



Medical Informatics

Mining and organizing data to improve healthcare quality is quickly becoming a reality, and it is big business.

By Amy Scanlin, MS

Improving healthcare at a reduced cost has long been a goal of the medical community. One way of accomplishing this is with medical informatics, or the use of vast quantities of data aggregated into usable searches, organized results and, in turn, improved procedures. In the coming years, medical informatics is anticipated to be a breakthrough strategy yielding higher-quality healthcare that can help to improve patient care, prevent medical errors and reduce costs — all with a

query and a push of a button. And, it has never been more exciting and, in some cases, closer with the increased usage of electronic health records (EHR), the ability to pull data from health plans and health systems, the growing area of genomics and more.

The latest push for medical informatics came about when the *2012 National Healthcare Quality and Disparities Report* (published each year as mandated by Congress to focus on

national trends in the quality of healthcare provided to the American people) indicated that healthcare quality and access are suboptimal, particularly for minorities and low-income individuals. In response, the U.S. Department of Health and Human Services called for an action plan. That plan, in part, is the need for improved data access for the underserved (which is often incomplete or collected in too low a quantity to be meaningful) to improve healthcare quality.¹

But, for medical informatics to work, policy decision makers, researchers and end users need to determine how best to collect, store, compare and improve upon digital data collection, analytics and reporting. Some challenges of how best to implement policy for maximizing comparative effectiveness research include data quality, data representation, data completeness, data timeliness, governance, technology, privacy, sustainability and issues of workforce development.²

Data Collection

The term “data” and what it encompasses has undergone an evolution, according to the American Medical Informatics Association (AMIA). The field of genomics, the methods of physician, device and hospital reporting, and even individual self-reporting, all factor into the very broad term that researchers are using to improve health outcomes of patients. What’s challenging, though, is that there is no single national healthcare database or clearinghouse of information from which to sort. Instead, data is collected in a variety of ways and in a variety of databases, and each provides estimates for the populations for which they serve. For instance, data is collected via health plans, health systems, inpatient, outpatient and emergency departments, and others.

To help solve this challenge, metadata registries are now being used. Metadata registries are a way of collecting data without collecting the actual data itself. They store data elements that include both semantics (the meaning of a data element with precise definitions) and representations (the definition of how data is represented in a specific format).

One such registry is the Agency for Healthcare Research and Quality’s (AHRQ) U.S. Health Information Knowledgebase (USHIK). Prior to the formation of USHIK, a major barrier to EHRs was a lack of standardization of codes, even for things as simple as gender, marital status and race. Getting that content right, so that researchers looking at data could compare apples to apples, was of the utmost importance. The creation of USHIK was a joint effort by the Centers for Medicare and Medicaid Services (CMS), as well as other agencies, with CMS taking the lead, working to meld the informatics initiative with the standards of HIPAA to find the right vocabulary that makes content understandable at both ends of the transmission process.

Determinations had to be made for which aspects would be addressed in hammering out the patient safety common format standards, such as “their definitions, their names, how they are represented (their code sets), the vocabularies in which they come, the base standard from which they are derived, and the organization that maintains the particular vocabulary or the code set such as for ICD-10 in the United States,” explains Michael Fitzmaurice, PhD, senior science adviser for information technology at AHRQ. (In the U.S., the National Center for Health Statistics maintains the ICD-10 code set.)

In the coming years, medical informatics is anticipated to be a breakthrough strategy yielding higher-quality healthcare that can help to improve patient care, prevent medical errors and reduce costs.

While data is not input directly into USHIK, says Fitzmaurice, there are “several organizations that input intellectual property about the data into the database, and [there are] a handful of states that input information about their data into USHIK.” These include standards developing organizations, federal and state organizations, specific harmonizing initiatives and others.³ “We have found a unique way of meeting those who provide intellectual property head on in a way that satisfies their needs,” adds Fitzmaurice. And, the voluntary reporting of their data is protected in court so “while it is proprietary data, we do have good success in obtaining the data dictionaries. They have freely given us the data dictionaries, and we put them into USHIK. They could pull their data out at any time if we didn’t treat it right. But, we do good things with the data, and so far they like it. It is mutually beneficial. Researchers can turn out robust findings.”

