Mark L. Braunstein

Practitioner's Guide to Health Informatics

Springer

Practitioner's Guide to Health Informatics

Mark L. Braunstein

Practitioner's Guide to Health Informatics

Springer

Mark L. Braunstein
College of Computing
Georgia Tech
Atlanta, GA
USA

ISBN 978-3-319-17661-1 ISBN 978-3-319-17662-8 (eBook)
DOI 10.1007/978-3-319-17662-8

Library of Congress Control Number: 2015936165

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

As a teacher I try to never forget the impact that we can have—sometimes in just a single conversation—on a young person's life.

Dr. David M. Kipnis died a few months before I began writing this book. He was chairman of the Department of Internal Medicine at Washington University. Toward the end of my incredibly demanding internship he unexpectedly summoned me to his office. To say the least, I was quite nervous. We had little interaction that busy, often hectic year during which I was subsumed by the care of very sick patients, often basically on my own and in a major hospital setting.

He immediately asked something like: "Do you know why you were accepted to this program?" I had no answer (and feared I was about to hear something awful about my performance). He went on to explain that he, personally, had selected me because of my highly unusual combination of interests in medicine and computing, something he said would be very important over the coming years.

That comment particularly struck me because I never had an interest in medicine until, several years earlier, Dr. Leroy S. Lavine, a prominent physician and my cousin by marriage and whom I greatly respected, asked me what I planned to do after graduating from MIT. It was the sixties, a crazy time, and I told him I had no idea. He advised me that my strong interest in computing should be directed toward medicine because it would be an important and growing field over the course of my career. Even having never previously considered medicine, I followed his advice.

Kipnis then went on to say that I was not a traditional candidate for a prestigious program like Washington University—I'm sure it was clear from my medical school record that I had spent far more time programming than learning to be a doctor—and that his concern about accepting me had been whether I could actually be a good doctor. I imagine I was pale white by then, so he let me off the hook by saying I had actually done very well, and urged me to continue, even suggesting I could earn a faculty appointment in his department to pursue medical informatics, once I was done with my training.

Similar to the conversation with Dr. Hiram Curry which I'll relate in the Introduction, I told him I was honored by his interest in me, but I wanted to take a year off to see if I could finish the work on the pharmacy system I had developed at MUSC. His response was something

I'll also never forget. It was something like this: "I'd like to see you stay in medicine but no one achieves a great deal in life unless they pursue their real passion."

One year became two, and I never went back. In truth, health informatics was my passion and I owe it to Drs. Kipnis, Lavine, and Curry for helping me figure that out. As a result, I am honored to be able to dedicate this book to them.

PS: In my current role at Georgia Tech, I'm often approached by students (sometimes medical students from nearby Emory) with similar conflicts about their future direction and I invariably tell them the Kipnis story and give them the same advice he gave me so many years ago.

Why I Wrote This Book

As a medical student at the Medical University of South Carolina (MUSC) in the early 1970s I fell under the spell of the late Prof. Hiram B. Curry, a former general practitioner who said that job was so hard that he went to Harvard to study neurology! Years later, he founded MUSC's cutting edge academic department of family medicine. I needed a summer job and he was looking for students to help find families for his new clinic, so I arranged a Friday afternoon interview. After he described the job, I gathered the courage to say that I wanted to do something else—computers in medicine. Instead of laughing, he gave me a copy of Dr. Larry Weed's then-new book, *Medical Records, Medical Education and Patient Care*. I read it—twice—over the weekend. Returning to his office Monday morning, I said excitedly that Weed was right and computerized problem-oriented medical records were the future. Over the next few years we developed one of the first fully operational ambulatory electronic medical record (EMR) systems.[1,2] Today it might even be described as an electronic health record (EHR) because it encompassed virtually all of a patient's care. With the advice and counsel of Dr. William Golod, Dean of the MUSC School of Pharmacy, and John D. James, RPh whom he brought in from industry to run our dedicated, on-site pharmacy, we developed a particularly rich subsystem with advanced clinical functionality for the time, including interaction screening and monitoring patient compliance based on refill intervals.[3] We had numerous visitors and the pharmacy component of our system attracted a great deal of interest. With the school's help, two colleagues and I started a company to create a commercial, standalone version of the pharmacy system. Both Kaiser (starting with their Southern California region under the guidance of Al Carver, someone to whom I owe a great debt for taking a chance on a very young, very green entrepreneur) and the U.S. Military Health System

[1]Office of Technology Assessment 1977. Policy Implications of Medical Information Systems. http://ota-cdn.fas.org/reports/7708.pdf.

[2]Braunstein, ML, Schuman, SH and Curry, HB 1977. "An On-Line Clinical Information System in Family Practice," J Fam Pract, 5:617–26.

[3]Braunstein, ML and James, JD 1976. A Computer-Based System for Screening Outpatient Drug Utilization, J of Am Pharm Assoc, NS16:82–85.

(Tri-Service Medical Information System or TRIMIS) installed it successfully, and our tiny company attracted the interest of a much larger, public company that eventually acquired it. As a result, I left MUSC and ended up spending the next three decades or so in the commercial health information technology (HIT) sector.

Since 2007, when my last company was acquired, I've been teaching health informatics at Georgia Tech. In 2012, I published *Health Informatics in the Cloud*, a short guide to the field written with nontechnical readers in mind. Based on it, I developed what may have been the first Massive Open Online Course (MOOC) in the field and, to my amazement, a third of the 20,000 students who enrolled in its two sessions were either physicians, nurses, or other healthcare providers. Many more were in other positions in the healthcare delivery industry. Over this same period of time the U.S. has achieved widespread adoption of electronic records and patient-facing healthcare tools, but these technologies often still have limitations. Many providers are unsatisfied with them and don't feel there are benefits that warrant the pain of learning to use them well. A key reason for this is that they don't talk to each other, so the focus has now largely shifted from adoption (which is where it was when I wrote the earlier book), to interoperability, how to make these systems talk to each other and how to use the digital "big data" derived from them to improve health care through analytics. The clear interoperability "crisis" has spawned, with astonishing speed for health care, the development and even acceptance of "radical" new and far better technical approaches to data sharing.

This convergence of events convinced me there was a need to update the earlier book substantially while maintaining it as a practical guide to the field. What started as a rewrite morphed into a very different book, written more specifically for busy healthcare providers but still suitable for all nontechnical readers. I hope it makes the potential of health informatics in patient care far clearer and more exciting to providers. For all readers, I hope it will provide a sense of where we are on what has been a long journey that still has much further to go. Most importantly, I hope it will excite you to learn how health informatics—if properly conceived, implemented, and used—can help move us to a more effective, efficient, and safer healthcare system.

For the most part, this book is not technical. I've highlighted the sections that do go into technical detail so readers with no interest in that can skip ahead. Doing so should not impede your ability to grasp the key concepts I hope to convey. At the same time, for those of you who want more technical detail, I have provided many links and references to related information that is almost all freely available on the Internet.

Atlanta, GA
Winter 2015

Mark L. Braunstein

Acknowledgments

My Georgia Tech health information technology colleagues, Marla Gorges, Phil Lamson, Steve Rushing, Rudy Snyder, and Margaret Wagner Dahl, provided invaluable help to me to find and correct numerous errors, omissions, and deficiencies in the draft version of this book. Dr. Eric Dahl, Associate Dean for Administrative Initiatives at the University of Georgia, College of Public Health, was kind enough to find time in his busy schedule to carefully review the entire text and provide many valuable comments and suggestions to improve its readability and clarity. Mary Boyd did an excellent job of final copyediting of the text.

Of course, any remaining problems are my responsibility alone.

Contents

List of Figures

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Introduction

After almost two decades of advocacy, the health care system might finally be ready to take full advantage of information technology to improve quality and efficiency.

— Don E Detmer, MD, MA, FACMI, FACS, Issues in Science and Technology, 2009[1]

This book is about *health informatics*—the applications of information technology to *healthcare delivery*. This is distinct from *bioinformatics*, a related field with which it is often confused; bioinformatics is about using computational models and other methods to analyze and understand complex intracellular biochemical mechanisms. The term *biomedical informatics* is often used in medical schools to denote the merger of these two fields into one training and/or research program.

This book is based on the premise that achieving a more efficient and effective care-delivery system depends, at least in part, on the widespread adoption and *proper use* of information technologies.

To appreciate that, it is critical to explore the role of data in healthcare. Healthcare delivery is clearly a very data-dependent activity and yet, historically, healthcare providers have not paid enough attention to recording high-quality data. To my knowledge, this conundrum was first pointed out in the mid-nineteenth century by Florence Nightingale, the founder of modern nursing, but also a social reformer and statistician who is now recognized as a pioneer in evidence-based nursing.[2] It was brought to wide attention again a century later by Dr. Lawrence Weed who, perhaps more than anyone, started the movement that has now led to widespread adoption of electronic records, many of which utilize some form of his problem-oriented medical record (POMR) format.

Despite these calls, the U.S. healthcare system has long resisted adopting the information technologies that are in common use in other industries. Partially as a result, the U.S. has many problems delivering efficient, high-quality care, particularly to patients with chronic disease. Arguably, we spend twice as much on healthcare as we need to. The Organisation for Economic Co-operation and Development (OECD) collects data from over 30 developed countries and its comparisons for healthcare are

[1] http://issues.org/25-4/detmer/.
[2] http://ebn.bmj.com/content/4/3/68.long.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_1

widely cited. For example, although the French outlive Americans, they spend around half of their gross domestic product on healthcare as compared to the U.S.[3] There are many other countries with similar costs and results as compared to the U.S.

The basic reasons for this are clear. As we'll discuss in more detail later on, the U.S. healthcare system is highly fragmented and many patients, particularly those with multiple chronic diseases, receive care from many providers practicing in many organizations. Data sharing to coordinate that care has historically been very problematic. Producing sharable digital data from care delivery—*and actually sharing it*—is arguably the single most important contribution health informatics can make to better healthcare in our country. However, physicians and other care providers receive little, if any, training on information technology or health information systems. I believe that proper selection and use of informatics tools depends on an informed provider community—a major reason for this book. Such an informed community would also drive vendors to produce better, more useable systems that are targeted at the actual needs of providers involved in care delivery.

Given its mission, this book is arranged to tell the informatics story from a healthcare delivery point of view. We begin with background on the U.S. healthcare system and look somewhat more deeply into some of its problems. A full treatment is impossible in a short book, so I provide references and suggested additional readings along the way. From there we discuss the role the federal government is playing to foster adoption of information technology. Its motivation for doing this is that it pays for a substantial part of U.S. healthcare and increasingly recognizes that improving the efficiency and quality of the system through informatics could go a long way toward reducing future healthcare costs. We then discuss the core technologies that all contemporary health informatics systems and tools rest on. We'll do this as nontechnically as possible and from the perspective of what a practicing provider or other nontechnical reader needs to know.

With this background, we'll then explore how health informatics is being used in real-world applications from electronic records for providers and patients, to managing the health of large, geographically dispersed patient populations, to understanding and improving public health. As part of this, we'll consider the usability issues that most commonly lead to provider dissatisfaction with EHRs and we'll look at some promising approaches to overcoming them. All EHRs are not created equal and it is important to understand the nature of these differences in order to choose one wisely.

We conclude by showing how data can be aggregated from these systems and analyzed to gain new knowledge and even improve the health delivery system that generated the data. This is the essence of what the Institute of Medicine of the National Academy of Sciences calls a "learning healthcare system."[4] Any serious study of the field should include at least key parts of their publications on that topic.

---

[3] http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT.

[4] http://www.iom.edu/Reports/2007/The-Learning-Healthcare-System-Workshop-Summary.aspx.

In a table at the end of the book, I summarize the fascinating "secondary uses" of digital clinical data for analytics that I've discussed throughout the book. Of course, there are many other examples than those of which I am aware. The exploitation of the tsunami of digital health data is growing and already widespread, so there may well be other, even better, examples of which I am unaware.

While healthcare providers are the primary target for the book, it should serve equally well for other nontechnical readers. As a result, I do at times use a few sentences to explain clinical concepts that I know are already familiar to most healthcare providers.

To follow this rapidly changing field from a real-world perspective, I suggest you subscribe to *iHealthBeat*, a free, daily e-newsletter published by the California Health Foundation.[5] Even briefly reviewing it each day will give you a good understanding of what's happening in healthcare and health informatics. Here are some blogs that can also help you stay up-to-date.

http://geekdoctor.blogspot.com/
John D. Halamka, MD, MS, is Chief Information Officer of the Beth Israel Deaconess Medical Center, Chief Information Officer and Dean for Technology at Harvard Medical School, Chairman of the New England Health Electronic Data Interchange Network, Chief Executive Officer of MA-SHARE (the Regional Health Information Organization), Chair of the U.S. Healthcare Information Technology Standards Panel, a practicing emergency physician and a skilled farmer.

http://blogs.gartner.com/wes_rishel/
Wes Rishel is a vice president and distinguished analyst in Gartner's healthcare provider research practice. He covers electronic medical records, interoperability, health information exchanges and the underlying technologies of healthcare IT, including application integration and standards.

http://www.emrandhipaa.com/
John Lynn has written over 1,500 articles on various topics in healthcare IT. He was the EMR manager for the University of Nevada Las Vegas' Health and Counseling Center. In this capacity, he led the conversion from paper charts to a full electronic medical record. He has also worked on a variety of EMR consulting projects.

http://histalk2.com/
Anonymous, the author, says he works for a non-profit hospital that has vendor relationships with some of the companies he writes about. He says that his objectivity (and potentially his job security) could be compromised if vendors or anybody else worked that connection to muzzle him.

[5] http://www.ihealthbeat.org/.

http://hitsphere.com/
This is an attempt to create a single, comprehensive site with blogs, news, tools and other items of interest to the healthcare IT community. It was founded and is operated by Shahid N. Shah, an enterprise software analyst who specializes in healthcare IT. Over the past 15 years he has been Chief Technology Officer (CTO) for Cardinal Health's Clinical Technologies and Services unit (now CareFusion), CTO of two EMR companies, a Chief Systems Architect at American Red Cross, Architecture Consultant at the National Institutes of Health (NIH) and Senior Vice President of Healthcare Technology at COMSYS.

http://blogs.msdn.com/b/healthblog/
Bill Crounse, MD, Senior Director, Worldwide Health for the Microsoft Corporation, is a board certified family physician and helps guide the company's efforts in the healthcare space.

http://www.informationweek.com/author-bio.asp?author_id=1000
My own blog about various topics in health informatics with a focus on interoperability.

Note: Some of these blogs may accept advertising from the HIT industry.

The Current Situation

In 2009, Congress committed to supporting the adoption and meaningful use of electronic health records.... since that time, the adoption of EHRs has accelerated rapidly, but research has repeatedly demonstrated that adoption has not been consistent across all physicians.

—Dawn Heisey-Grove, MPH and Vaishali Patel, PhD, ONC Data Brief No 21, 2014[1]

Given current EHR adoption levels and provider survey data, I assume that many, if not most, healthcare provider readers of this book are using EHRs, may well be frustrated with them in one respect or another and are seeking advice or "wisdom" to help alleviate those frustrations or chose a better EHR in the future. I hope this book will help, but first, some perspective on where we've been and where we are today.

As we've said, healthcare informatics serves healthcare delivery—the translation of medical knowledge into the care of actual patients. As a result, its key components are:

- **Electronic medical records** created and used by licensed professionals
- **Personal health records** created and managed by patients
- **Health information exchange** used to share digital information

The term **electronic health record** means the merger of all available health data about patients into one integrated record. It is important to recognize that, given today's mobile, wearable and wireless technologies, a new source of this data can be the patients themselves. Major companies such as Apple, Samsung and Microsoft increasingly see health and fitness devices as an important adjunct to their ubiquitous smartphones and mobile devices and have introduced wearable devices to collect health and activity related data. They may also see aggregation of diverse patient-generated data as a business opportunity. This is a main purpose of software hubs like Apple's HealthKit, Samsung's Health Board and the Microsoft Health mobile applications (apps). Other companies are offering novel sensing

[1] http://www.healthit.gov/sites/default/files/oncdatabrief-physician-ehr-adoption-motivators-2014.pdf.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_2

technologies with the objective of making health and fitness data collection as ubiquitous, easy and multidimensional as possible. What to do with all this new data is still an open issue that we'll discuss later on.

No matter what the source, creating an integrated and comprehensive patient record requires **health information exchange**. It is a critical tool for achieving coordinated care among the many providers who may be seeing the same multi-chronic disease patients. It can empower patients to access and even contribute to their records, and it can make health data accessible for so-called "secondary uses" such as research on both population and public health.

Health information exchange requires core technologies to assure **privacy** and **security** of health data and to establish **trust** that the data is being shared as intended. All of this is governed by the Health Insurance Portability and Accountability Act of 1996 (HIPAA), a strict law with severe penalties, including jail time. Increasingly, objections are being raised to the difficulties this law has created in obtaining health data for important research purposes.[2]

Historically, health information exchange has required **standards** so that data is represented in a consistent way in the diverse systems involved and can more easily and accurately be meaningfully combined. As we'll see, these standards are a large and diverse part of health informatics and attempt to define the way data is represented, packaged and communicated for exchange. As they have evolved over time they have grown increasingly complex. Some would argue too complex. How best to overcome this complexity is a current, and I believe critically important, issue we will discuss in some detail later on.

First, of course, there must be digital data to share and for decades, obtaining it had been a virtually insurmountable problem. As I write, it is clear that the federal government's Health Information Technology for Economic and Clinical Health (HITECH) program has led to an explosion of EHR adoption. As recently as 2008–09, according to twin papers in the *New England Journal of Medicine*, only some 4 percent of physicians and 1.5 percent of non-federal hospitals (the United States Department of Veterans Affairs (VA) and U.S. Military Health System have long had electronic records) had a *clinically significant* EHR.[3, 4] While one recent report argues the current adoption levels would have been achieved without HITECH (albeit a couple of years later),[5] these levels are reportedly now around 70 percent for eligible providers and above 90 percent for hospitals, a huge increase that I believe was directly, and likely in large part, the result of HITECH. Moreover, as of the most recent November 2014 report by the Office of the National Coordinator for Health IT (ONC), the federal agency established to implement HITECH, some 337,861 Medicare-eligible and 166,670 Medicaid-eligible health professionals (refer to the glossary for an explanation of eligible professionals) and

[2] http://issues.org/25-4/detmer/.

[3] http://www.nejm.org/doi/full/10.1056/NEJMsa0802005.

[4] http://www.nejm.org/doi/full/10.1056/NEJMsa0900592.

[5] http://www.nber.org/papers/w20553.

some 4,789 hospitals had registered for the Meaningful Use program to encourage EHR adoption. This is over 95 percent of both groups which can then be reimbursed for the cost (perhaps more of it or less depending on how much they actually spend) of implementing an EHR if it qualifies and is used according to the rules of the program.[6]

Many are surprised to hear that, when this program was announced in 2004 and despite decades of failed efforts, some experts pushed back saying that we would end up deploying a panoply of noninteroperable systems and by doing so would, in essence, create a further barrier to information exchange. It is certainly correct that we've ended up deploying many different EHR systems. A visit to the ONC dashboard site reveals that there are around 700 different systems that providers have reported adopting.[7] Only around 30 of these are "primary" vendors with at least 1,500 installations, but even that is a large number considering that each vendor typically represents clinical data in their own, often unique and typically proprietary, way.

