can reveal or suggest important information about operations. When detail is needed, the CDR offers more answers and solutions.

Developing technology always emerges along a continuum. There are Analytics companies that can produce patient-level detail and Registries with top-tier Analytics. The distinction is in range and focus, and this industry will continue to mature and develop.

Health systems will need both Analytics and Outcome-Focused Solutions in the future, and, of course, we weigh in on the CDR being the comprehensive answer to Performance Improvement. Health care developments predict that the risk environment for providers will put a heavy premium on action in addition to knowledge, and the CDR will have the tools to power that process. And, once real research starts happening through these databases—as it should and will—the tracking of individual identified patients will be essential.

[Download your free copy of the ICLOPS Insider's Guide to PQRS 2015 Reporting: How to Succeed in the Value-Based Health Care Environment.](#)

Founded in 2002, ICLOPS has pioneered data registry solutions for improving patient health. Our industry experts provide comprehensive [Solutions](#) that help you both report and improve your performance. ICLOPS is a CMS Qualified Clinical Data Registry.

[Contact ICLOPS for a Discovery Session.](#)

Image Credit: [Peter Skadberg](#)