Once data is collected, it has to be analyzed through algorithms to provide information on health trends, readmissions, healthcare costs and any number of subjects a user would like to search. “Congress charged AHRQ to look at patient safety, and we focused on hospital reporting standards in common formats, what questions were being asked and what answers

were supplied, and how they should be coded,” explains Fitzmaurice. “We supplied the formats, hospitals provided the patient safety data, and we combined them looking for commonalities — kind of a ‘daisy chain.’” They then put the name, definition and attributes about the data into USHIK so the user can view side-by-side comparisons to see how they compare or how they don’t compare. This is particularly useful to researchers because, when combining the actual data, they need to know that the data means the same thing — particularly the same thing as the concepts they are investigating.

Applications of Data Use

There are many examples of how researchers, administrators and clinicians are making use of the information contained in health informatics databases for improved patient care and improving costs.

A newer area of study for AHRQ is the Multiple Chronic Care Research Network (MCCRN), which is looking at patients with multiple comorbidities.

USHIK. Those who rely on USHIK include EHR vendors, state public health departments, physicians, researchers, developers and policymakers. One example of how USHIK is being used is with meaningful use stage II (which must be met by EHR vendors in order to continue to participate in the Medicare and Medicaid EHR Incentive Programs). USHIK contains the clinical quality measures, how they are calculated, the data elements used in the calculation, and the codes that the data elements can use such as ICD-9 or ICD-10 codes and SNOMED code for something like diagnostics, says Fitzmaurice. With this, an EHR vendor may want to compare what is in the EHR data dictionary with what needs to be in the data dictionary to produce the Medicare and Medicaid clinical quality measures for the incentive payment program. Or, there could be a vendor that may want to compare what is required for certification so that their EHR can produce the clinical quality measures for meaningful use stage II. Those vendors “can come to USHIK as a one-stop shop and get the clinical quality measures,” explains Fitzmaurice.

AHRQ Common Formats. The Quality Assessment and Performance Improvement (QAPI) Act of 2005 was implemented to require hospitals to track adverse patient events. However, recent reports by the HHS Office of the Inspector General (OIG) indicated that hospitals fail to identify most adverse events. In response, the OIG recommended that AHRQ and CMS help hospitals improve their ability to track adverse patient safety events by disseminating information on AHRQ’s Common Formats. The AHRQ Common Formats define a systematic process for reporting adverse events, near misses and unsafe conditions, and allow a hospital to report harm from all causes. While hospital use of the AHRQ Common Formats is voluntary, CMS recently stated in a memo that hospitals that use them and are adept at the analysis that they permit will be in a better position to meet the QAPI requirements.⁴

Advanced analytics for analyzing hospital readmission rates. Hospital readmission rates became an even more critical financial hurdle to overcome in 2012 when the government began withholding 1 percent of base Medicare reimbursement from hospitals with excessive readmissions, and again in 2013, when the penalty climbed to 2 percent, with an expected 3 percent in 2014. But, according to a report by Health Data Management, “by leveraging advanced analytics, organizations can identify which conditions are the best candidates for quality improvement initiatives.” The analytics “would take into account the cost of the interventions required to have an impact on readmission rates compared with the total revenue reductions that a hospital would experience if readmission rates land them in the bottom quartile, subjecting the hospital to financial penalties as prescribed by the Affordable Care Act.”

Looking at the cost of interventions that could prevent readmissions and comparing those with the cost of the readmissions, as well as the penalties assessed, analysts can determine which conditions have the highest readmission rates, for whom and which interventions can be implemented with success and, in turn, where their resources can be best spent for the greatest return in reducing readmission rates — a kind of proactive readmission approach.⁵

Multiple Chronic Care Research Network. A newer area of study for AHRQ is the Multiple Chronic Care Research Network (MCCRN), which is looking at patients with multiple comorbidities. “It’s a big initiative across HHS,” says Richard Ricciardi, PhD, RN, health scientist at AHRQ. “We are looking at ways to improve quality, the patient experience and value. Two-thirds of all claims data are related to multiple chronic conditions (MCC). These are high-utilizers with high costs.”

Questions that AHRQ is seeking answers to are: How is it best to treat these patients? Which is the most important

disease to start with? How will the medications interact with each other? And, “How can we best approach this as a team?” explains Ricciardi. Working with patients with MCCs is a complex process, not linear, and clinicians need to know “how best to improve outcomes for the patients, improve their quality of life and, in layman’s terms, get the biggest bang for the buck.”