As a result, interoperability—the ability of diverse systems to share data—has risen to the level of public, national debate. Providers who have adopted EHRs are complaining that they can't share data.[8] Policy papers are being published to suggest how we should proceed.[9] Congress is complaining about the lack of interoperability and laws mandating it have been introduced.[10] Of greater interest to this discussion, interesting and arguably more facile and less expensive new technologies are now being promoted and taken seriously as a means to achieve interoperability.

Given all of this, I believe health informatics is now at a "tipping point" from which it will succeed or fail to achieve the lofty goals envisioned by thought leaders for decades. Getting past the interoperability barrier is arguably the key challenge now before us; but it is followed closely by how to usefully aggregate and analyze vast new quantities of digital health data. Your interest in informatics at this point in time is well placed and important. A more educated population of care providers and others in the healthcare domain is essential to gaining momentum and to moving us in the right direction.

We begin with some essential background on U.S. healthcare.

---

[6] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/May2014_SummaryReport.pdf.

[7] http://dashboard.healthit.gov/quickstats/pages/FIG-Vendors-of-EHRs-to-Participating-Professionals-2014.html.

[8] http://www.nytimes.com/2014/10/01/business/digital-medical-records-become-common-but-sharing-remains-challenging.html.

[9] http://healthit.gov/sites/default/files/ptp13-700hhs_white.pdf.

[10] https://www.congress.gov/bill/113th-congress/house-bill/4015.

Unique Complexity

U.S. health care is a highly complex enterprise with a "cottage-industry" structure (i.e., many small-scale, interdependent service providers that act independently creating "silos" of function and expertise). This siloed system is sorely mismatched to the nation's overriding health challenge, namely, providing coordinated, integrated, continuous care to more than 125 million Americans who suffer from chronic disease. ... While this framework has supported rapid advances in medical science and the development of increasingly precise diagnostic tools and therapeutic interventions, it has been indifferent to, if not discouraging to, innovation directed at harnessing advances in medical knowledge and precision diagnostics to improve the quality and efficiency of health care.

— Jerome H Grossman, MD, The Bridge, National Academy of Engineering, 2008[1]

Healthcare providers are often so immersed in the highly unusual U.S. healthcare system that they take it for granted and accept its extraordinary complexity as the norm. In fact, no other country on earth has a healthcare system quite like ours. Its complexity is the core reason for some very difficult and long-standing problems we'll now discuss.

For our purposes, it is useful to divide that complexity into its administrative and care-delivery components. Our system of employer-funded care adds substantial complexity to billing and associated administrative processes. Each employed patient who enters the waiting room of a healthcare provider may well have an employer-sponsored health insurance plan that is unique or at least highly customized to their employer and it can even be further tailored by individual choices made by that employee. The plan determines what will be covered by insurance, how much the insurance will pay for each specific test or procedure and how much of the cost is the patient's responsibility. Unusual, experimental or expensive tests and procedures usually require pre-approval. Increasingly, under the Affordable Care Act, even unemployed or poor patients may have commercial insurance plans that are offered by increasingly competitive insurance exchanges operated by their state or the federal government.

[1] http://www.nae.edu/File.aspx?id=7417.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_3

Older or poor patients are often insured by Medicare or Medicaid, each of which have their own rules and procedures and which vary from state-to-state for Medicaid. Medicare patients may also have supplemental insurance plans.

The sum of all this is a very complex process of determining insurance coverage and properly billing the insurance companies and patients for their respective parts of the bill. Providers and patients struggle to manage this complexity and it accounts for a substantial part of healthcare cost. Providers and patients are helped in part by automation of medical billing and related processes which, while widespread for many years, are often still not able to deal accurately or comprehensively with the entire process. The complexity of our payment system and the shifts that have taken place in it can be appreciated by visiting an interactive graphic on the California Health Foundation web site.[2] Another more widely cited interactive graphic by the Institute of Medicine, the part of the National Academy of Sciences that focuses on healthcare, estimates the price of excessive administrative costs in U.S. healthcare at $190 billion per year.[3]

In this discussion we'll focus on the other component—the complexity of our fragmented care-delivery system. It is useful and important to what follows that we take the time to step back and reflect not only on the complexity of care delivery but on some of the root causes of that complexity. This will provide the background for appreciating the potential role health informatics can play in improving care quality and efficiency.

In what follows, we'll also focus mainly on the design of our ambulatory healthcare delivery system. Specifically, it is not well suited to the management challenges presented by our most common clinical problems—chronic diseases. Before exploring the system design issues, we need to review chronic diseases. These are, of course, quite common, particularly in the elderly. Virtually all Medicare patients have a chronic disease and two-thirds have more than one. This should also not be surprising since chronic diseases can cause other chronic diseases, particularly if they are not well controlled. Hypertension can lead to coronary artery disease. Diabetes can lead to chronic kidney disease. As a result, unless they are managed well on a continuous basis—*something most parts of our healthcare system are not designed to do*—many chronic disease patients will, over time, develop multiple problems and their resulting expensive complications. This impact on cost is illustrated by the fact that the 37 percent of Medicare patients who have four or more chronic diseases account for 74 percent of all Medicare costs.[4]

Patients often cause or contribute substantially to their chronic diseases. In part because of lifestyle and behavioral issues that are largely outside the control of the

[2] http://www.chcf.org/publications/2014/07/data-viz-hcc-national.

[3] http://resources.iom.edu/widgets/vsrt/healthcare-waste.html?keepThis=true&TB_iframe=false&height=729&width=871;.

[4] http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/Downloads/2012Chartbook.pdf.

healthcare system, the U.S. is experiencing a chronic disease epidemic.[5] Chief among the lifestyle-induced risk factors for chronic disease is obesity. Only some 25 years ago, according to the Centers for Disease Control and Prevention (CDC) no state in the U.S. had an obesity rate *as high as 15 percent*. Today, no state has a rate that *low* and some states are well above 30 percent. Projections are that, if nothing changes, some state rates may exceed 40 percent in the coming years. The prospect of changing patient behavior through technology is the primary motivation behind the flood of wearable, mobile and wireless technologies and devices we discussed earlier in *The Current Situation* section. We'll discuss it in more detail in the *Patient Engagement and Empowerment* section.

We turn now to the structural issues in U.S. healthcare that make chronic disease management so problematic. This is not complete coverage of the issue. For that, you might want to read at least the relevant parts of the 2012 Institute of Medicine (IOM) publication *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*[6] and *The Growing Burden of Chronic Disease in America,*[7] a seminal 2004 paper by Gerard Anderson and Jane Horvath from the Johns Hopkins School of Public Health. This is a complex issue with many dimensions. For our purposes we'll focus on two that most directly suggest the potential role of health informatics:

We're Overly Specialist-Oriented: Most of the routine care of chronic disease is best done by a primary care physician (PCP) heading a team that can *proactively and continuously* manage the whole patient. According to the Commonwealth Fund (a private foundation that "aims to promote a high performing health care system"):

- People who have access to a regular PCP are more likely than those who do not to receive recommended preventive services and timely care for medical conditions before they become more serious and more costly to treat.
- Having a regular doctor is also associated with fewer preventable emergency department visits and fewer hospital admissions, as well as with greater trust in and adherence to physicians' treatment recommendations.
- Among low-income patients, access to primary care is associated with better preventive care, better management of chronic conditions and reduced mortality.[8]

[5]Anderson G and Horvath J 2004. The growing burden of chronic disease in America. Public Health Rep. May-Jun; 119(3): 263–270.

[6]http://www.iom.edu/Reports/2012/Best-Care-at-Lower-Cost-The-Path-to-Continuously-Learning-Health-Care-in-America.aspx.

[7]Anderson G and Horvath J 2004. The growing burden of chronic disease in America. Public Health Rep. May-Jun; 119(3): 263–270.

[8]Abrams, M et al. 2011. Realizing Health Reform's Potential: How the Affordable Care Act Will Strengthen Primary Care and Benefit Patients, Providers, and Payers. Commonwealth Fund pub. 1466.

Yet, according to 2009–10 data from the CDC, "the supply of generalist physicians per capita (472 per 1 million population) was 26 percent lower than the supply of specialty physicians (636 per 1 million population)" and "the annual visit load for generalist physicians (3,521) was 30 percent higher than the annual visit load for specialty physicians (2,704)."[9] Data from the OECD suggests that, at 12.3 percent, the percentage of U.S. physicians who are in primary care (which OECD terms "generalist physicians") is among the lowest in the world.[10] The percentage of PCPs varies (up to around 20 percent) in other reports depending on which providers are included in that category. At present, there are concerns about the ability of our primary care system to handle the increased patient load created by the estimated 10 million Americans who now have health insurance coverage provided by the Affordable Care Act.[11] However, a recent report suggests it's too early to know how severe this problem is.[12] In any case, there is growing interest in increasing the involvement of patients in their own care and expanding the role of other providers such as nurse practitioners, physician's assistants and even medical school graduates who have yet to do a residency to address this potentially serious shortfall.[13]

There does seem to be consensus that, under a traditional physician-centric practice model, we're ill equipped to provide enough primary care to the ever growing population of chronic disease patients. As a result of that model "while the average Medicare beneficiary sees between six and seven different physicians, *beneficiaries with five or more chronic conditions see almost 14 different physicians in a year and average 37 physician visits annually.*"[14] This is an enormous problem in a healthcare system that is not well equipped or designed to share data. I recently interviewed Dr. Gerard Anderson, the primary author of the paper quoted earlier (and which I strongly suggest you read)—he is not very optimistic things will change anytime soon.

This leads to what was arguably the *key* rationale for the federal programs to spur EHR adoption: How can these many specialists effectively coordinate care if they are using paper records? The answer is that, for the most part, they can't. Again turning to the IOM, but this time to their earlier 1999 book *To Err Is Human: Building a Safer Healthcare System*, "One oft-cited problem arises from the decentralized and fragmented nature of the health care delivery system—or "nonsystem"—to some observers. When patients see multiple providers in different settings, none of whom has access to complete information, it becomes easier for things to go wrong."

---

[9] http://www.cdc.gov/nchs/data/databriefs/db105.htm.

[10] http://stats.oecd.org/Index.aspx.

[11] http://www.nytimes.com/interactive/2014/10/29/upshot/obamacare-who-was-helped-most.html?ref=health.

[12] http://www.kaiserhealthnews.org/stories/2014/may/14/obamacare-primary-care-demand-shortage.aspx.

[13] http://www.latimes.com/opinion/op-ed/la-oe-0808-caplan-doctor-shortage-20140808-story.html.

[14] Anderson G and Horvath J 2004. The Growing Burden of Chronic Disease in America. Public Health Reports, May–June, 119:263-270.

In fact, "20 percent of the time specialists see patients without having their medical records."[15] A September 2014 ONC report states that: "When seeking care for a medical problem within the last year, about one in three individuals reported experiencing one or more gaps in information exchange. Common gaps included: recounting medical history because the healthcare provider had not received records from another provider or having to bring test results to an appointment."[16]

The economic results are serious and could get worse when, according to another powerful paper I strongly suggest you read: "The single greatest cause of rising healthcare spending in the U.S. is the growing prevalence of chronic disease."[17]

We Reward the Wrong Things: Managed well, the complications of chronic disease that often require expensive high-tech care may be avoided or at least delayed for years. Managing chronic disease itself is a relatively low technology, long-term process that can be economically characterized as investing now for a future return of lower healthcare costs. However, under our current system we don't reward that. Rather, we reward procedures and tests (particularly if they involve state-of-the-art, expensive technology) more than we reward routine primary care. In essence, we reward the care of acute, life-threatening problems while neglecting the care of the far more common and, in aggregate over time, far more expensive routine chronic diseases. Dr. Brent James, Director of Intermountain Healthcare's Institute for Health Care Delivery Research, calls the treatment of these serious, acute problems "rescue care" (a term that may have first been used in a 1986 paper by Dr. Albert Jonsen[18])—saving those who would otherwise die soon—and his data show that the U.S. has the world's best survival rates for these problems. However, he goes on to say that, otherwise, many countries out-perform the U.S. and says that this is "primarily attributable to healthier behaviors, better public health and a heavy emphasis on easily accessible primary care."[19]

The Institute of Medicine agrees that we have serious problems and says: "America's health care system has become far too complex and costly to continue business as usual. Pervasive inefficiencies, an inability to manage a rapidly deepening clinical knowledge base and a reward system poorly focused on key patient needs, all hinder improvements in the safety and quality of care and threaten the nation's economic stability and global competitiveness. Achieving higher quality care at lower cost will require fundamental commitments to the incentives, culture

[15]2008 Commonwealth Fund International Health Policy Survey of Sicker Adults.

[16]http://www.healthit.gov/sites/default/files/consumeraccessdatabrief_9_10_14.pdf.

[17]Kumar, S and Nigmatullin, A, 2010. Information Knowledge Systems Management 9:127-152.

[18]http://www.medicine.mcgill.ca/epidemiology/courses/EPIB654/Summer2010/Pricing/Jonson_Law%20Medicine%20-%20Health%20Care_1986.pdf.

[19]http://www.mdanderson.org/education-and-research/resources-for-professionals/clinical-tools-and-resources/clinical-safety-and-effectiveness-educational-program/selected-lectures/csetraining-managing-clinical-processes.pdf.

and leadership that foster continuous 'learning,' as the lessons from research and each care experience are systematically captured, assessed and translated into reliable care."[20] This concept of a "learning healthcare system" advocated by the IOM leads directly to the role that analytics of aggregated health data can play in improving care and the delivery system. The role of analytics will be the topic of the last two sections of the book, but I will provide other examples in context in some earlier sections.

Some key PCP groups agree. Robert B. Doherty, Senior Vice President for Governmental Affairs and Public Policy at the American College of Physicians, has publicly said that "We need to move away from the piecemeal approach: how many visits you can generate, how many tests you can order."[21] In testimony to Congress, he urged more funding for a new model of care called the patient-centered medical home (PCMH).[22] A complete treatment of the PCMH would be lengthy but, for our purposes, its key attributes will suffice. According to the American Academy of Family Physicians (AAFP) they are:

- Access to a personal physician who leads the care team within a medical practice
- A whole-person orientation to providing patient care
- Integrated and coordinated care
- Focus on quality and safety

The AAFP feels that informatics is a critical part of this and says that: "A patient-centered medical home automates business and clinical processes, depends on clinical decision support (CDS) tools and connects patients with the health care team."[23] In the same congressional testimony we quoted earlier, Mr. Doherty went on to explain the key role of financial incentives when he said that: "Medicare discourages primary care physicians from organizing their practices to achieve optimal results for patients by failing to reimburse for care coordination, health information technology, e-mail consultations and other proven methods for improving quality and access." This was the theme of a Tom Friedman article in the May 25, 2013, edition of *The New York Times* in which he says that to succeed under a pay-for-value system "doctors and hospitals need instant access to data about patients—diagnoses, medications, test results, procedures and potential gaps in care that need to be addressed."[24]

The results of a healthcare system with these problems is far too often poor quality at high cost. Under current trends in reimbursement, the focus of hospital-

[20] http://www.iom.edu/Reports/2012/Best-Care-at-Lower-Cost-The-Path-to-Continuously-Learning-Health-Care-in-America.aspx.

[21] http://www.managedcaremag.com/archives/0712/0712.docpay.html.

[22] http://www.acponline.org/advocacy/advocacy_in_action/state_of_the_nations_healthcare/assets/statehc06_4.pdf.

[23] http://www.aafp.org/practice-management/pcmh/overview.html.

[24] http://www.nytimes.com/2013/05/26/opinion/sunday/friedman-obamacares-other-surprise.html.

based care is increasingly "rescue care," high-technology treatment of acutely ill patients who require the specialized equipment and skills found only in hospitals. Here, as Dr. James said, the U.S. leads the world. United States hospitals are generally accredited by the Joint Commission. In 2013, it reported that:

> "The number of hospitals recognized by the Top Performer on Key Quality Measures program has increased by 77 percent from last year, with 1,099 hospitals earning this achievement for 2013. These hospitals represent 33 percent of all Joint Commission-accredited hospitals reporting accountability measure performance data for 2012. Each of the 1,099 Top Performer hospitals met three performance criteria based on 2012 accountability measure data.
>
> A recognized hospital must: 1) achieve cumulative performance of 95 percent or above across all reported accountability measures; 2) achieve performance of 95 percent or above on each and every reported accountability measure where there are at least 30 denominator cases; and 3) have at least one core measure set that has a composite rate of 95 percent or above and (within that measure set) all applicable individual accountability measures have a performance rate of 95 percent or above."[25]

Despite this optimistic view, a recent analysis of four 2006–12 studies estimates that there are "210,000 preventable adverse events per year that contribute to the death of hospitalized patients."[26] This echoes the IOM's earlier estimates and indicates that, despite any progress, we still have a long way to go both in terms of measuring and addressing concerns about patient safety in U.S. hospitals. Information technology is sometimes sold or even seen by hospital management as a panacea but, as with any tool, it may help but will not, alone, provide a real solution absent meaningful process and workflow changes.

Improper medical decision making is clearly not limited to hospitals or to acute, life threatening problems. The AAFP reported in 2011 that "as primary care physicians, we spend the majority of our time caring for patients with chronic diseases, but data suggest we achieve the standard of care for chronic diseases and preventive care only 50 percent to 60 percent of the time."[27] The Assessing Care of Vulnerable Elders (ACOVE) study published by RAND and Pfizer in 2004 similarly concluded that "vulnerable elders receive about half of the recommended care, and the quality of care varies widely from one condition and type of care to another."[28] This suggests the key role that CDS, a technology we'll discuss later, might play in improving care quality and efficiency.

Even today, well into a new era of electronic patient record system adoption, publicly reported quality data from the Indiana Health Information Exchange shows that many physician groups have achieved acceptable HbA1c levels for only around half of their diabetic patients and a smaller number of practices has performed the

---

[25] http://www.jointcommission.org/assets/1/6/TJC_Annual_Report_2013.pdf.

[26] http://journals.lww.com/journalpatientsafety/Fulltext/2013/09000/A_New,_Evidence_based_Estimate_of_Patient_Harms.2.aspx.

[27] http://www.aafp.org/fpm/2011/0500/p27.html.

[28] http://www.rand.org/content/dam/rand/pubs/research_briefs/2005/RB9051.pdf.

test for only around half of their patients. Keep in mind that these are practices using informatics tools specifically designed to improve performance against these metrics and that are voluntarily reporting quality metrics that they know will be made public.[29] It would be reasonable to be concerned that elsewhere performance might be even worse. In the next section we'll see some data from the National Committee for Quality Assurance (NCQA) that suggests this may well be the case.

Summary: Chronic disease accounts for most U.S. healthcare spending and the rates continue to grow. The U.S. healthcare system excels at the treatment of acute, life threatening conditions, but it produces mediocre results for managing chronic diseases. The solutions involve different models of care such as the PCMH that utilize a team to leverage precious PCP resources and to produce a more continuous, coordinated approach to care driven by quality metrics. Actually doing this requires the use of sophisticated health informatics tools. Making it work in the real world requires new financial incentives that reward these new models when they produce measurable quality results.

In the next section we'll see how closely aligned the design of the new federal programs and incentives are to this view of healthcare.