In 2010, AHRQ began looking at comparative effectiveness and infrastructure studies to develop guidance to facilitate a research network. One question they had was how they could build the infrastructure so that research can be better conducted. They paired databases in an effort to enable more exploratory research, and linked research papers to collectively start putting it all together.

MCCRN looks at various tools, including information technology, for improving patient outcomes. One example is a study by Dr. Henry Fisher, at the Denver Health and Hospital Authority, who conducted an interventional study on bidirectional text messaging for diabetic patients. Patients were reminded by text message to send in their data, and the data they sent in via text was reported to their healthcare team. Many patients found that it helped them to keep their diabetes under control because they felt like someone cared and was reaching out to them. “It’s a good reminder that mobile technology has potential to improve outcomes,” says Ricciardi.

Another area under review is building databases for dual-eligibles or those with mental health issues as part of their comorbidity, and how the healthcare team can help these patients stay on track with taking their medications and other health improvement interventions.

MONAHRQ. MONAHRQ, or My Own Network, powered by AHRQ Learning Network, is a free tool that uses hospital discharge data for measurements and comparisons by hospitals, communities, counties and states. Researchers can look at quality of care for emergency room, inpatient and outpatient settings; quality ratings and avoidable hospital stays by specific conditions; procedures and hospitals; as well as financials of specific hospitals or by county. Users also can compare data with results found in Hospital Compare, a website hosted by CMS.

“MONAHRQ uses state discharge data sets, which are pretty uniform across the states as far as the date, procedure and diagnosis code,” says Susan Schow, Pathways to Excellence program director at the Maine Health Management Coalition. Schow was previously at the Maine Health Data Organization and was instrumental in getting MONAHRQ up and running in the state. “It is great! You can take all the data sets and query, without having any special data language, program or query code. It is wonderful for small programs with limited funds but great needs. It can drill down to a very discrete level.”

Not only does MONAHRQ offer functionality today, but

that functionality can power new ideas and better capabilities. Just thinking off the cuff, Schow adds that MONAHRQ would be useful as a tool for verifying required event reporting. “The more people who are aware of this, the more widely it can be used to help inform on healthcare — to make policy,” she explains. “It’s free to anyone — researchers, policymakers, hospitals — and it is great because you don’t have to pay to analyze data.”

Certainly, a big question when it comes to data mining is how to protect patients’ privacy when their personal information is collected with the data.

“It became very clear when we saw the functionality of these data streams that we needed to transform ourselves,” says Karynlee Harrington, executive director at the Maine Health Data Organization. “We need to take that cost data and expand it to join with quality data. We want to go a lot further than we originally planned! I have a vision of taking the MONAHRQ cost information and querying it down to hospital information in a geographic area, and then query the utilization out of it.”

Patient Privacy

Certainly, a big question when it comes to data mining is how to protect patients’ privacy when their personal information is collected with the data. In many states, patients have an all-or-nothing opt-in or opt-out option of providing their health information to data warehouses. In other states, there is more control on what information is shared and how.

A small study conducted at the Weill Cornell Medical College in New York found that while most patients support the idea of sharing health information, “78 percent would prefer to explicitly approve the sharing of all types of information, and most prefer restricting information by clinician (83 percent), visit (81 percent) or information type (88 percent).” Another study found that 70 percent of patients are either somewhat or very concerned about the privacy of their medical information in light of the new health information exchanges (HIEs). Even physicians who feel that HIEs are valuable tools are concerned about privacy. As more states launch HIEs, the issue of privacy and how best to protect it will be extremely

important. Federal and state agencies will need to address their privacy policies and technical standards.⁶

“Privacy is a big issue,” says Dave Page, PhD, a professor in the Department of Biostatistics and Medical Informatics and the Department of Computer Science at the University of Wisconsin at Madison. Through medical informatics, “we are combining data from many different sites, and each hospital and provider owes [its] patients privacy.” While data is de-identified, in some cases, “noise” is added or data is changed just slightly to further protect patient data. “If you jiggle the data a little bit, then you can’t distinguish one patient from another,” explains Page. In bigger models, the idea is “differential privacy,” or the idea of de-identifying patients and only asking certain questions of the database so that the data would not be significantly different whether or not that person was included. Page cites the well-known case in which Harvard professor and researcher LaTanya Sweeney was able to re-identify Massachusetts Governor William Weld from his anonymous hospital discharge records as a high-profile example of the importance of ensuring the privacy of data.