[29] http://www.ihie.org/public-reporting.

EHR Adoption and Meaningful Use

Two basic arguments justified intervention by the federal government in 2009 to promote the adoption and meaningful use of health information technology. The first was a conviction that information technology could improve health and health care for the American people. The second was that major problems inhibit the spread of health information technology in ways that create the need for government remedies.

— David Blumenthal, MD, MPP, NEJM, 2011[1]

As we've seen, there is a disconnect between what is required to successfully manage the chronic diseases that drive most healthcare costs and the structure and incentives of the U.S. healthcare system. We've also suggested that health IT could be a key tool for restructuring healthcare delivery to help address these issues. We'll review that again here, from an engineering perspective, before turning to the policies the federal government has promulgated in recent years to encourage adoption of health information technology and new models of patient-centered care.

From an engineering perspective, chronic disease care presents a data logistics problem. Figure 1 presents a "network depiction" of the complex, highly specialist-driven care of chronic disease patients from the perspective of the typical PCP's practice. You know from our earlier discussion that the average patient with multiple chronic diseases—the 37 percent of Medicare patients who have four or more chronic diseases and account for 74 percent of the costs—is seen each year by the 14, mostly specialist providers, depicted at the lower left. In aggregate, all the multiple chronic disease patients in a typical PCP's practice are seen by 86 providers, again mostly specialists, depicted behind the group of 14 providers. Of course, *each* of those specialists focuses only on the particular organ or body system they have special knowledge of and trained to treat. In total, across their entire practice, the average PCP has referral relationships with the 229 providers depicted overall in this illustration.

This is reminiscent of the network of specialized suppliers to a manufacturing company. For an automobile manufacturer one supplier makes seats, but not radiators, while another makes dashboards, but not tires. Somehow, it all needs to

[1] http://www.nejm.org/doi/full/10.1056/NEJMsr1110507.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_4

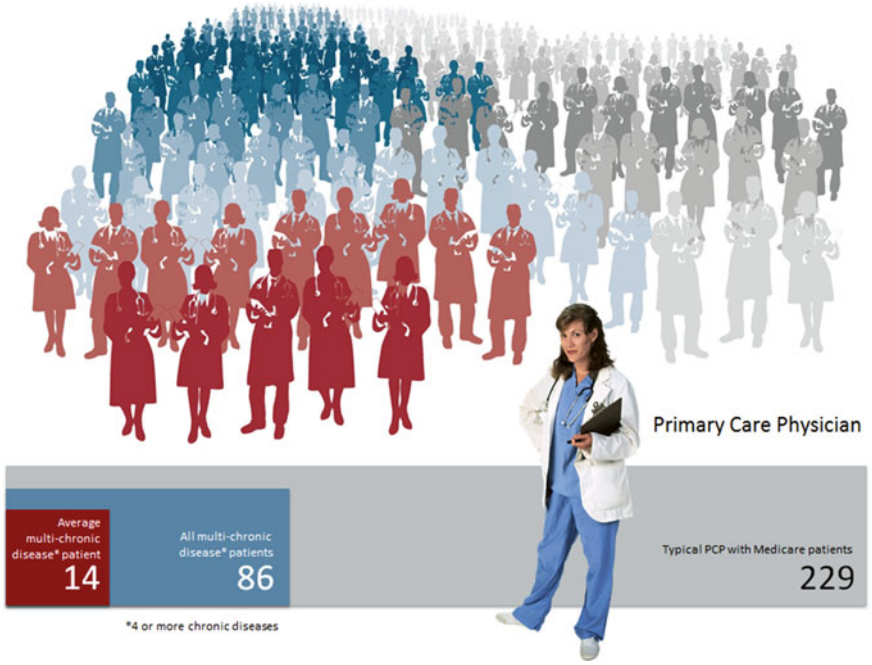


Figure 1: This presents a network depiction of the typical PCP's practice. It illustrates that 14 providers are involved in the care of the average multichronic disease patient (e.g. a patient with four or more conditions); on average 86 providers are involved in the care of all the multichronic disease patients in a typical PCP's practice; and that, in total, the typical PCP has referral relationships with 229 other providers. Trying to manage such a complex care network on paper has been a primary reason for the lack of well-coordinated care of chronic disease. (Source Author)

come together seamlessly to produce a great car. The manufacturing industries long ago recognized the need for data sharing and automation to help coordinate their supplier network, starting with the automobile industry in the 1980s. Until very recently—and this only started to change as a result of the federal programs we'll discuss in this section—the healthcare industry mostly tried to operate its complex care network using paper records and fax machines—very limited and constricting communication systems reminiscent of the low-technology industrial past.

The resulting problems are not theoretical. Patient surveys show a positive correlation between the number of providers caring for a chronic disease patient and the likelihood of that patient being seen by a provider who has incomplete or missing records. The multiple chronic disease patients we've been discussing make many visits to multiple physicians; they also fill over 50 prescriptions per year because, for most chronic diseases, medications are the main therapy.[2] Misuse of

[2]Anderson G and Horvath J 2004. The Growing Burden of Chronic Disease in America. Public Health Reports, May–June, 119:263-270.

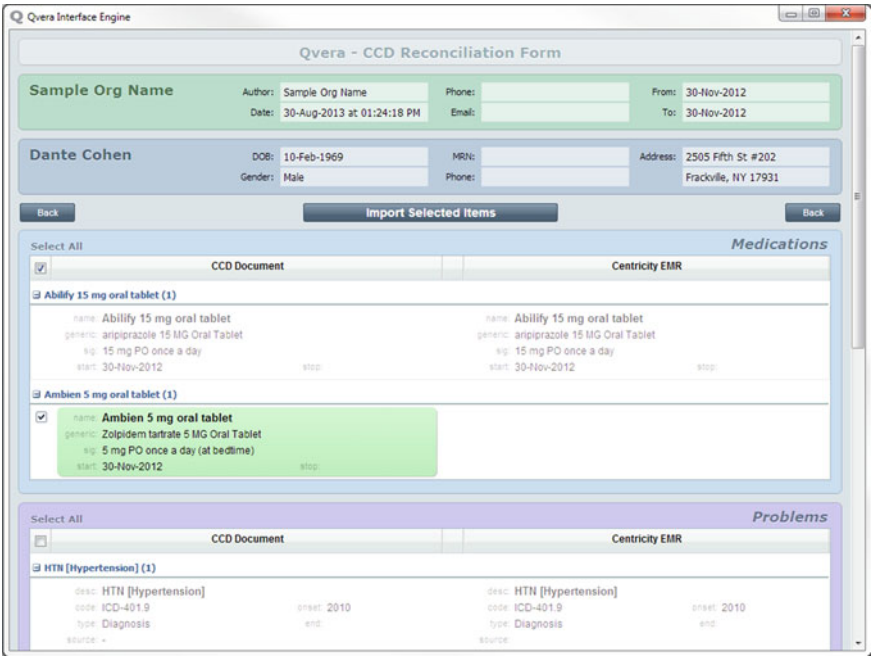


Figure 2: Automatic reconciliation of the medications in a physician's EMR with those from an electronic patient summary from another practice highlights Ambien as a medication from that other practice, but not on this physician's version of the patient's medication list. (Courtesy Qvera, All Rights Reserved)

medications, including in particular, duplicates because of poor care coordination, is a major problem that accounts for nearly a third of all hospitalizations. Surveys in 2008 and 2011 by the Commonwealth Fund show progress, but many patients from all countries (28 percent in the U.S. in 2011) still report that no provider has reviewed all of their medications with them in the past year.

One reason for the absence of review is that physicians may not know the medications that all the other physicians have prescribed for a multichronic disease patient. Medication reconciliation is the solution and is a clear opportunity for health IT, once the underlying records are digital. Figure 2 is a screenshot from an actual commercial HIT system showing the medications in this physician's EMR on the right and those from other EMRs on the left. Both records include the prescription for Abilify (aripiprazole) 15 mg but the physician's EMR does not have a record of the Ambien (zolpidem) prescribed elsewhere. It is, of course, important for all this patient's physicians to have a complete medication record and, in this case, the two drugs illustrated do interact causing increased dizziness, drowsiness and difficulty concentrating and, less often, impairment in thinking and judgment.

As we'll discuss later on, e-prescribing is a key requirement of the new federal adoption programs. Because electronic prescriptions are clear and legible and an EMR can often spot potential problems as prescriptions are written, e-prescribing

has been shown to dramatically reduce medication errors within a year of being adopted. A widely quoted 2010 study showed that medication errors decreased nearly sevenfold, from 42.5 per 100 prescriptions (95 percent confidence interval (CI), 36.7–49.3) to 6.6 per 100 prescriptions (95 percent CI, 5.1–8.3) within a year of adopting e-prescribing ($p < 0.001$).[3] Hopefully, someday soon each provider will have a full medication record, including whether patients are filling their prescriptions and then refilling them at the proper intervals. In a conversation with Surescripts, the nationwide e-prescribing network, I was told that patient privacy concerns, a topic we'll discuss in detail later on, are an impediment to routinely and automatically providing this information.

New Incentives: Given their potential, as illustrated by the examples of care coordination and medication reconciliation, why have electronic records not been adopted much sooner? One reason may be that healthcare providers have a financial incentive to do all available tests and procedures once someone is sick, and have had little or no incentive to keep their patients well. Until recently there was no incentive —except for special situations like health maintenance organizations (HMOs)—for providers to invest their own money in HIT systems that might actually reduce income by helping avoid unnecessary or duplicative tests and procedures.

The needed incentives are often referred to as "pay-for-performance" or PFP. In the typical PFP, system providers are rewarded for doing the tests and procedures that scientific evidence suggests are beneficial for either managing or preventing chronic disease. For diabetes, this might be an annual HbA1c determination and appropriate measures to achieve good control. The need to do better is illustrated by NCQA data from 2012 (reported in 2013) which shows that between 8.6 and 17 percent of diabetic patients (depending on their payer) do not get regular annual HbA1c testing. Moreover, when tested, significant numbers of these patients are in poor control with HbA1c values of 9 percent or higher: 28.5 percent of commercial HMO patients versus 35.2 percent of commercial preferred provider organization (PPO) patients; 29.3 percent for Medicare PPO patients versus 27.1 for Medicare HMO patients; and 44.7 percent of Medicaid patients.[4] Other quality measures that are often part of a PFP system might be *screening* patients for obesity and smoking and *counseling* them to lose weight or quit smoking.

A more advanced PFP system rewards providers who are able to produce higher care quality at a lower cost. There is evidence that these programs can work. The physician group practice (PGP) demonstration, a well-known Medicare pilot in 10 practices, is of particular importance in that regard. All of the groups achieved improvements in 25 of 27 quality metrics for important chronic diseases. Four practices earned a bonus for outstanding performance. Marshfield Clinic earned half of the total bonus and, in explaining how they did it, their CEO cited "a well-developed electronic health record" and went on to describe how it reduced

[3]Kaushal R *et al* 2010. Electronic Prescribing Improves Medication Safety in Community-Based Office Practices. J Gen Intern Med. Jun 25(6): 530–536.

[4]http://www.ncqa.org/Portals/0/Newsroom/SOHC/2013/SOHC-web_version_report.pdf.

unnecessary duplication of services by making information available to all providers caring for each patient.[5]

Accountable care organizations (ACOs) are a key feature of the Affordable Care Act. The ACO program design is referred to by Medicare as a shared savings plan and is, in part, derived from the PGP demonstration project.[6] There are currently around 600 public and private ACOs being formed or in operation.[7] Starting in 2011, 32 select Pioneer ACOs began piloting an even more advanced payment model.[8] The first two years were also a shared savings payment program but with generally higher levels of shared savings and risk than in the ACO program. In year three of the program, Pioneer ACOs that have shown a specified level of savings over the first two years became eligible to move a substantial portion of their payments to a population-based model. However, this is a challenging model and, as of this writing, only 19 providers (listed in the reference) remain in the Pioneer ACO program due to inconsistent results and financial losses in a program that rewards success but also penalizes failure to save money while delivering acceptable quality.[9] Like the Marshfield Clinic, Pioneer ACOs must use advanced health IT to manage entire populations of chronic disease patients in a coordinated manner, to exchange data, to report on results, to engage patients and to coordinate care. So, new incentives are necessary but, alone, insufficient. To operate successfully under a PFP or similar model physicians need advanced clinical health informatics. We turn now to the federal programs designed to foster their adoption.

Health IT Adoption Programs: In 2004 in his State of the Union Address, President George W. Bush made universal adoption of electronic records a 10-year national goal. This was not a minor part of the address. The text posted on whitehouse.gov has this headline: "By computerizing health records, we can avoid dangerous medical mistakes, reduce costs and improve care."[10] He tasked a new Office of the National Coordinator for Health IT with achieving the goal. In 2009, the Obama administration's HITECH Act provided funding of $20–$30 billion (the exact amount depends on adoption levels) to reimburse hospitals and "eligible providers"[11] (a term defined differently by the Medicare and Medicaid programs) for adopting an electronic health record system. The adoption program has three codependent components: EHR Certification, Meaningful Use and incentive payments. We'll spend the rest of this section discussing these at a high level. I provide references for those wanting more detail.

---

[5] https://www.marshfieldclinic.org/about-us/quality/medicare-savings.

[6] http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/Downloads/ACO_Quality_Factsheet_ICN907407.pdf.

[7] http://healthaffairs.org/blog/2014/01/29/accountable-care-growth-in-2014-a-look-ahead/.

[8] http://innovation.cms.gov/initiatives/Pioneer-ACO-Model/.

[9] https://data.cms.gov/dataset/Pioneer-ACO/izub-xmpg.

[10] http://georgewbush-whitehouse.archives.gov/infocus/technology/economic_policy200404/chap3.html.

[11] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/eligibility.html#BOOKMARK1.

- EHR Certification defines the minimal acceptable requirements for an EHR that, if used according to the requirements of Meaningful Use, would qualify an eligible provider for an incentive payment.
- Meaningful Use defines how eligible providers must use their certified EHR to be eligible to receive incentive payments.
- Incentive payments from either Medicare or Medicaid compensate hospitals and eligible providers for implementing a certified EHR and achieving Meaningful Use.

EHR Certification: Specified requirements for commercial EHR systems are an important development. For decades the "grand challenge" for health IT here in the U.S. has been interoperability—the ability of the hundreds of commercial EMR products and tools to exchange and meaningfully share data. Even if we had universal adoption, without at least some degree of interoperability, we would still not achieve coordinated care among the many providers who collectively care for the same patients. The EHR Certification process supports a basic interoperability capability. Many critics say it sets a bar that is too low. Others view it as an important first step. Later on, we'll look at a proposal to achieve a far more robust interoperability framework. For now, EHR Certification requires the recording of key demographic and clinical data, tools to measure and improve care quality, reporting the recorded data and protecting data confidentiality, integrity and availability. We'll now look at each of these in some detail.

The key clinical data that must be collected includes:

- Vital signs
- Smoking status
- A current problem list
- An active medication list
- An active medication allergy list
- Laboratory test results

The system must compare proposed new prescriptions to any relevant drug formulary (an institutional or payer created list of recommended or approved medications) and it must also be capable of generating lists of patients, such as those that are overdue for preventive care.

The reasons for some of these should be clear from our prior discussions. Keep these requirements in mind as we later discuss Meaningful Use, the EHR usage requirements placed on eligible providers.

Just recording data isn't sufficient. Certified EHRs must provide tools to use that data to improve care quality through functions like these:

- Electronic prescribing
- Drug-drug and drug-allergy interaction checks
- Medication reconciliation
- Computerized provider order entry
- Patient reminders
- Patient-specific education resources

- Automated measure calculation
- Ability to calculate and submit clinical quality measures (CQMs)

Earlier, we discussed the key role medications play in managing chronic disease and the many problems they currently create, so note that the first four requirements relate directly to improved medication management. Also, note that the EHR must be able to calculate and report quality measures. Remember that as we discuss the key role that quality measures play in determining Meaningful Use.

Of course, none of this can be done legally or will be accepted unless EHRs can insure **privacy** so that patient data is accessible only to people to whom the patient grants access; they must also provide **security** from unauthorized access. Other systems for information exchange must provide a means to establish **trust**—the ability to know that the persons or entities with which information is being exchanged are who they claim to be. The technologies to help meet these challenges will be a later topic.

Vendors become certified through a formal testing process developed by the National Institute of Standards and Technology and administered by one of several companies. You might wonder how this testing works. One of the key clinical data collection requirements we listed was that a certified EHR must "maintain a current problem list" but what does that mean? Medical problems are usually coded using the International Classification of Disease (ICD). A vendor can demonstrate that its EHR meets this criterion by showing that it can store test ICD codes and supplement them with problem status and the date diagnosed. The testing would include asking the vendor to change a problem's status, for example from active to resolved, and demonstrating that this change is posted and displayed at appropriate places throughout their EHR.

Quality reporting is a particularly interesting and challenging area generally done through **process** and **outcome** measures. We'll get into more detail later but process measures tally the things that are believed to produce a favorable outcome. Outcomes are the actual results of care and, while usually preferable, may not be readily available. The measurement of each diabetic patient's HbA1c level is useful to explain the difference since it can serve both as a process and an outcome measure for diabetes care. Hemoglobin is the oxygen-carrying molecule in our erythrocytes. The level of its A1c variant is proportionate to the amount of serum glucose—the molecule that is not properly regulated in diabetes—entering the red blood cells. The more glucose, the higher the HbA1c level, but this increase occurs over time so HbA1c is proportional to the *average* blood glucose level roughly over the prior couple of months. This is very useful since the goal of diabetes therapy is to keep that same *average* glucose level within normal ranges over time. The blood glucose level is volatile based on diet, exercise and other factors. However, while it may go up and down, the HbA1c level will remain static so long as the average glucose level remains unchanged. Good control, the desired outcome, is defined as an HbA1c level under 9 percent but some organizations, such as the Mayo Clinic, use 7 percent as the benchmark. As a result of the usefulness of this test to measure control of diabetes over time, for example, between office visits, HbA1c testing at appropriate intervals is the process metric and is generally accepted as a useful and necessary part of diabetes care.

Just as this book was going to press ONC released the 2015 Edition Health Information Technology Certification Criteria, a complex 431 page document that specifies an updated EHR certification program designed to support the final Stage 3 of Meaningful Use. I did not have time yet to fully digest it and it may well change but one particularly interesting section (pages 206–7) divides health information exchange into "data category" (e.g., a request for just a patient's medications) and "all" requests (e.g., a request for a complete patient summary). It further specifies that EHRs will need to provide an API-based response in XML or JSO N formats (similar to what the new HL7 Fast Healthcare Interoperability Resources (FHIR) standard proposes) to the former request but an HL7 Consolidated Clinical Document Architecture (CCDA) XML formatted clinical document to the later. We will be discussing both FHIR and CCDA in detail later on, so keep this in mind when we do. Also, later on, we will be discussing The Argonaut Project that seeks to potentially replace CCDA documents with "bundles" of FHIR resources. Some observers feel that this proposal is an attempt to position CCDA as a "bridge" until FHIR is sufficiently developed to respond to "all" requests.[12]

Meaningful Use: This is a very important and highly visible, sometimes even controversial, program divided into three stages that phase in over a period of years. Stage 1 focuses on **data capture and sharing**, which depends on EHR Certification; Stage 2 adds **advanced clinical processes** and is similar to, but more ambitious and sophisticated than Stage 1; Stage 3 is not yet clearly defined but aims for **improved outcomes**, which depends on more advanced EHR functionality such as CDS to guide providers and tools to assist patients. As a result, Stage 3 is quite ambitious and has recently been pushed out at least one year to 2017. A new and potentially important report recommends that Stage 3 introduce a national interoperability framework. We'll discuss that in more detail later in this section but a concern expressed by some is that doing this might push the date even further out. The majority of providers and hospitals participating in Meaningful Use have achieved Stage 1 and are now focused on Stage 2 (only a small percentage having achieved at present).

Eligible providers submit quality metrics in three categories to demonstrate that they have achieved Stage 1: **Core objectives**, **menu set objectives** and **clinical quality measures**. The specific requirements were updated in 2014 and there are now 13 mandatory core objectives. One of these is particularly interesting and we'll discuss it in more detail. First, providers must also submit five of nine **menu set** objectives including at least one public health objective. These are summarized in portable document format (a PDF) if the details interest you.[13] They must also submit nine out of 64 of the CQMs covering at least three of these six National Quality Strategy domains:

1) Patient and family engagement
2) Patient safety

[12] http://federalregister.gov/a/2015-06612

[13] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/EP_MU_TableOfContents.pdf.

3) Care coordination
4) Population/public health
5) Efficient use of healthcare resources
6) Clinical process/effectiveness

Providing patients with a clinical summary of their visits has always been part of Meaningful Use, but the specifics have changed over time.

Stage 1, Pre-2014: Providers without an electronic means of providing a clinical summary (such as a patient portal, a web site designed for patient interaction with their provider) could, upon request, provide the patient with a printed version within three business days of the visit.

Stage 1, Post-2014: 50 percent of unique patients who are seen during a reporting period must have online access and the *ability* to view, download and transmit (VDT) their health information to a third party within four business days of the information being available to the eligible professional.

Stage 2: A summary must be provided within a business day after the visit. If an electronic means to deliver it is available, it can be provided using that, rather than in printed form. 50 percent of unique patients must have online access to tools providing VDT capability *and* 5 percent of unique patients seen in each reporting period *must actually* view, download or transmit to a third party, their health information.

Providers can exclude patients not seen using an EHR from calculation of these percentages. The clinical summary can be provided through electronic technologies we'll be discussing later on. In addition to a patient portal, these could be done via a personal health record (PHR) (a web site for patients to use to record and access their health data), using secure e-mail (such as healthcare-specific Direct secure e-mail-based technology for information sharing), via digital media such as compact disc (CD) or Universal Serial Bus (USB) stick or as a printed document. This interesting requirement is also summarized in a PDF you may wish to read.[14]

The VDT requirement is directly linked to the need for coordinated care we discussed in the previous section. Again for emphasis, in Stage 2 at least 5 percent of a provider's unique patients seen in a reporting period *must actually* view, download or transmit their clinical summary to a third party.[15] A particularly important use case for this is Transitions of Care where, for example, a patient goes from a hospital to a nursing home or back to their home. Errors often occur at these transition points because information is not passed on completely and accurately. Later, we'll look in detail at the electronic clinical summary specified at transitions of care. This is the same Continuity of Care Document (CCD) electronic clinical summary that was used to perform the example of automated medication reconciliation we saw earlier.

[14] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/13_Clinical_Summaries.pdf.

[15] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/Stage2_EPCore_7_PatientElectronicAccess.pdf.

You may wonder why providers are measured on activities actually done by their patients. Research and experience show that patients are far more likely to embrace electronic self-care tools if their provider encourages them to do it and even provides some help in getting started.[16] As a result of VDT, many more provider practices will be offering encouragement and even the needed tools integrated with their EHR.

What might Stage 3 be like? That is now becoming clearer because, just as this book was going to press, CMS released its proposed Stage 3 rules.[17] The public comment period will end on May 29, 2015. We will discuss the proposal at a high level but, first, some background. Once the incentive payments have been made, ONC loses a great deal of leverage and there is also doubt about its continued funding by Congress. Stage 3 might thus be the last opportunity for ONC to raise the bar on interoperability. I indicated earlier that many observers feel that this bar was set too low in the first two stages, so a lot of attention is being paid to the details of Stage 3. In April 2014, the Agency for Healthcare Research and Quality (AHRQ) released *A Robust Health Data Infrastructure*, a proposal by a task force named JASON, a group of some 30–60 scientists who have been advising the government on matters of science and technology since 1960. It argues for a set of interoperability requirements that should be placed on EHR vendors under Stage 3. Early on, it states that a "meaningful exchange of information, electronic or otherwise, can take place between two parties only when the data are expressed in a mutually comprehensible format and include the information that both parties deem important. While these requirements are obvious, they have been major obstacles to the practical exchange of health care information." We'll return to some of the technical details later on but, in essence, the report proposes a national health IT architecture that provides "a migration pathway from the current legacy software used to store and process EHRs to the future system of broad interoperability. This pathway could be provided by published application program interfaces (APIs) mandated through the Centers for Medicare and Medicaid Services (CMS) Stage 3 Meaningful Use program." [18] In essence, this means that the data now often locked up in proprietary EHRs could be accessed using a standard programming approach that would create a platform over which apps could be written for providers, patients, payers and any other entity with a legitimate need and permission. The obvious analogy is the apps running on a smartphone that can access data stored on that phone or elsewhere. The implications for practicing physicians could be truly transformational, as we'll see later on.

In October 2014 the final report of the JASON Task Force convened by CMS and ONC supported most of the JASON report's recommendations.[19] Later on

[16] http://www.healthit.gov/sites/default/files/recservicelinecasestudy_helpingprovidersengagepatients.pdf.

[17] https://s3.amazonaws.com/public-inspection.federalregister.gov/2015-06685.pdf

[18] http://healthit.gov/sites/default/files/ptp13-700hhs_white.pdf.

[19] http://www.healthit.gov/facas/sites/faca/files/Joint_HIT_JTF_JTF%20HITPC%20Final%20Report%20Presentation%20v3_2014-10-15.pdf.

we'll discuss the Argonaut Project which may settle the matter long before then through an industry collaboration of health IT vendors and leading healthcare organizations. Here is a brief summary of the proposed Stage 3 rules adapted from a March 24, 2015 blog post by Dr. John Halamka.[20]

Provider Requirements:

e-Prescribing: An increase in the percentage of prescriptions that must be e-Prescribed to 85 % for eligible providers and 25 % for eligible hospitals. Controlled substances can be included in calculating the percentages in states where electronic prescribing of these drugs is legal.

Clinical Decision Support: Implement five clinical decision support tools tied to four quality measures; and use drug–drug and drug-allergy interaction alerts for the entire EHR reporting period.

Computer-based Physician Order Entry (CPOE): Must be used for at least 80 % of medication orders, 60 % of lab orders, and 60 % of diagnostic imaging orders. Order entry by "scribes" counts toward these goals.

Patient-facing Requirements:

Electronic Record Access: 80 % of patients must be able to access their records either through the View/Download/Transmit (VDT) function or through an ONC-certified API; and 35 % must have access to patient-specific educational resources. Only patient access but no patient action is required to meet these objectives.

Active Engagement: 1) 25 % of patients must access their records either through VDT or through an ONC-certified API; 2) 35 % of patients must receive a clinically-relevant secure message; and 3) providers must incorporate patient-generated data from "non-clinical" settings (such as home health) for 15 % of patients. The last of these is a new direction for Meaningful Use and is a technology challenge we will explore in detail later on.

Interoperability Requirements:

Health Information Exchange: Providers must send an electronic summary for 50 % of transitions of care (TOCs) and referrals; they must receive an electronic summary for 40 % of TOCs and referrals; and they must perform medication/allergy/problem reconciliation for 80 % of TOCs and referrals.

Public Health and Clinical Data Registry Reporting: "Active engagement" is required for: 1) immunizations; 2) syndromic surveillance; 3) reportable conditions case reporting; 4) public health registries; 5) nonpublic health registries; or 6) electronic lab reporting. Eligible providers must choose three from the first five of these, and eligible hospitals need to choose four from the entire list. Given the variability in public health reporting capabilities across the country, CMS defines "active engagement" broadly to include either registering, testing, or actually reporting public health and registry data.

Incentive Payments: Beyond these technical programs, there is also a clear need for incentives that can be divided into at least three components: 1)

[20] http://geekdoctor.blogspot.com/

reimbursement for the expense of the system; 2) some **ongoing financial incentive** to use the system properly and 3) incentives to **change practice models** to those that increase efficiency and quality.

The Medicare and Medicaid incentive payments programs were designed to do the first. The Medicare payment program may do the second through **penalties** for providers and hospitals who don't achieve Meaningful Use. It is also possible, or perhaps even likely, that private health insurance companies may start funneling patients to providers who have achieved Meaningful Use and the prospect of that is yet another incentive. The third needed incentive is often referred to as **pay-for-performance**, which implies further rewarding providers for improving care efficiency and quality through the use of health IT and utilizing new, more efficient and effective care models.

The Medicare and Medicaid incentive payments are based on achieving the stages of Meaningful Use. The amount providers can earn is tied to the number of Medicare patients in their practice. To qualify for Medicaid incentive payments, 30 percent of a practice must be in that program (except for pediatricians where the threshold is 20 percent). The details for Medicare are more complex.[21] In both programs, payments are tied to achieving the stages of Meaningful Use, however in the Medicare program, the earlier a provider starts, the more they can earn. There will likely also be reductions in Medicare payments for providers that haven't achieved Meaningful Use by 2015 although, like all federal programs, this could be changed by Congress as the deadline approaches. There are no penalties under the Medicaid program.

Results to Date: These programs have succeeded to a greater degree than many observers thought possible. As of the most recent report (November 2014), some 337,861 Medicare-eligible and 166,670 Medicaid-eligible health professionals (refer to the glossary for an explanation of eligible professionals) and some 4,789 hospitals had registered. This is over 95 percent of both groups. Of these, 276,517 Medicare-eligible professionals and 133,457 Medicare-eligible professionals had received Stage 1 incentive payments. In total, $26 billion in incentive payments had been made including $8.7 billion to Medicaid-eligible professionals and $17.3 billion to Medicare-eligible professionals.[22, 23]

At this same time, 6,219 Medicare-eligible professionals and 342 eligible hospitals (including rural critical-access hospitals with under 25 beds) had received payments under Stage 2. These Stage 2 numbers require some explanation. To be eligible, providers must have completed two years of attestation for Stage 1, something only 56 percent of eligible hospitals and critical-access hospitals and 42 percent of eligible professionals had done as of November 2014. By then, 16,455 eligible professionals and 1,681 eligible hospitals had attested for Stage 2. So even though the numbers are relatively small, by November 2014, nearly 80 percent of

[21] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Basics.html.

[22] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/November2014_SummaryReport.pdf.

[23] http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/May2014_SummaryReport.pdf.

Stage 2-eligible hospitals and 60 percent of Stage 2-eligible professionals had attested to Stage 2.[24] CMS and ONC say they are optimistic these trends will continue, but it is too early to be certain.

The Stage 2 numbers might have been even higher if not for the requirement to send an electronic clinical summary (typically a CCD), at least 10 percent of the time, to the location to which a patient is being referred in a transition of care. If the location receiving the patient does not have a 2014 Certified EHR, they might not be able to accept the electronic document. Therefore, CMS issued the 2014 Flexibility Rule that in part, allows providers and hospitals to attest to Stage 1 for a third year without moving to Stage 2 if they have been unable to meet the 10 percent requirement for sending an electronic summary because other settings of care with which they work, and to which they refer patients, have not implemented 2014 certified EHRs and therefore cannot receive the electronic summary.

Hospital adopters are divided into three groups based on the sophistication of their EHR implementation. As of 2013, some 70 percent of non-federal hospitals had achieved one of these: 10.9 percent with a Basic EHR without Clinician Notes, 33.7 percent with a Basic EHR with Clinician Notes and 25.7 percent with a Comprehensive EHR that includes electronic entry for virtually all clinical orders, visualization of virtually all results and a comprehensive suite of CDS tools.[25]

By 2013, 78 percent of office-based physicians used *any type* of EHR system, up from 18 percent in 2001. Of course, these systems vary substantially in their capabilities. Figure 3 shows the percentages of physicians with an EHR capable of providing for selected Meaningful Use Stage 2 objectives as of the end of 2013. Up-to-date information is available on an interactive HIT adoption dashboard at the ONC web site.[26]

A 2011 survey by the National Center for Health Statistics, part of the CDC, examined the demographics of physician EHR adopters. As you might expect, younger providers are more likely to adopt but the difference isn't striking. What is striking (and has been in all prior studies) is the difficulty smaller practices have in adopting, presumably due to lack of financial as well as technical resources since many of these are in rural, poor or underserved areas. Some 86 percent of practices with 11 or more physicians were adopters, with solo practitioners at 29 percent and those in the middle at around 60 percent. To help these smaller, often rural practices, ONC funded 62 Regional Extension Center (REC) programs to provide special assistance to primary care practices with fewer than 10 providers. As of this writing, over 147,000 providers are enrolled (including 41 percent of primary care practices nationwide and 51 percent in rural areas) and more than 124,000 are "live" on an EHR. Of these, more than 70,000 have demonstrated Meaningful Use Stage

[24] http://www.healthdatamanagement.com/news/77-Percent-of-Eligible-Hospitals-Have-Attested-to-Stage-2-MU-49626-1.html.

[25] http://www.healthit.gov/sites/default/files/oncdatabrief16.pdf.

[26] http://dashboard.healthit.gov/.

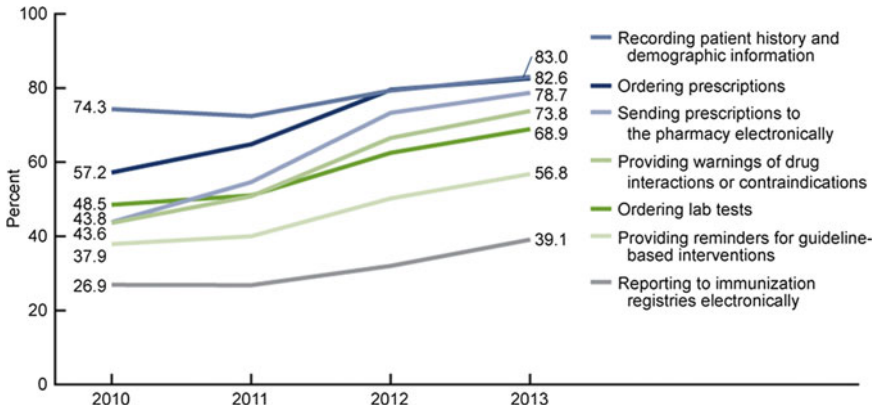


Figure 3: This graph shows the percentage of physicians with an EHR capability sufficient to meet selected Meaningful Use Stage 2 objectives (http://www.cdc.gov/nchs/data/databriefs/db143.htm). (Source CDC)

1. Eighty-five percent of REC-enrolled providers are live on an EHR versus 62 percent in the general provider population.[27]

Practice ownership is also a significant factor, with HMO physicians (usually employed and provided an EHR by the HMO) at 100 percent adoption, and physician-owned practices at around 50 percent. Once they adopt, substantial majorities of providers report administrative, financial and, most importantly, clinical benefits. Around 70 percent of nonsurgical specialists report improved clinical communication—a key result given the problems created by our fragmented approach to managing chronic disease.[28] However, in a late 2013, Rand/American Medical Association (AMA) survey, physicians complained about "poor EHR usability, time-consuming data entry, interference with face-to-face patient care, degraded clinical documentation (as a consequence of template-based notes) and inefficient and less fulfilling work content." They also complained about the inability to exchange health information. Despite these shortcomings, many of which we will discuss in detail later, only 18 percent of providers said they would "prefer using paper medical records instead of electronic records."[29]

Clearly an optimist would say that the cup is half full, while a pessimist might complain that it's half empty, but the shift to wider use of electronic records in the last five years has clearly been dramatic. Later on, we'll discuss approaches that offer the potential to overcome many of the common and well-founded physician complaints about EHR design. First, we'll discuss the core technologies of health IT and learn more about the barriers to interoperability and the exchange of health information.

---

[27] http://www.healthit.gov/providers-professionals/regional-extension-centers-recs.

[28] http://www.cdc.gov/nchs/data/databriefs/db98.pdf.

[29] http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR439/RAND_RR439.pdf.

Technologies for Sharing Health Information

Electronic health information exchange (HIE) allows doctors, nurses, pharmacists, other health care providers and patients to appropriately access and securely share a patient's vital medical information electronically—improving the speed, quality, safety and cost of patient care.

—HealthIT.gov

We know that coordination among the many providers who care for a patient with chronic disease is of paramount importance in improving quality and efficiency. In the age of the Internet, in which we're all used to information anytime and anyplace, you would probably think that, once healthcare records are digital, exchanging them would be easy. The reality is quite different and health information exchange is another topic that illustrates the often skewed incentives within our complex healthcare system.

The need for HIE is apparent. Beyond care coordination, HIE is a key tool for understanding and improving the health of patients at the practice or an even broader level (population health), it can become a platform for involving patients directly in their own care (patient engagement), and it can also be used for data aggregation for public health and other secondary uses, including research. However, interoperability remains a major technical issue because of the multiplicity of EHRs in use in the U.S. There are hundreds of certified EHRs and, although a dozen or so represent around 75 percent of all installations, the rest are spread among many vendors. You can see this for yourself at ONC's dashboard,[1] a site I recommended earlier in the book. For the most part, these systems were not designed to record data in a standard fashion or to share data with other systems using some common convention—they are not interoperable. While EHR certification creates some degree of standardization and interoperability, building

[1] http://dashboard.healthit.gov/quickstats/pages/FIG-Vendors-of-EHRs-to-Participating-Professionals-2014.html.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_5

technologies to facilitate more extensive data sharing remains a major challenge. The JASON report's call for a standard data model with access through the same data sharing technologies (APIs) used elsewhere on the Internet would clearly revolutionize the HIE landscape and it appears to be rapidly gaining support and momentum, but it is futuristic at this point, and still possible that it won't happen at all. For now, we'll discuss current approaches to using technologies usually developed more specifically for healthcare.

Accurate patient identification across providers and health systems is a fundamentally important HIE subchallenge. Even if data is standardized and accessible, it is critical to know if John Smith at hospital A is the same or a different patient from John A Smith at hospital B. France solves this problem by issuing a special health smartcard. Other countries have a national health ID number. American distrust of a strong central government makes a national identity number or card a controversial political issue that is unlikely to be resolved anytime soon. For now, the only solution is technology designed specifically to deal with this issue. In most cases, cross-enterprise HIE depends on the use of an enterprise **Master Patient Index** (MPI) based on sophisticated matching algorithms designed to identify patients as accurately as possible. Even the best MPI isn't 100 percent accurate, so some means must be provided for human intervention to deal with identity issues the software can't accurately determine (well-designed software will know this and bring those cases to the human's attention).