Clinical Informatics Subspecialty

The schooling of health informatics is big business in and of itself. In 2011, the American Board of Medical Specialties approved clinical informatics as a board-certified medical subspecialty. The informatics specialist is able to determine which data will be most useful in clinical decision-making, and how that data should be delivered to physicians. In an article published in *American Medical News*, AMIA’s Board of Director Chair Nancy Lorenzi, PhD, stated: “It is entirely appropriate and timely to certify clinical informatics as a specialized area of training and expertise in an era when more and more clinicians are turning to data-driven, computer-assisted clinical decision support to provide care for their patients. Clinical informatics blends medical and informatics knowledge to support and optimize healthcare delivery.”¹¹

“The NLM [National Library of Medicine] has been funding education programs on medical informatics for over 40 years,” says Page, whose program at the University of Wisconsin at Madison has been supported by NLM for 12 years. In fact, it is such a growing trend that it is estimated that 50,000 new medical informatics specialists are needed to meet the growing demand.⁷ “Our students come from a variety of majors,” adds Page, “from computer science, biochemistry, MDs, nurses and statisticians. It’s a very large field and goes well beyond data analysis.”

The Future of “Big Data”

“I’m trying to think about what this field will look like in five to 10 years,” says Page. “I’m excited by the predictive models and what they mean for personalized medicine. I’d like

to be able to predict who is most at risk for a heart attack, diabetes or cancer so that we can take action early. But, there is also the problem of adverse drug events. Is a new drug causing some specific subset of the population an adverse reaction that we didn’t see in clinical trials because we were only looking at 1,000 people? Can we predict who is most likely to have an adverse event, and can we also identify for whom this drug is going to perform? If we can, we can incorporate these findings into EHRs so they can build predictive support. Then, we can build in pop-up alerts that this patient is at an increased risk for a heart attack, etc. For the most part, we are not at that point yet. [But,] as we collect more thorough genomic sequencing, data will greatly improve our ability to produce better medicine. I think that cancer research is where we’ll see the biggest impact both short and long term. By genotyping a tumor, we can see the result of the disease, and companies are already looking at this.”

“We are proud of this work,” adds Ricciardi. “It requires future thought as to where we are going and how best to integrate all this care in a way that uses the right provider at the right time to improve the healthcare delivery system. Science administration is not easy, and it is important to invest our public’s money wisely and engage the right people to provide the right guidance.”

With an emphasis on “big data,” medical informatics is changing the way we are thinking about making decisions. Says Page: “With personalized medicine and predictive analytics, we can do a better job of improving outcomes for patients.” ♦

AMY SCANLIN, MS, is a freelance writer and editor specializing in medical and fitness topics.

References

1. Agency for Healthcare Research and Quality. 2012 *National Healthcare Quality Report*. Accessed at www.ahrq.gov/research/findings/nhqdr/nhq12/highlights.html.
2. American Medical Informatics Association. *Health Policy Meeting Briefing Book*. Accessed at www.amia.org/sites/amia.org/files/Invitational-2012-Health-Policy-Meeting-Briefing-Book.pdf.
3. U.S. Department of Health and Human Services Agency for Healthcare Research and Quality. USHIK 101 — A Quick Course in Metadata and Metadata Registries. Accessed at ushik.ahrq.gov/new_user.jsp?faq=q1&enableAsynchronousLoading=true.
4. Centers for Medicare and Medicaid Services. AHRQ Common Formats — Information for Hospitals and State Survey Agencies (SAs) - Comprehensive Patient Safety Reporting Using AHRQ Common Formats, March 15, 2013. Accessed at www.cms.gov/Medicare/Provider-Enrollment-and-Certification/SurveyCertificationGenInfo/Policy-and-Memos-to-States-and-Regions-Items/Survey-and-Cert-Letter-13-19.html.
5. Health Data Management. Leveraging Advanced Analytics to Successfully Reduce Hospital Readmissions. Accessed at himss.files.cms-plus.com/FileDownloads/106420_0213.pdf.
6. Dolan, PL. Patients Seek Strong Say in How Their Data Are Used. *American Medical News*, Oct. 3, 2012. Accessed at www.amednews.com/article/20121003/business/310039997/8.
7. Dolan, PL. Clinical Informatics Now a Subspecialty. *American Medical News*, Oct. 10, 2011. Accessed at www.amednews.com/article/20111010/business/310109962/6.