Another subchallenge is figuring out where a specific patient's clinical data is stored. Multichronic disease patients seeing many different specialists can have parts of their record scattered across many EHRs. Suppose one of their specialists or a PCP (should they even have one) wants an integrated view of their care. Where do they go to get all the needed information? Special **Document Locator Systems** (DLS) are part of some HIEs and facilitate this type of query by indexing where a patient's images, lab tests and other important clinical information are stored. Typically the receiving systems which store the data send a transaction in a special format to the DLS each time a new package of information is received and stored. The DLS uses these messages to build indices which can have a tremendously positive impact on the ability to accurately and rapidly find all of a patient's clinical data. The new Meaningful Use Stage 2 VDT requirements mean that the patient might create their own integrated record by obtaining electronic summaries from each of her providers. We'll look at the concept of a patient-generated and controlled record or data repository in more detail later.

Integrating the Healthcare Enterprise (IHE) is a non-profit organization based in Illinois but with global reach. Its goal is to promote health information exchange (HIE)—I'm not sure if these similar acronyms are a coincidence – through specifications, tools and services for interoperability. IHE defines three models for document exchange and the role and function of the DLS differs in each of them. In the first, there is no DLS so it depends on the records expertise of the data provider.

In the second, a central DLS serves as an archive. The third uses multiple, interacting systems. The models are:

Direct Push: clinical content in the form of documents and metadata (the identifying data about a document such as a patient identifier, document identifier or author information) is sent directly to a known recipient or published on media (such as a CD or USB stick) for delivery.
Centralized Discovery and Retrieve: a centralized document locator is used to discover the location of documents enabling retrieval of the document from a custodian who has registered existence of the document with the centralized locator.
Federated Discovery and Retrieve: a collection of peer entities are enabled to query each other to locate documents of interest, followed by retrieval of specific documents.[2]

For our purposes, HIEs can be classified in four different ways:

Scope: What geographic area do they cover?
Status: Where is each HIE in the process from an idea to full functionality or even innovation?
Architecture: How are the systems that support the HIE organized?
Functionality: Of all the things that are possible, what does each HIE do?

HIE Scope: This can vary substantially. An HIE can serve a single healthcare enterprise (health system) that may consist of one or more hospitals, associated physician practices and other entities such as nursing homes or rehabilitation facilities. These entities may have EMRs from different vendors, so interoperability can be an issue even within a single health system. Some health systems extend their HIE to practices that refer patients to them in an effort to make that easier and, thereby attract more referrals. This is often called a service-area HIE. In an effort to better position the health system to contract under outcome-based reimbursement arrangements, such as an ACO, it is increasingly common for the system to acquire these outlying medical practices and they would almost definitely be included in the enterprise HIE.

There is a clear business case for HIE in a healthcare enterprise—grow revenue by making it easier for physicians to refer patients to the enterprise and more convenient for patients to get all their care within it. Increasingly such connected enterprises offer both physicians, and their patients, special electronic tools (including apps) to interact with them for functions such as scheduling and the reporting of lab test results. In most business activities, we assume that "bigger is better" and thus, more sustainable. With HIE, the business case typically gets weaker once it expands beyond a health enterprise. As a result, for the most part, regional, statewide and nationwide HIE remains an economic more than a technical

[2] http://www.ihe.net/Technical_Framework/upload/IHE_ITI_White-Paper_Enabling-doc-sharing-through-IHE-Profiles_Rev1-0_2012-01-24.pdf.

challenge. We'll discuss two notable exceptions, with different but effective models, later in this section and there are a few others, but not nearly enough so these success stories are worth careful consideration.

HIE Status: eHealth Initiative is a broadly based, collaborative, not-for-profit organization focused on improving healthcare quality and efficiency, including support of HIE. As part of its activities, it tracks the status of HIE implementations across the country against its seven stages of HIE maturity—start-up, getting organized, planning, piloting, operability, sustainability and innovation. There has been a recent and rapid shift toward more mature HIEs, a clear sign of progress, but one that comes during an era of substantial federal financial support for HIE. Many individuals and groups are concerned about what will happen as that support ends and are attempting to develop longer term strategies.[3]

HIE Architecture: There are typically no clear boundaries separating the common HIE technology architectures. For our purposes, the main architectures are **centralized** HIEs which involve a master repository of patient data, **federated** HIEs where all patient data remains where it was recorded and **hybrid** HIEs where clinical data is either stored locally with centralized services to help locate it or, in an alternate approach, there are "data lockers" where any centrally stored clinical data is segregated by institution and each contributing institution controls access to its data. These architectures closely align with the IHE models for document exchange that we just discussed. The new Direct e-mail-based HIE technology specifically supports IHE's Direct Push approach to HIE in which providers send clinical documents to known and trusted entities. We'll look in some detail at Direct later on because of its simplicity and low cost which make it particularly interesting to small physician practices outside a healthcare enterprise.

HIE Functionality: The Office of the National Coordinator for Health IT defines a fourth HIE classification—function or purpose. The first category in this classification is **directed exchange, which is** typically used to send and receive information among providers for the purpose of care coordination. The second is **query-based HIE,** the purpose of which is data aggregation among providers, typically for reporting, public health or research purposes. The third category is **consumer-mediated** exchange involving data aggregation for the empowerment and increased involvement of patients, the key objective of the VDT requirement.

Indiana Health Information Exchange (IHIE): IHIE is the premier U.S. example of a large-scale (statewide), centralized HIE with multiple components and services. IHIE's **Indiana Network for Patient Care**—a single, virtual, community-wide health record—delivers more than a million transactions per day from a data warehouse that contains five billion pieces of clinical information for some 10 million patients. Each patient's data is presented in a familiar format that mimics a typical EHR. **Docs4Docs** provides a web portal to make it easy for physicians to obtain documents such as lab test results or reports on images they've ordered. **Quality Health First** is a population-based quality reporting system we'll look at

[3] http://www.nga.org/files/live/sites/NGA/files/pdf/1103SUSTAININGHIETOOLKIT.PDF.

later on in the section on population health. **ImageZone** is a repository that allows physicians, from their office computers, to see any image-based study done on a patient anywhere in the IHIE coverage area. **ACO Services** is a newer offering that provides tailored management tools for groups of providers (which might include one or more hospitals and a number of physician groups) managing care quality and cost within a pay-for-performance reimbursement system.

IHIE's ability to provide these sophisticated services depends on aggregating data from many sources that are not inherently interoperable. The key technical component is **data governance**, a sophisticated engine that bridges differences in data syntax (that is, data structure, the way data is represented) and semantics (the meaning of that data, no matter how it is represented) across these sources to create a single database in a standardized format. The goal is to make it appear as though everyone is using the same electronic record system.

Kansas Health Information Network (KHIN): KHIN is a robust, statewide and private, not-for-profit HIE with a unique sustainable business model. Established in 2010 by the provider community—the Kansas Medical Society, the Kansas Hospital Association and the Wichita Health Information Exchange—as of January 2015, it had over 950 member organizations including 102 hospitals and their affiliated physicians, 160 independent physician practices, three payers and numerous other healthcare provider organizations. Of these, 358 are sharing data and 275 are in testing.

We'll discuss models for obtaining patient consent for data sharing in the next section but KHIN manages a patient opt-out process at the state level. In this model a patient's data is shared unless they specifically request that sharing not take place. To assure adequate patient notification, providers are required to change their notice of privacy practices to indicate they participate in KHIN and to post notices and provide state-developed brochures that advise patients of the opt-out process. To date, approximately 550 (less than 0.4 percent) of patients have opted out. Kansas has also passed legislation that supports health information exchange by reducing or eliminating provider liability concerns related to information sharing.

As a result of this broad provider support and patient participation, KHIN has over 1.4 million unique patients in its databases and KHIN members can access data on over five million patients through the connections KHIN has with other HIEs and health systems. Available data includes CCD clinical summaries as well as discrete data including problem lists, reports and notes, procedures, allergies, immunizations, encounters, demographics, medications, labs as well as insurance and provider information.

In the last four months of 2014, KHIN supported 199,443 patient queries. Use of the HIE is encouraged by a 2.5 percent financial incentives from Blue Cross and Blue Shield of Kansas to providers when they access KHIN patient records and a .05 percent incentive payment to hospitals for connecting to KHIN with an ADT (Admit, Discharge, Transfer) Health Level 7 (HL7) data feed. Kansas law allows payers to become full members of the HIE which then provides aggregate data back to them for care coordination, risk adjustment and risk stratification.

KHIN electronically reports data to all of the Kansas state immunization, cancer and infectious disease public health registries and has reported over 1.5 million messages to the CDC's national syndromic surveillance system (BioSense 2.0,[4] a system we'll discuss in more detail in the section of aggregating data) via the first real time HL7 connection to it. KHIN also reports positive lab results for reportable diseases to EpiTrax, the Kansas communicable disease surveillance system. KHIN is also a member of eHealth Exchange, the proposed national HIE operated by Healtheway, a non-profit, public-private collaborative,[5] and is actively pursuing connections with other eHealth Exchange members.

KHIN provides a free personal health record (MyKSHealth) to all patients that receive health care in Kansas. Moreover, CCD patient summaries automatically populate each patient's PHR after a care episode so they have an integrated record of their care across all of their providers. Patients can also deconstruct their CCDs to create integrated problem lists or medication lists across providers. MyKSHealth offers Direct messaging between patients and their providers and the ability for patients to add their own data and track it across time.

KHIN provides real-time alerts (such as emergency department visits or hospitalizations) to providers and payers. KHIN has developed quality metric reporting and has a secondary data use policy in place that allows outside organizations to request access to the data for approved uses such as population health and quality reporting.

KHIN also had a unique approach to initial funding which came from medical liability insurer, KaMMCO (Kansas Medical Mutual Insurance Company), as well as from ONC. As a result of its robust services and incentives for use of the HIE provided by payers and KaMMCO, KHIN is now a fully sustainable health information exchange with membership revenues (from both providers and payers) covering all of its expenses, including technology vendor costs and operations.

Other than HIEs operated by large organizations such as the Kaiser Permanente HMO or the VA, it is difficult to find examples of centralized HIE at the scope and sophistication of IHIE or KHIN. Why have Indiana and Kansas been able to do it when so many others have not? Part of the answer was a model for startup financing. We've discussed the model in Kansas. In Indiana it was the Regenstrief Foundation which supported the development of the technology and the HIE over a period of many years. Sam Regenstrief invented the front-loading dishwasher and the story of his decision to support digital healthcare and the technologies and companies that evolved from it is told in *From Dishwashers to Digital Medical Records*, an entertaining publication from BioCrossroads, an economic development organization in Indiana.[6]

[4] http://www.cdc.gov/biosense/biosense20.html.

[5] http://healthewayinc.org/.

[6] http://www.biointellex.com/wp-content/uploads/2014/08/BioCrossroads-HIT-Report-Feb-2011-Final.pdf.

CONNECT: There are a number of commercially supported, centralized HIE technologies. As we've implied, developing a sustainable business model for HIE beyond a health enterprise is challenging, so to encourage more expanded coverage, the federal government sponsored the development of CONNECT, an open-source solution for centralized (or other models) of HIE. CONNECT is a very robust, complex system that uses web services (technologies similar to those recommended by JASON and which simplify information exchange between computers on the Internet) to facilitate connectivity with other systems. Its architecture is modular with some elements required, some optional (typically those subsystems a hospital might already have) and others customizable. It also provides a gateway to the national eHealth Exchange.

CONNECT adoption is not widespread but has been growing. Many adopters are federally supported health systems or organizations such as the CDC, VA or Department of Defense (DOD). Others are state HIEs or private entities such as the Marshfield Clinic, mentioned earlier as the major success story from the Physician Group Practice demonstration project. You can get more information on adoption (and about CONNECT, but that information is highly technical) by visiting the CONNECT web site.[7]

Direct: In so-called federated exchange, all patient data typically stays within the source EHR. This facilitates provider participation and simplifies data governance, but it also creates technical challenges. Earlier in this section we listed Direct Push as one of IHE's technical three approaches to document sharing. Direct (not to be confused with IHE's use of the term) is a technical framework that allows a provider to send electronic documents to another known and trusted provider. The concept was first put forward in December 2009 in a blog post written by Dr. David McCallie, Senior Vice President of Medical Informatics at Cerner Corporation, one of the major suppliers of large-scale systems to hospitals and health enterprises.[8] More recently he served as cochair of the JASON task force.

Five years later, in an interview for my blog, he explained that back in 2009, he "was concerned that we didn't have a national standard for simple encrypted e-mail-like exchange between providers." He added that: "It occurred to me that starting with a simple 'push' model of exchange would greatly simplify the governance and policy decisions necessary for what we wanted—universal exchange—as ubiquitous as the fax machine."[9]

Work began in 2010 on defining such an HIE approach. The original "use case" (i.e., the scenarios around which HIT systems are developed) was replacing the fax machine as the way two physicians commonly share data, just as Dr. McCallie had hoped. This effort produced Direct. Today, as a result of the requirement that federally funded HIEs include a Direct service, sending a patient record to another

[7] http://www.connectopensource.org/adopters.

[8] http://blogs.gartner.com/wes_rishel/2009/12/01/guest-david-mccallie-on-simplifying-interop/.

[9] http://www.informationweek.com/healthcare/electronic-health-records/direct-5-years-of-simplifying-health-information-exchange/a/d-id/1317777?.

provider increasingly uses a Direct service provided by a state HIE or even by an HIE operated by a local or regional health enterprise.

In addition to the requirement that HIEs support Direct, Meaningful Use 2014 now also mandates that EHRs offer integrated support of Direct messaging. However, it is currently up to each vendor to interpret what that means. Cerner's PowerChart provides what I consider an excellent example of meaningful and useful integration, as shown in Figure 1. The key innovation illustrated here is workflow and process integration. Without it, a PCP wishing to achieve Meaningful Use and who wants to refer the current patient to a specialist provider, might have to first export a special electronic document (i.e., CCD) from that patient's electronic record to support care of that patient in what is termed a transition of care. We'll discuss the CCD in more detail as part of the section on interoperability standards but, for now, think of it as the electronic equivalent of a patient's clinical summary, including their basic demographics, problem list, medications, lab tests and other key health data. They might then have to leave the charting session; go to an e-mail program connected to a special Direct server (called a Health Information

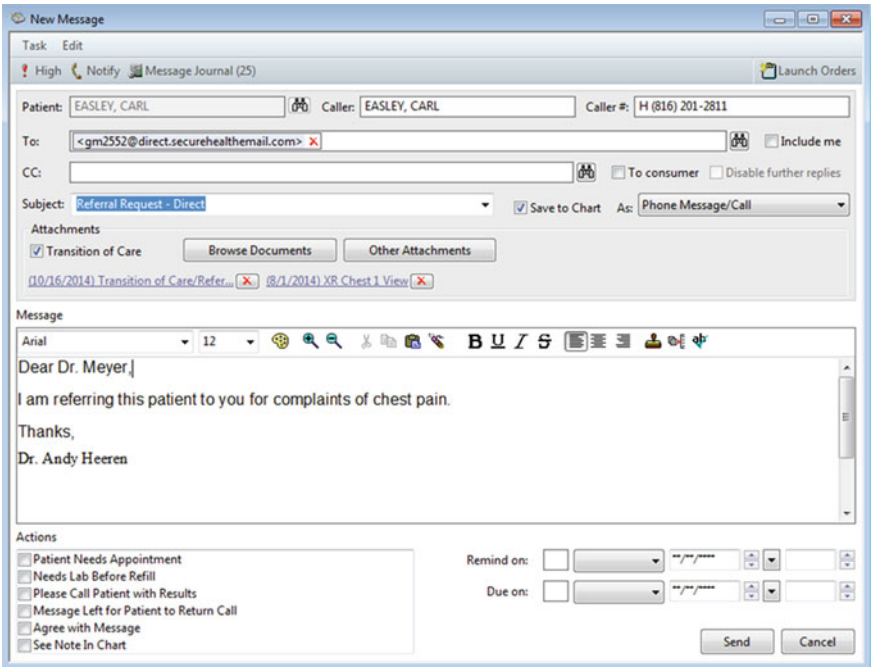


Figure 1: Direct HIE is now an integrated component of Cerner's PowerChart EHR. As shown here, a primary care provider can initiate a Direct message to a specialist to whom they wish to refer the patient. This is done from within the charting session and the necessary electronic CCD is automatically produced and attached to the Direct e-mail message. In this example, the referring provider is also sending a chest x-ray. This might well eliminate a duplicate procedure when the patient sees the specialist, lowering healthcare costs. (Courtesy Cerner Corporation, All Rights Reserved)

Service Provider or HISP); and then initiate the e-mail, attach the document, address the e-mail and send it.

As shown in Figure 1, that entire process can now be done *within an electronic charting session*, and many parts of it can be automated since the patient's identity, the reason for the e-mail and the identity of the sending provider are all known or can be simply indicated by a checkbox.

The importance of this kind of workflow and process integration cannot be overemphasized if electronic record systems are to improve efficiency in addition to providing digital records. It is also worth noting that, in this integrated scenario, many opportunities for incorrect data entry or other mistakes are eliminated, improving care quality and patient safety.

The Direct HISP is the e-mail input/output station which also provides each physician it serves with a special Direct e-mail address to establish **trust** so that all providers, patients or other entities using Direct know their e-mails will only go to the intended recipient, and that all recipients are who they say they are. The trust process is policy-based, so each HISP decides what is required before it assigns an address to each provider. This creates an issue of establishing trust among hundreds of HISPs across the country that we'll discuss later in this section. Many EHR vendors now provide a HISP to their customers.

The format and content of the record attached to a Direct e-mail message can be whatever the sending physician chooses. It could be a scan of a paper record, it might be a PDF created from a digital record or it could be the special XML-formatted Transitions of Care patient summary illustrated in the attachments section in Figure 1.

XML uses tags to identify each data item to a computer (an example of meta-data) and is widely used on the Internet to specify the meaning of a data item. We'll look at it in more detail in healthcare later on, but a simple example might be: <CodingSystem>NDC</CodingSystem> in which the tags in brackets tell the computer that "NDC" is the name of a coding system (actually the National Drug Code for medications). The first of these tags, <CodingSystem>, specifies the name/purpose and the beginning of that part of the XML. The same tag preceded by a forward slash, /, indicates the end of that part of the XML. While an individual data field is usually simple, an entire CCD is complex. The more technically inclined reader can see that in an example posted on the Internet.[10]

However, whatever it is, the record is encrypted and attached to a secure e-mail, so the information can't be viewed in transit, meeting the **security** requirement for health information exchange. The encryption of the record before it is sent and decryption at the receiving end is done using the Public Key Infrastructure (PKI), a common technology you've used if you've ever done banking, bought securities or paid bills via the Internet. We'll look at PKI in more detail later on.

Privacy, a third requirement for HIE, is typically met for the specific purposes of routine patient care when patients consent to sharing their records with other

[10] http://www.ehrdoctors.com/CCD_Example_Meaningful_Use_Phllip_Jones.xml.

physicians involved in their care, This consent usually also encompasses operational use of the data (e.g. billing) and its use for certain reporting purposes (e.g. quality reporting) by the medical practice. This is yet another issue we'll look at in more detail later on.

Workflow and Process Automation: The recipient provider gets the e-mail in their special Direct inbox and can review it if it's a human-readable document such as a PDF or paper scan. If it's an XML-formatted CCD, the record could be parsed and put right into the patient's chart at the recipient provider's office. Parsing it and merging it into an existing record is another example of workflow and process automation with the potential to save time and reduce the chance of mistakes. Figure 2 illustrates this process and a reconciliation step through which the data in the incoming record is grouped and compared by the computer with similar data in the recipient provider's EHR, further improving the process for example, by helping to eliminate duplicate prescriptions (a common problem in complex multiple chronic disease patients being seen by several providers) and by updating out-of-date or incomplete information at the recipient providers end, further improving care coordination, eliminating duplicate tests and procedures and possibly even avoiding clinical errors.

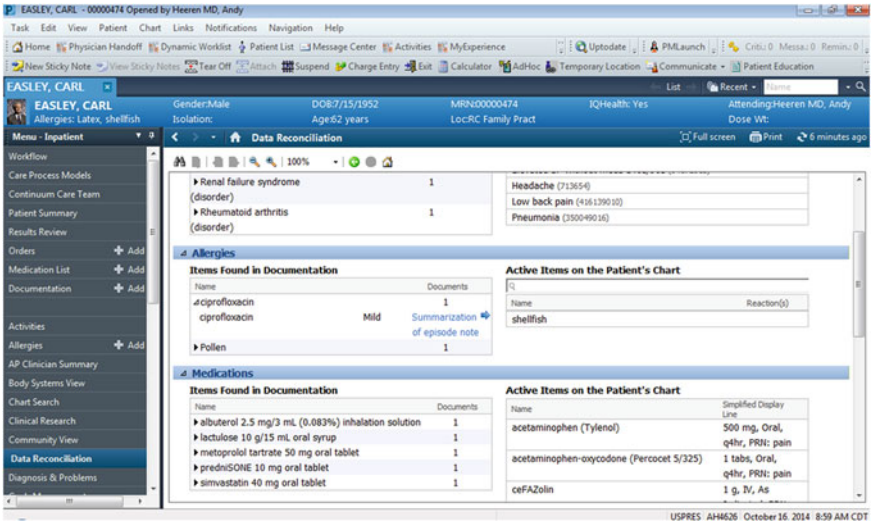


Figure 2: If the Direct attachment is a computer-readable document, such as an XML-formatted CCD, it is possible for the receiving computer to parse the information into discrete data items and offer to reconcile them with similar information in the receiving provider's EHR. An example, as shown here, is Cerner's integration of Direct e-mail into their PowerChart EHR. Problems, allergies and medications from the CCD are grouped side-by-side with the similar data from the receiving provider's EHR. This has the clear potential to save time, eliminate duplicate data entry and avoid mistakes due to inaccurate, incomplete or out-of-date records when patients are seen by multiple providers. (Courtesy Cerner Corporation, All Rights Reserved)

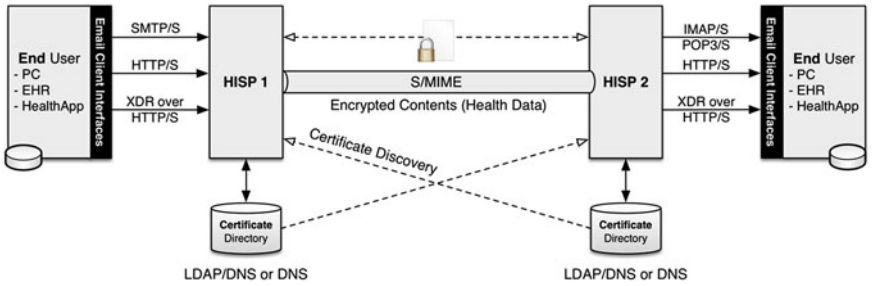


Figure 3: The typical multi-HISP Direct scenario uses many existing web technologies for e-mail transport and certificate management. In this diagram the sending practice is on the left and the recipient practice is on the right. Note in particular, that they each are associated with their own HISP. While XDR is shown here as an alternative web services based transport mechanism, we won't discuss it further. (Courtesy Dr. Myung Choi, Georgia Tech)

Direct Technical Concepts: *Nontechnical readers may want to skip this paragraph*. The Direct e-mail-based transport technologies are relatively simple if both providers are using the same HISP. In the more likely and more complicated case (Fig. 3), where they are connected to separate HISPs, S/MIME is used to encrypt the attachment containing the patient record, SMTP is used to send secure messages from an e-mail client like Microsoft Outlook and HTTPS is used for sending webmail (such as Gmail). The sender's HISP can discover (find) the recipient provider's X.509 certificate (their public key) using the Domain Name System (DNS) or the Lightweight Directory Access Protocol (LDAP), two technologies for managing distributed resources over a network. Once it has the recipient's public key, the HISP uses it to encrypt the message. It then uses SMTP to send the message to the recipient's HISP (having used DNS or LDAP to find it based on the recipient's e-mail address). The recipient's HISP decrypts the message, using the recipient's private key, which is normally stored in the HISP as a service to its registered providers. The recipient then retrieves the message like any other e-mail, but using the secure versions of either the Post Office Protocol (POP) or Internet Message Access Protocol (IMAP) e-mail retrieval protocols (if using an e-mail client) or HTTPS (if using webmail).

Additional Direct Technologies: Most clinical documents require associated metadata to be useful and it would ideally be packaged along with the document into a single file for attachment to the Direct e-mail. To facilitate this, IHE defines the Cross-Enterprise Document Media Interchange (XDM) profile standard that provides a format for zipping together both documents and metadata into one file that utilizes agreed-upon conventions for directory structure (directories are like

physical folders, a place to store objects) and the location of files within them. Because of this, an XDM file can be unzipped at the receiving end and software there will know where to look for its contents. This simple approach also allows for the use of encryption provided by most zip software utilities.[11] The IHE Cross-Enterprise Document Reliable Interchange (XDR) standard is an alternative to Direct e-mail-based transport (as shown in Figure 3). In Direct via XDR, EHRs, PHRs and other healthcare IT systems use Simple Object Access Protocol (SOAP) for transport, rather than e-mail. SOAP offers advantages with respect to security, reliability and assurance that messages get through. The technical details are beyond the scope of this book but are provided in the referenced document.[12]

Meaningful Use Stage 2 recognizes both of these IHE standards for provider-to-provider sending and receiving and allows for messages in one or more of these three ways:

- Standard, secure Direct e-mail only
- Direct e-mail with XDM
- XDR-based message exchange[13]

Technical Challenges with Direct: Direct is still relatively new and has some technical limitations. The first is message disposition notification (MDN). In certain situations it is very important to know that a Direct message has been received. For example, clinical laboratories are using Direct to send results back to ordering physicians. An abnormal result may require urgent attention so it is important to know it was received. Confirmation of delivery is available, although currently not very robust, but it is accomplished using the MDN technology.[14]

Trust among HISPs is another challenge. Earlier, we said that the HISP establishes trust by verifying that the providers to whom it assigns e-mail addresses are who they say they are; and it also verifies issues such as whether they are licensed in their practice state. That's fairly easy to do on a local basis. Suppose, however, that physicians in two states or two regions of the country want to use Direct. They will probably be connected to different HISPs, each of which has established its own policies for establishing trust. Nationwide, there could be hundreds of HISPs. Will each of them have to investigate and contract with all the others to make sure they can trust each other? That quickly gets unwieldy since the number of HISP-to-HISP relationships grows exponentially as the number of HISPs grows. DirectTrust[15] was

[11] http://wiki.directproject.org/file/view/2011-03-09%20PDF%20-%20XDR%20and%20XDM%20for%20Direct%20Messaging%20Specification_FINAL.pdf.

[12] http://wiki.directproject.org/XDR+and+XDM+for+Direct+Messaging+Working+Version#XDR and XDM for Direct Messaging Specification.

[13] http://www.himssehra.org/docs/EHRAStage2SecureHealthTransportCertificationandMeaningfulUse.pdf.

[14] http://wiki.directproject.org/file/view/Implementation+Guide+for+Delivery+Notification+in+Direct+v1.0.pdf.

[15] http://www.directtrust.org/.

established to provide a solution by establishing a **trust community** in which the entities involved in Direct are accredited so that all the other entities can trust that they are following policies and practices agreed to by the entire DirectTrust community. At present, the network serves over 28,000 healthcare organizations and over 450,000 individual Direct users who are exchanging hundreds of thousands of Direct messages and attachments nationwide each month. These numbers are growing rapidly. Use cases include transitions of care from the hospital to long-term care, reporting sexually transmitted diseases to public health, referring VA patients to local private providers and even reporting data collected by sensors and devices in patients' homes.

Another challenge is **pull**, automating the *retrieval* of data using Direct. So far we've only discussed the fax machine replacement use case where one provider sends a patient chart to another, for example, as part of referring a patient for specialty care. Suppose a physician wants patient data that another physician has in his EHR? Suppose a patient has started recording her blood glucose at home and her physician wants to retrieve the most recent readings to see how well controlled they are? Suppose a patient has established a personal health record and wants it automatically updated whenever there is new information in any of his EMRs? These are all examples of **pull** where the recipient, not the sender, seeks to initiate the message. Accomplishing that is still in the future, but a group of experts has proposed how it might be done using some new standards within Direct for requesting information and sending back a response. These would be specially formatted XML documents similar to constructs that already exist in quality reporting. Another interesting aspect of this proposal that's not specifically limited to Direct, is that all of each patient's data would be stored in a repository that would receive and respond to these requests, presumably based on each patient's information sharing preferences. Many people advocate some form of a patient-controlled data repository to simplify health data privacy issues.[16]

The Harvard Boston Children's Hospital was awarded one of four Strategic Health IT Advanced Research Projects (SHARP) grants by ONC to explore key issues constraining the further development of HIT.[17] It developed an app platform, originally called SMART but now called "SMART on Fast Healthcare Interoperability Resources" (FHIR), which is similar in many respects to what is proposed in the JASON report. It has posted several sample apps, including Growth-tastic!,[18] a working prototype of a pull-based Direct service that also uses some cutting-edge technologies including Blue Button+ and FHIR. We'll discuss both of these technologies later on but FHIR is very similar to the approach for information exchange recommended in the JASON report and is the likely candidate for implementing it.

[16] http://www.projecthealthdesign.org/media/file/Standard-Model-For-Collecting-And-Reporting-PGHI_Sujansky_Assoc_2013-07-18.pdf.

[17] http://www.healthit.gov/policy-researchers-implementers/strategic-health-it-advanced-research-projects-sharp.

[18] http://growth-pull.bluebuttonpl.us/.

The basic idea is that a family can subscribe and receive updated information on the development of a child as it is recorded in the pediatrician's chart. The app also provides professionally designed visualizations of the data in classic growth chart and more user-friendly "parent" formats.[19] Growth-tastic! was demonstrated as an integrated part of five distinct EHR platforms at the 2014 Healthcare Information and Management Systems Society's (HIMSS) Interoperability Showcase (the largest annual meeting devoted to health informatics) and attracted a great deal of attention. This is a clear indication of the future landscape if a national interoperability framework, as proposed by the JASON report, were to become a reality.

Summary: Health Information Exchange is essential for care coordination, managing populations of patients, public health and research. However, there are many challenges that have proven hard to overcome, despite many years of effort. Perhaps most importantly, it is difficult to establish a sustainable business model beyond the confines of a healthcare enterprise that has its own proprietary business reasons for supporting an HIE. There is also an inevitable tradeoff between functionality and cost. Centralized models, like IHIE, deal with data interoperability issues and can provide extensive services, but at a relatively high cost. New, virtually free models, such as Direct, are not yet providing services nearly as rich as IHIE but we've discussed how that may be changing. The JASON proposal could be a game changer for HIE and for healthcare in general. It suggests a day, several years out in the future, where the place where health data is stored and the proprietary ways in which it is represented may become largely irrelevant. It also suggests that innovations, once developed, might work with most, if not all, proprietary EHR systems. This could spur investment in those innovations by simplifying market access. It also suggests that providers who practice in multiple hospitals might be able to use a single app of their choice to access data for all their patients in a consistent way that they feel is most useful for their specialty and practice style. A specific example of this for retrieval of images is provided as a use case on the FHIR Resources site.[20]

[19] http://sandbox.smartplatforms.org/showcase?app=growth_charts@apps.smartplatforms.org&patients=7777701,7777702,7777703,7777704,7777705.

[20] http://www.hl7.org/implement/standards/fhir/imagingstudy.html.

Technologies to Assure Privacy, Security and Trust

In this electronic era it is essential to safeguard the privacy of medical records while insuring our privacy laws do not stifle the flow of information fundamental to effective health care.

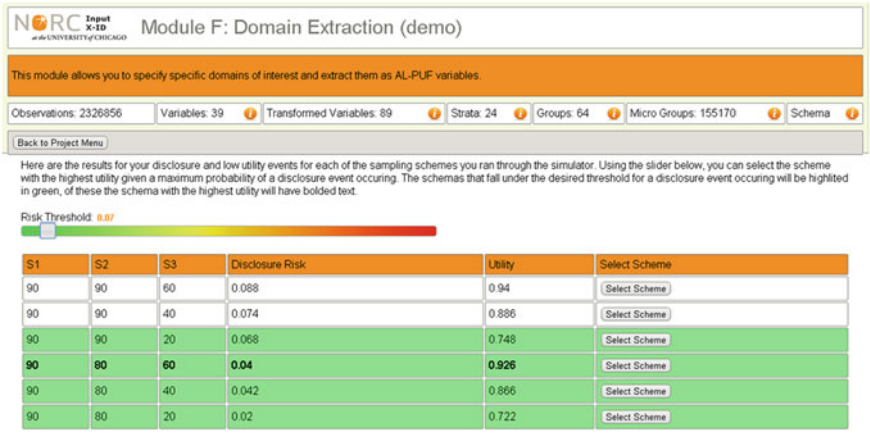
—Senator Edward M. Kennedy, The New York Times, 2007 [1]

We've repeatedly emphasized that successful management of chronic diseases in our highly specialized and fragmented healthcare system requires digital data that can be shared to coordinate care. However, patient-specific healthcare data is highly sensitive and is protected by HIPAA which calls for severe penalties for failures to properly secure so-called protected health information (PHI). PHI is clinical data about a specific patient that includes information that can be used to identify the patient. To facilitate its use beyond patient care where the individual's identify must be known, PHI can be de-identified in one of two ways. An expert can determine that there is a "very small risk" of re-identification (the "expert determination method") or 18 specific fields can be removed from each patient record (the "safe harbor method").[2] Increasingly, given the power and sophistication of modern computers, even de-identified data may be subject to re-identification. If, as is anticipated, genomic sequencing becomes inexpensive enough that it is a routine part of patient records, then de-identification may be impossible without removing it.

Analytic techniques are increasingly being applied to problems through healthcare. An interesting research effort at the University of Chicago's NORC (originally the National Opinion Research Center but now an independent social research center) has developed X-ID, an innovative and more adaptive approach to the expert determination method of de-identification using sophisticated statistical techniques.[3] The goal is to allow entities with PHI to balance the

[1] http://www.nytimes.com/2007/07/03/health/policy/03hipaa.html?pagewanted=all.
[2] http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html#standard.
[3] http://xid.norc.org/.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_6



S1	S2	S3	Disclosure Risk	Utility	Select Scheme
90	90	60	0.088	0.94	Select Scheme
90	90	40	0.074	0.886	Select Scheme
90	90	20	0.068	0.748	Select Scheme
90	**80**	**60**	**0.04**	**0.926**	Select Scheme
90	80	40	0.042	0.866	Select Scheme
90	80	20	0.02	0.722	Select Scheme

Figure 1: The X-ID system assists an expert to arrive at the right balance between the utility of a clinical dataset and the risk of re-identifying the patients within it. Risk can be minimized by subtle alterations of the data, but the more that is done, the less useful the data might be for a specific research purpose. The tool helps the expert find the right balance between these two conflicting objectives. It is operated using a simple slider control which produces different options each of which is characterized by a "Disclosure Risk" and a "Utility" score. (Courtesy NORC, University of Chicago)

risk of re-identification with the analytic utility of their data for a specified purpose. This is done by creating very small (compared to the overall dataset size) "micro-groups" of patients and publishing summary statistics (means, proportion, counts) at the micro-group level. In addition, the micro-group level data is altered sufficiently to protect individual patient data to a desired level, even to an intruder with significant information about a specific patient, while still supporting the analysis desired by the end user of the data. The specific alterations depend on data type: the mean is changed for continuous variables, such as the cost of caring for each patient; the proportion of patients is changed for a specific factor, such as diabetes, and counts are changed for other factors, such as the micro-group size). The system, as shown in Figure 1, allows its users to find the right tradeoff between the risk of specifically determining a data element about a specific patient and the value of the dataset for a known analytic purpose.

Increasingly, sharing of health data is not limited to providers. Patient engagement is a key strategy for preventing and managing disease. In a 2010 national survey by the California Health Foundation, patients reported that access to their data improved their care. Those using a specific tool—a personal health record—reported they knew more about their own health, they asked their physician a question they would not otherwise have asked, they felt more connected to their physician and they did something to improve their health. However, the survey also showed that patient adoption is still low with only one in 14 having used a PHR. Further, adopters are predominantly young, highly educated, higher income patients, a group that does not include most of the multichronic disease patients

with whom engagement is particularly important to improving outcomes and reducing costs. A big part of the reason seems to be concerns about the privacy and security of health data. Sixty-eight percent of the respondents said they are very or somewhat concerned about misuse of their digital health data.[4] Keep in mind that all PHRs are cloud-based web tools, so it is likely that any well publicized problems with data misuse on the Internet translates into patient concerns about health data being stored there. At least partially as a result, a 2013 survey showed that while 76 percent of patients want their provider to use an EHR and 70 percent want them to share the data with the other providers caring for them, similar percentages of patients expressed concerns about data privacy and security.[5]

The Domains: Protecting sensitive health data divides into three subdomains. **Privacy** in healthcare means that only individuals or entities authorized by the patient may access their data. So, a patient's record may be stored digitally in the EHR of a hospital but that doesn't mean anyone in the hospital may view it. Only authorized people can legally access it and then, only for specific, legitimate purposes. This can be a real concern when celebrities or other very important people are hospitalized. **Security** is preventing unauthorized access to the data by outsiders, such as hackers or criminals intending to use it for purposes such as generating false healthcare claims. This is a particular concern where data is being shared and must be transmitted outside of the entity where it was created. Just as with financial records, health data has value and must be protected in transit and in all places where it resides. This was an area of interest in the JASON report which called for specific protection in both the storage and transit scenarios. **Trust** means knowing that the individuals or entities with which data is being shared are who they say they are.

Assuring privacy, security and trust is a necessity for digital patient records and, in particular, for sharing those records over an HIE so we'll discuss each of them in more detail. First, as a result of their critical importance, this is an ONC Standards and Interoperability Framework workgroup priority area. These workgroups focus on a number of key health technologies and issues. The process is intended to be grounded in real-world issues—called Use Cases—that are first identified and then molded into a consensus document which drives the development of standards.

Privacy: This begins with patients. Only they or their designee(s) can consent to the use of their health information. There are a number of models for obtaining this consent. Using data without asking for consent is the **no-consent model,** but this isn't reasonable or even legal in most healthcare contexts. In the **opt-out model** consent is assumed unless the patient specifically revokes it. We mentioned earlier that KHIN uses this model, but it would be very unlikely and probably not even legal for an individual healthcare provider to use it. The converse is the **opt-in**

[4]http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/C/PDF%20ConsumersHealthInfoTechnologyNationalSurvey.pdf.

[5]http://www.healthit.gov/facas/sites/faca/files/HITPC_Consumer_Privacy_and_Security_2014-11-04.pdf.

model where no data is shared unless the patient consents. Physician's offices normally obtain patient consent under this model. CurrentCare, the very successful state-wide Rhode Island HIE we'll discuss later as an exemplar of the use of Direct, uses this model. Typically Opt-in is all or none, so patients can't share only some selected parts of their data. Despite this inflexibility, at least one study suggests as many as 90 percent of patients will consent under it. The final model, **opt-in with patient-specified restrictions model,** provides more flexibility and would probably be the preferred method if it could be managed in practice.

In routine patient care, consent is generally received to share the entire patient record with anyone involved in that patient's treatment. With this permission, data can generally be used for **treatment, payment and healthcare operations** (TPO), the exclusion that permits providers, insurance companies and other healthcare entities to exchange data for these specific purposes without obtaining permission for each use. For example, TPO allows quality reporting to be done without obtaining each patient's explicit approval. Research is not covered under TPO, so patient consent must be obtained for the use of medical record data for that purpose. This can be complex to do in practice, as we'll now discuss.

Data Segmentation: Often in a research scenario, it is only necessary for patients to share specific, relevant data but, under an opt-in with patient-specified restrictions model, how do patients indicate what data they want to share or not share? Providing this mechanism is the **data segmentation problem**. Suppose, for example, a patient has hypertension and depression among his medical problems. The hypertension is not well controlled and he has undesirable side-effects from the medications and is interested in joining a clinical trial for a new and promising drug. He wants to share the hypertension part of his data to apply for the trial, but is sensitive about the depression and wishes to restrict access to it. How can that be specified? The solution is easier if someone with clinical expertise and the patient's trust, such as the physician, does the specification on the patient's behalf. But, suppose the patient has his record in a PHR and wants to apply for the trial himself? Presented with a list of medications he may not recognize the names or know for sure which ones are for which condition. That form of sharing would be greatly facilitated if the PHR or EMR understood clinical relationships.

Figure 2 is a simple depiction of just such a record for our hypothetical patient with hypertension and depression (along with some other problems) from Dr. Jonathan Nebeker who was recently appointed the Deputy National Chief Medical Information Officer for Strategy and Functional Design at the VA. Here, the physician has clicked on hypertension and the interventions, observations and patient goals related to that problem are highlighted. Given an underlying representation of relationships such as this, the patient could simply click on hypertension and opt-in to sharing the related data under the patient-specified restrictions model. It is important to emphasize that almost no current EHRs have the necessary understanding of these relationships among the various data types they store. As we'll see later on, at least one data standard Systematized Nomenclature of

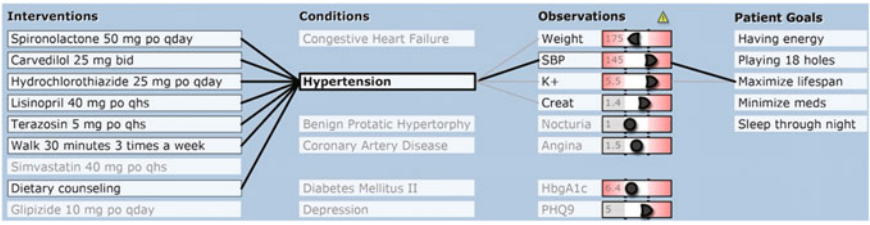


Figure 2: Dr. Johnathan Nebeker's design proposal for an EMR based on its understanding clinical relationships. (Courtesy Dr. Johnathan Nebeker)

Medicine (SNOMED) does attempt to codify these types of clinical relationships. The FHIR data model which we'll discuss later, also includes the concept of relationships. The following discussion of Decision Support for Data Segmentation (DS2) suggests another approach.

This discussion is a bit technical, so you may wish to skip it if segmenting patient's data into private or shared parts isn't of interest to you. The DS2 project at the University of Illinois' ONC-funded SHARP project[6] aimed at security for healthcare data may be an indication of future direction in this area. It derives clinical relationships conceptually similar to those assumed by the Nebeker proposal and encompassed by some existing data standards but without the need for data to be recorded in those standards. It is yet another indication of the increasing use of analytic techniques now that digital health data is increasingly available.

The redaction of a condition and its related clinical facts sometimes leaves residual facts through which the redacted condition can be inferred. To address this, the DS2 prototype software (not yet a production system ready for real-world use) uses machine learning and other techniques to redact targeted conditions automatically *along with* certain co-occurrences and comorbidities. A prototype for use in an HIE was created in collaboration with the Illinois Health Information Exchange. It consists of three modules: **SimpleConsentFormDemo** submits a specially formatted electronic patient consent document to the HIE (using the Clinical Document Architecture, or CDA, we'll discuss later on). **XDSProxy** intercepts the patient's electronic record (in a Continuity of Care or CCD format we mentioned earlier) and performs automatic redaction based on the privacy preferences in the consent document. **SafetyChecker** is a drug interaction checking program to demonstrate effective use of the redacted CCD.

A typical use case might involve a request by a physician for a patient's record (CCD). Using XDSProxy, the HIE will automatically return redacted

[6] http://sharps.org/wp-content/uploads/DS2-Policy.pdf.

CCDs based on patient privacy preferences. The requesting physician—knowing that the CCD may have been redacted—can still initiate a drug interaction check using SafetyChecker, since it is aware of the redaction issue. It responds (with a true or false result) to requests to compare a proposed medication list with all known interactions with medications in the *nonredacted* CCD without the physician seeing any redacted medications.

SimpleConsentFormDemo obtains patient consent at one of three levels:

- Level 1 relationships are known to be equivalent to or closely related to the target condition. For example, if "HIV" or any clinical fact known to indicate or treat HIV is present, then an HIV flag would return "True."
- Level 2 relationships include correlated concepts, such as comorbidities or co-occurrences, but only if certain conditions are met. For example, an HIV flag might be configured to return "True" when HIV comorbidities are present, but only if the record begins with a Level 1 concept.
- Level 3 relationships include concepts of significant complexity based on specific clinical rules. For example, if "HIV" is targeted, then a Level 3 flag might be set when two or more indirectly related concepts that likely suggest HIV are present. Alternately, it might be set if a comorbidity is present and is consistent with an HIV-related laboratory result such as a CD4 count within a particular range.

At present, SafetyChecker is limited to medications but it could potentially cover additional clinical domains. SimpleConsentFormDemo directives supported by the prototype are relatively simple and are limited to indicate a preference for disclosure or nondisclosure of specific conditions, such as HIV and mental health information.

As with privacy consent, there are several data segmentation models. A completely **patient-controlled model** is the norm for PHRs where there is no obvious alternative other than a friend or family member who may be more familiar with medical terminology than the patient. We earlier mentioned a **provider-assisted model** but that requires that busy providers spend time on the task and, even if they are willing, this may not work well if the patient's data is, as is often the case, spread among a number of providers. There are other approaches including **organization-controlled models, hybrid models and innovative tools** that are typically aimed at providing patients with more support. All of these ultimately try to simplify the complexity of the underlying clinical data which, in turn, relies on better data structures than are typically found in today's EHRs.

Public Key Infrastructure: *This is a technical subject but much of what is covered should be understandable to most readers.* Security and trust are interrelated, particularly with respect to the technology that is most commonly used to assure them: Public Key Infrastructure.

For our purposes, you need to master three key concepts: 1) **public and private keys**, 2) the difference between the **message** (sensitive information) and the digital **signature** derived from and attached to it and 3) the function of the two organizations involved in issuing the keys. The organizations are the **Registration Authority** (RA) which confirms the identity of the people or entities involved in the exchange of information. The functionally separate **Certificate Authority** (CA) actually signs and issues the digital certificates that are bound to that identity. One actual organization can serve both functions.

Public and private keys are used together to facilitate the secure transfer of information and for other purposes such as protecting licensed software or digital media. You can think of them simply as two numbers that are mathematically related, but in a complex way. One of them is public and freely available. The other is private and must, at all times, be secured. PKI rests on the core assumption that, given the freely available public key, it is prohibitively expensive and time consuming to calculate the matching private key. If you've ever done banking, paid bills or purchased anything on the Internet, you've used these keys.

Earlier, we referred to PKI in Direct where it can play two roles. First, as you know, the message attachment can contain PHI in any format the sending provider chooses. As a result, it must be protected in transit, so it is encrypted by the *sender* using the *receiver's* public key. *This is a critical and sometimes confusing point.* Anyone can "discover" their intended recipient's *public* key and can use it to secure information they wish to send to the owner of that key. Since only its owner should have access to the matching *private* key, only they can decrypt the data and view it. That's the essence of PKI for security. For healthcare providers subject to the HIPAA law, it is critically important to protect any private keys in their possession. If you read the technical description of Direct, you may recall that the HISP normally stores this key for its associated providers, in part to help secure it. If a private key is stored in computers located in a provider's office they could be stolen if so-called "malware" infected those computers.[7] If the keys must be stored on site, a security consultant can recommend proper procedures to protect against data theft and audit that they are being followed properly.

The second role that PKI can play in Direct is the creation of a "digital signature" via a calculation performed on the message before it is sent. This assures that the signature is unique for each message so it can be used for two

[7] http://www.symantec.com/connect/blogs/how-attackers-steal-private-keys-digital-certificates.

purposes—to validate that the message was unaltered in transit *and* that it actually came from the purported sender. *What follows can be confusing so pay close attention.* Remember the *attachment* was encrypted using the *recipient's public key*. The *signature* is encrypted using the *sender's private key*. Since anyone can obtain the sender's public key (after all, it is public), anyone can decrypt the signature. Remember, the digital signature is *not* the message; it is *derived from* the message. Once decrypted, it can be used to assure that the message wasn't altered in transit by again performing the calculation on the message and making sure it matches the result (the signature) that was derived from the original message before it was sent. Also, the fact that the signature can be decrypted using the sender's *public* key assures that it was encrypted by the sender using their *private* key, thus validating the source of the message. It is important to recognize that, nowhere in this scenario, was the sender's private key or the protected health data compromised.

Whether or not you followed the digital signature discussion, as a result of it, the recipient can trust that the message actually came unaltered from its sender. But how do the sender and recipient each know that the other is who they say they are? Could a hacker be masquerading as a doctor in order to get access to health data and use it to file fraudulent claims? Assuring trust to prevent this is the job of the RA. This isn't a particularly technology-driven process, rather it is driven by policies intended to assure that the background of people or entities is thoroughly verified before they are issued a pair of certificates. If you're a licensed health provider you are already familiar with such a process for licensure or credentialing.

Once an RA has verified identity, certificates can be issued. That's the job of the CA. They issue the certificates and typically encrypt public keys using *their private keys s*o people know/trust that certificates were, in fact, issued by them.

This section is definitely optional, but should be of interest to readers who wonder how PKI is implemented on web sites, including those used by providers and patients. As we said earlier, you've probably used PKI on the web. You can easily explore this yourself on any site whose URL begins with HTTPS. Depending on your browser, there is an icon (usually near the URL of the web site) that appears to indicate that the site is secure. Clicking on the icon should bring up the site name and a confirmation that the connection is encrypted. Again, depending on the browser, there is usually a button to obtain more detail, including the name of the CA that issued the site's certificate. As we just indicated in the technical discussion of PKI, the browser can verify the issuer of the certificate using the CA's public key, so browsers are normally shipped with those keys to use for this purpose. Remember,

once again, this is a *public* key so anyone can have it. Finally, clicking one or more other buttons brings up the site's public key, showing that it is really public.

If, for example, you try this with PayPal you should be able to verify that their CA is VeriSign; their certificate expires each year on April 2nd and, with a bit of effort, you can find PayPal's public key. *Note:* The public key is expressed in hexadecimal, a 16-digit numbering system that usually uses the symbols 0–9 to represent values zero to nine, and A,B,C,D,E,F (or alternatively a–f) to represent values 10 to 15. Its exact representation varies somewhat from browser to browser.

PKI is used for many more purposes than we've mentioned so far, including securing software until you've entered the vendor's code indicating you've paid for it, securing wireless networks and digital rights management to protect movies and music. In reality, PKI is more complex than described here. The validation of the certificate should be more robust, something called extended validation. To reduce the computational load, particularly on mobile devices, there are so-called hybrid approaches that create keys just for a specific session, but this is probably not sufficiently secure for healthcare. Finally, there is the concern about advances in computing. As we said, PKI rests on the assumption that, given the public key, it is too expensive and time consuming to calculate the private key. Proposed advances such as quantum computers might change this and could lead to the need to make major modifications for encryption to remain useful.

Summary: For digital health data to help coordinate care in our highly specialized and fragmented system, it must be shared. To support population health management, public health and research, it must be aggregated for analysis. When it is being stored, it must be protected from unauthorized or inappropriate access or use. Whenever it is moving, there are concerns about respecting the patient's privacy preferences, securing it from unauthorized disclosure and knowing that the persons or entities who sent it and with whom it is being shared are who they claim to be. We've now gained a basic understanding of the technologies and policies required to achieve all of these important objectives.

We've not yet dealt with a question that is increasingly being asked: have we gone too far with this? Are the complexities we've created in the interest of protecting individual privacy, security and trust now standing in the way of other important societal objectives? Don E. Detmer, MD, MA, Professor of Medical Education in the Department of Public Health Sciences at the University of Virginia, School of Medicine, is one of the most distinguished people in academic health informatics and was coeditor of the seminal 1997 IOM book *The Computer-Based Patient Record: An Essential Technology for Health Care.*[8] He has become

[8] http://www.nap.edu/catalog.php?record_id=5306.

concerned about this issue and has said that a "long and sometimes heated debate on the appropriate approach to the privacy of patients and confidentiality of their data has stalled the development of a framework that protects confidentiality while supporting the legitimate use of data for improving quality, research and public health. Unquestionably, privacy is a valuable and valued social good. But so too are altruism, health and freedom. Currently, health information policy seems to be giving too much weight to privacy at the expense of freedom and health."[9]

This is going to be a complex issue to resolve. It is neither black nor white and in the end, is about achieving the right balance between individual privacy and the needs of society. Better informatics tools such as DS2, could help by making patients more comfortable and amenable to sharing data as well as providing them with a practical means for specifying their sharing preferences.

Finally, what about the data itself? How is it represented, how is it packaged into documents and how are those documents actually transmitted? We now turn to that, one of the oldest topics in health informatics.

[9] http://issues.org/25-4/detmer/.

Data Standards

At the level of the health care organization, the lack of common data standards has prevented information sharing between commercial clinical laboratories and health care facilities, between pharmacies and health care providers regarding prescriptions, and between health care organizations and payers for reimbursement.
— Patient Safety: Achieving a New Standard for Care Institute of Medicine, 2004

Data and interoperability standards are the virtually ubiquitous plumbing that underlies all contemporary health informatics systems and tools. Given the complexity of healthcare it should not be surprising that this is a large topic. In this section we'll cover the part of it that should already be familiar to most healthcare providers—**data standards**. In the next section we'll discuss **interoperability standards** which deal with how health data is packaged into electronic documents, such as the CCD mentioned previously, and then how it is transported and shared. Interoperability standards is a more technical subject and one with which you may be less familiar than the data standards that providers use on a daily basis for purposes such as billing. However, it is a technology that will be increasingly important in clinical practice over the coming years.

Why do we need data standards? Figure 1 helps to make this clear by using a seemingly obvious data element: the patient's gender. It also helps explain the difference between syntax, which is structure and semantics, which is meaning. Here, in system (e.g. EHR) A, male is represented by a "1" and female by a "0." In system B, it is reversed. Should you mix data from these two systems without some intervening "standard", gender would be impossible to determine accurately for reporting and other purposes. We say that these two systems differ *syntactically*. They may be using the same language of 1s and 0s but they can't interoperate without some intermediate translation process that maps one to the other or both to some common syntax. In system C, "M" is used for male and "F" for female. We say that system C differs *semantically* from systems A and B—it uses a different "language" to represent gender. Moreover, system C recognizes that gender may be ambiguous and represents that with a U, a concept the other two systems don't deal with. Interoperability between system C and the other two systems would require translation from its syntax to theirs or translation of both semantic representations to a common form—a standard. Since patients with gender U can't be represented in systems A and B, some accommodation for that would also have to be made.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_7

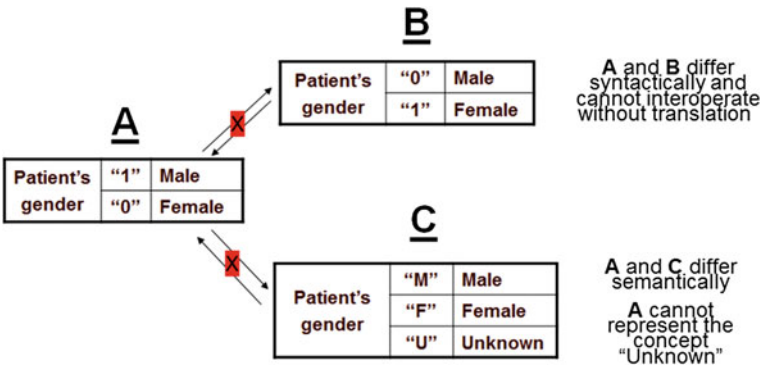


Figure 1: Semantic (meaning) and syntactic (structural) incompatibilities in the same or similar data stored in diverse systems create the need for standards. (Courtesy © Health Level Seven International)

If something as seemingly simple and obvious as gender can lead to this much complexity, it's not hard to understand what happens with concepts such as a patient's diagnosis that has many more possible values and is inherently more subjective.

Standards Evolution: Before we discuss the major data standards, it is important to understand that they have evolved over many years to encompass more aspects of medicine, to code them in more detail and to adapt as technology changes. To discuss each, we will divide this evolution into three dimensions: structure, purpose and technology.

Structure: Early data standards were lists such as medical diagnoses, laboratory tests or medications. We'll refer to such list standards as **classifications**. More recently, attention has been paid to coding more details about each identity and relationships among the entities. We'll refer to a standard that can code more detail including relationships as **ontologies**.

Purpose: Pre-computing, all of the early standards were for **data**. Physicians would use one of a number of classifications to describe patient problems. The clinical laboratory would have a classification for the tests they performed and the pharmacy for the medications they dispensed. As computers came into use in hospitals, the various departments needed to share information, but were using their own separate, specialized computer systems. This led to the early **messaging** standards. These meant a physician could order a lab test or medication at the nurses' station and that order could be routed electronically to the appropriate department to do the test or send the medication. Test results could also be returned electronically. The next evolution was standards for clinical **documents**. A message had typically been something like an order for a specific lab test or medication, but a complete summary of a patient's care might be required to manage the transition of care to the patient's physician upon hospital discharge. This is the role of document standards.

More recently, as computers have become more powerful, standards have been evolving to represent clinical **processes and workflows**. These have often been targeted at hospitals, and have proven hard to develop and implement. If this could be overcome, these standards could, at least in theory, lead to a reduction in unnecessary variations in patient care. This is yet another area where machine analytics may help. A new field called process mining is used widely in other industries to infer process and process variations from data that might be derived, for example in manufacturing, from sensors located throughout the factory floor. Early efforts are being made to do something similar based on EHR data. We'll see an example in the last section of the book but, as of now, this is research and is typically confined to one part of the hospital (often the emergency department). In time it would be of obvious value to be able to infer and visualize clinical processes across all departments, but this is much more difficult because they often use different systems that may represent information in different ways.

Technology: The final dimension of standards evolution is technology which is particularly applicable to messaging standards where the messages must be constructed from data standards in a manner that they can be understood by the systems that receive them. We just discussed why—departments in a hospital want to share information with each other so, for example, an order can flow from the physician caring for a patient to the lab and the result of that test can come back. Messaging standards serve that *purpose*, but now we're looking at the *technologies* through which those standards are implemented and formatted.

Early messaging standards were created using EDI/X12 – itself a standard that had evolved in other industries to automate business processes such as ordering, invoicing and payment. EDI/X12 was developed in the early days of computing when memory and storage were limited and expensive so it is quite intentionally compact and therefore cryptic. Here's an example:

OBX|1|SN|1554-5^GLUCOSE^POST 12H CFST:MCNC:PT:SER/PLAS: QN||^182|mg/dl|70_105|H|||F

With some effort, you should be able to tell that this line reports the results of a 12-hour post-prandial glucose test, but many of the details are not obvious. For example, what does "1554-5" mean? In fact, it is a special Logical Observation Identifiers Names and Code (LOINC) for that laboratory test,[1] but the message gives no indication of that. To understand each specific field typically requires a reference guide. More recently, messaging and document standards have been developed using XML, the more modern syntax common on the Internet that we've previously discussed as the format used in CCD electronic clinical summary documents. It is more verbose, but has the advantages of being more human-readable

[1] http://r.details.loinc.org/LOINC/1554-5.html?sections=Comprehensive.

and more easily rendered in a browser. This is the part of the lab test result we just looked at, but this time in XML:

```
<value xsi:type="PQ" value="182" unit="mg/dL"/>
<interpretationCode code="H"/>
<referenceRange>
<interpretationRange>
<value xsi:type="IVL_PQ">
<low value="70" unit="mg/dL"/>
<high value="105" unit="mg/dL"/>
</value>
<interpretationCode code="N"/>
</interpretationRange>
```

While it's still not easy to read, you should be able to tell that the value was 182 mg/dl (high), the normal range (70–105) is more obvious and it is clearer that this result is high (H).

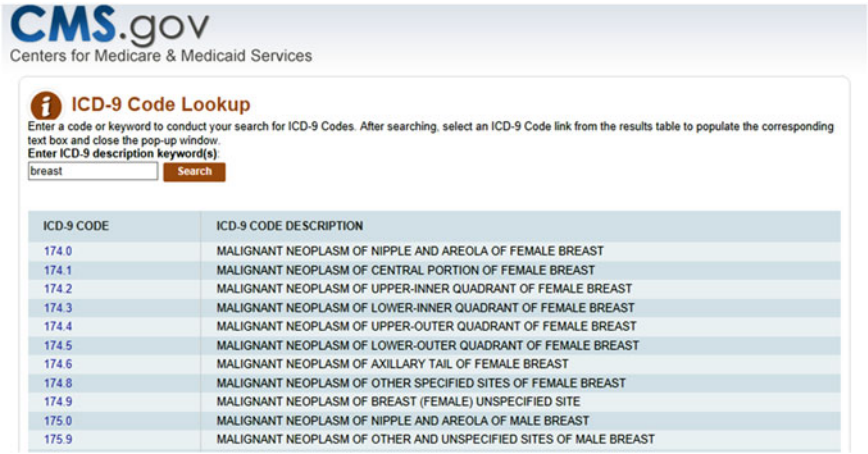
Data standards are the most widely used health standards and are a topic that is far too complex to cover in detail here. I'll provide an overview and you can look at the many resources, particularly on the Internet, for more details.

Key Data Standards: The five key data standards used in different combinations for different administrative, financial and clinical purposes are:

- International Classification of Diseases
- Current Procedural Terminology (CPT)
- Logical Observation Identifiers Names and Codes
- National Drug Code
- Systematized Nomenclature for Medicine

ICD and CPT are very widely used because they are required in most cases for medical billing. CPT and NDC are pretty much U.S.-specific. ICD-10, LOINC and SNOMED are internationally used ontologies capable of representing relationships among their concepts. NDC and CPT are classifications, although CPT increasingly incorporates subcodes to provide more details about each procedure. ICD is also becoming more complex as we'll discuss next.

International Classification of Disease: This is the oldest data standard, dating back to the 1800s and even to earlier centuries when researchers became interested in the causes of human mortality. Traditionally, it has been a list or classification of medical diagnoses maintained by the World Health Organization (WHO) and updated every 10 years. ICD-10, which was adopted in 1994, is the current version and is used by most countries, but we're still using ICD-9 in the U.S. The primary reason for the delay here is the substantial increase in complexity in ICD-10 so that it can represent many more clinical details. A major issue in standards design is how to achieve the right balance between complexity (capability to represent


CMS.gov
Centers for Medicare & Medicaid Services

ICD-9 Code Lookup

Enter a code or keyword to conduct your search for ICD-9 Codes. After searching, select an ICD-9 Code link from the results table to populate the corresponding text box and close the pop-up window.

Enter ICD-9 description keyword(s):

breast Search

ICD-9 CODE	ICD-9 CODE DESCRIPTION
174.0	MALIGNANT NEOPLASM OF NIPPLE AND AREOLA OF FEMALE BREAST
174.1	MALIGNANT NEOPLASM OF CENTRAL PORTION OF FEMALE BREAST
174.2	MALIGNANT NEOPLASM OF UPPER-INNER QUADRANT OF FEMALE BREAST
174.3	MALIGNANT NEOPLASM OF LOWER-INNER QUADRANT OF FEMALE BREAST
174.4	MALIGNANT NEOPLASM OF UPPER-OUTER QUADRANT OF FEMALE BREAST
174.5	MALIGNANT NEOPLASM OF LOWER-OUTER QUADRANT OF FEMALE BREAST
174.6	MALIGNANT NEOPLASM OF AXILLARY TAIL OF FEMALE BREAST
174.8	MALIGNANT NEOPLASM OF OTHER SPECIFIED SITES OF FEMALE BREAST
174.9	MALIGNANT NEOPLASM OF BREAST (FEMALE) UNSPECIFIED SITE
175.0	MALIGNANT NEOPLASM OF NIPPLE AND AREOLA OF MALE BREAST
175.9	MALIGNANT NEOPLASM OF OTHER AND UNSPECIFIED SITES OF MALE BREAST

Figure 2: The ICD-9 code set for breast cancer shows the site within the breast but not which breast is affected. (Source CMS)

medicine in all its details) and usability, the practical application of standards in the real world. The ICD-10 versus ICD-9 debate here in the U.S. is a clear example.

We can see the increased capability and complexity of ICD-10 by comparing the coding systems' capabilities for breast cancer. In ICD-9, as shown in Figure 2, we can know what portion of the breast is affected but, surprisingly, we can't know in which breast that is.

In contrast to ICD-9, laterality is represented in ICD-10—greatly increasing the number of codes—which is a big part of the objection to using it here in the U.S. In fact, ICD–10 really becomes an ontology capable of representing clinical relationships, as shown in Figure 3, where we can see that the patient has gout affecting the left shoulder but has not yet developed a complication (a uric acid deposit called a tophus) in that shoulder. The current U.S. target date for ICD-10 adoption is October 1, 2015, but given how close we are to ICD-11, there are proposals to skip ICD-10 entirely.

Current Procedural Terminology: This U.S. standard classification for medical procedures is maintained and updated annually by the American Medical Association and is required for virtually all reimbursement. CPT codes divide into three categories:

Category I codes represent widely performed procedures and are five-digit numbers divided into sections for anesthesiology, surgery, radiology, pathology and laboratory medicine, and medicine.
Category II codes are for the collection of quality and performance metrics and are four digits followed by an "F."
Category III codes are also four digits but are followed by an "I" and are temporary to allow for new or experimental procedures.

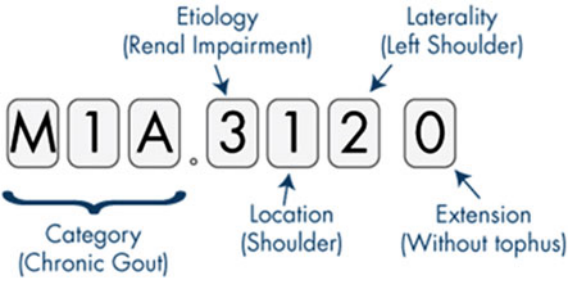


Figure 3: An ICD-10 code for chronic gout provides clinical detail about its etiology, location (and laterality) and the presence or absence of a tophus. (Courtesy AAPC)

For each code there are full, medium and short descriptions described at various levels of details for different purposes. Given its use in billing, a CPT code may provide details necessary to determine the proper charge. For example, there are codes a psychiatrist can use to indicate the length of a visit. The charge would, of course, typically be more for longer visits. Other codes are used to indicate how much dead, damaged or infected tissue has been removed to promote healing. Here again, removing more tissue could result in an increased charge.

CPT codes are simple, but given their critical role in billing (and subtleties such as these), selection of the right code is important and billing personnel are extensively trained to code correctly, usually to assure that the largest, but still legitimate bill is submitted.

Logical Observation Identifiers Names and Codes: LOINC was developed and is maintained by the Regenstrief Institute in Indiana, but is used virtually worldwide. Each code is for a laboratory test or a clinical observation and is a number of up to seven digits. While the codes themselves are deceptively simple, a great deal of information is detailed in their names which, as shown in Figure 4, are divided into five or six main parts separated by a colon.

Each of these parts can also potentially contain subfields separated by a carat (^). The LOINC code 1494-4 is a serum glucose determination 1.5 hours after the patient has consumed 100 grams of glucose. The code is simple, but here is its name:

Glucose^1.5H post 100 g glucose PO:MCnc:Pt:Ser/Plas:Qn

The first part is divided into two subparts (there can be three) separated by ^. We can clearly see it is a glucose measurement determined 1.5 hours after the patient has consumed 100 grams of glucose. The first subpart, "glucose," could also contain multiple levels of increasing specification separated by dots. The third and fourth parts—the time aspect (Pt means a point in time rather than a time range) and the system sample (Ser/Plas indicates that the test was performed on those components of a blood sample)—can be modified by a second subpart, again separated by a ^.

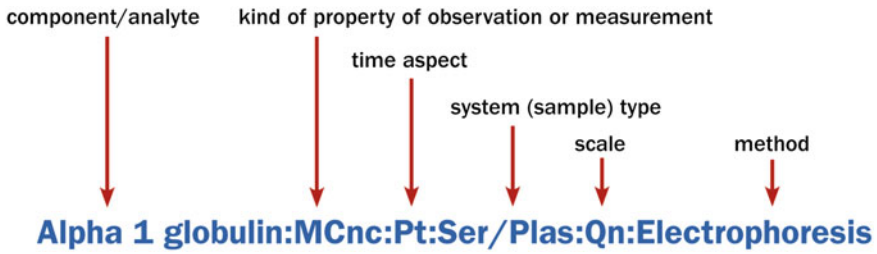


Figure 4: A LOINC name can have up to six parts, separated by colons. They can describe the component or analyte that was submitted for testing, the observation or measurement, whether the test is for a point in time or some interval, the sample that was submitted, the scale of the result and the testing method. In this example, Alpha 1 globulin (a serum protein constituent) is measured quantitatively (Qn) as mass concentration (MCnc) which is the ratio of mass to serum (Ser) volume, at a point in time (Pt) using electrophoresis. (Courtesy Regenstrief Institute)

Figure 5: The NDC specifies the labeler or vendor, the drug and the packaging. This set of information is useful to pharmacies for inventory control and can help identify a particular generic medication if, for example, a recall is indicated. (Courtesy Liberty Management Group Ltd., All Rights Reserved)

The National Drug Code: This is a U.S.-specific standard for medications and is maintained by the U.S. Food and Drug Administration (FDA). As shown in Figure 5, it consists of a simple 10-digit, three-segment structure to indicate the drug, the labeler/vendor and the packaging.

There are other commercial medication classification systems including First Databank, Micromedex, MediSpan, Gold Standard Drug Database and Multum. The National Library of Medicine also maintains RxNorm, a system of normalized drug names that is linked to these proprietary databases (which are often used in commercial pharmacy software to manage medication dispensing). In this role, RxNorm provides a relatively simple example of an interoperability tool to bring different semantic systems together since each of these commercial classifications effectively has its own "language."

SNOMED and SNOMED-CT (its subset for clinical medicine): The development of this most comprehensive of all existing health data standards began in the mid-1960s at the NIH as the vision of the late Dr. Arnold W. Pratt, a pathologist

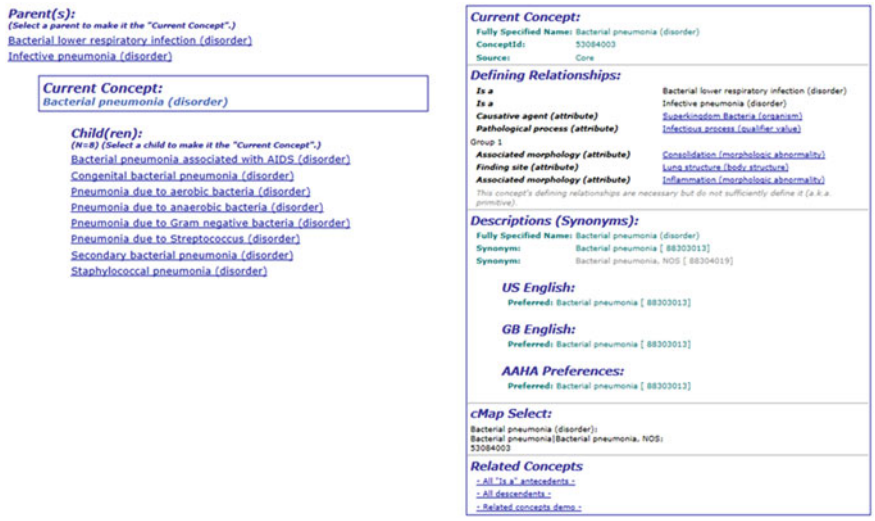


Figure 6: SNOMED-CT is a hierarchy that, in this example, represents the fact that bacterial pneumonia is a subset of infective pneumonia and bacterial lower respiratory infections. It also shows its Concept ID (53084003) and its relationships to the affected body structure (lung), infective agent (superkingdom bacteria) and clinical manifestations (consolidation and inflammation). (Courtesy the Veterinary Medical Informatics Laboratory at the Virginia-Maryland Regional College of Veterinary Medicine)

and the first head of NIH's Division of Computer Research and Technology. The original system was the Standardized Nomenclature of Pathology (SNOP) and the objective was computer encoding of pathologists' free-text dictated notes. It was always an ontology representing relationships among its concepts.[2] Today, it is multilingual, has expanded to all of medicine and is maintained by the International Health Terminology Standards Development Organisation (IHTSDO). SNOMED is huge and complex. Even its SNOMED-CT subset for clinical medicine has 311,000 concepts with 1.3 million relationships connecting them.

Concepts are a basic component of SNOMED-CT and have clinical meaning, a unique nine-digit numeric **ConceptID** and a unique human-readable **fully specified name (FSN).** Concepts can be represented as a hierarchy based on level of detail, as shown in Figure 6, which is from a publicly available SNOMED-CT browser that I recommend you visit.[3] In addition to showing the clinical hierarchy of Concepts on the left, it shows the ConceptID for the Current Concept–Bacterial Pneumonia, which is **53084003.** The box on the right, labeled Defining Relationships, illustrates two other basic components of SNOMED–**Relationship Links** expressed as "| is a |"

[2] http://dl.acm.org/citation.cfm?id=990414.

[3] http://vtsl.vetmed.vt.edu/.

and **Attribute Relationships** which can be a "finding site" or an "associated morphology."

We've barely touched the surface of SNOMED-CT. For example, there are 19 Top Level Concepts in the overall SNOMED-CT hierarchy, but we're only seeing three of them here–Organism (superkingdom bacteria), Clinical Finding (consolidation, inflammation) and Body Structure (lung). IHTSDO publishes a very readable starter guide to SNOMED-CT that I highly recommend if you have an interest in learning more.[4] What should be clear, however, is that a computer armed with a large clinical dataset coded in SNOMED-CT could perform very sophisticated searches and analyses using the hierarchy as a guide to medicine and biology. Direct coding of clinical notes into SNOMED-CT by providers isn't a realistic idea, but we'll see later on that computers are increasingly capable of automating that process.

Summary: We've now had an overview of how the structure, purpose and technology of standards have evolved over the years. You should be aware of the difference between simpler classifications such as NDC, CPT and ICD (at least prior to version 10) and a complex ontology, such as SNOMED. The standards community has been divided for decades over the choice between "perfection," a standard that can represent medicine in all its detail, and "practicality" standards that can actually be deployed and used in the real world. We'll look at this more specifically when we discuss FHIR, an attempt to create a new and more practical approach, and when we review EMR design challenges. We now know how health data is represented and we've had an introduction to the concept of packaging it into documents. Next, we'll explore how that is actually done.

[4] http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_StarterGuide_Current-en-US_INT_20140222.pdf?ok.

Interoperability Standards

... Truth is that there are lots of incentives for different economic actors in healthcare to make interoperability and standards more complicated. The reality is that sometimes these create barriers to entry for competitors. What happens though is at some point a tipping point occurs when suddenly it becomes in everyone's interest to make interoperability easier rather than harder.

— Barry Smith, PhD, HL7 Watch Blog, 2013.[1]

To be maximally useful in care coordination, standardized data—typically along with other nonstandardized data such as free text notes—is often packaged into standardized clinical documents and sent using standardized messaging formats. We briefly mentioned messaging standards earlier. Developed as hospital health IT systems, they came into widespread use in the late 1970 s and '80 s. Even back then, there were some early attempts to provide a single-vendor, fully integrated hospital-wide solution, but most systems were developed for a particular department—such as pharmacy, clinical lab or radiology—and it was common for each of these departments to independently select the system that they felt best fit their needs. As a result, hospitals were faced with the problem of communicating among these systems so, for example, patients could be uniformly registered in them and charges could be reliably collected from them. There were a number of early technologies employed to meet this need[2] that led to the founding of Health Level Seven in 1987 as a "not-for-profit, accredited standards developing organization dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services."[3]

The current standard, HL7 Version 3, was developed over many years and was first published in 2005. Prior to it, HL7 used EDI/X12 to format the messages. Earlier, we saw part of the following example of an HL7 v2 EDI/X12 formatted

[1] http://hl7-watch.blogspot.com/2013/05/fhir-lets-make-things-difficult-again.html.
[2] http://www.ringholm.com/docs/the_early_history_of_health_level_7_HL7.htm.
[3] http://www.hl7.org/.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_8

message—a lab test result as it might be sent back to the ordering physician. Note that here I've highlighted the three-letter name of each of the complete message's four lines (called segments). Each HL7 segment consists of one or more composites (also known as fields). A pipe (|) character is used to separate one composite from another. If a composite contains other composites, these subcomposites (or subfields) are normally separated by ^. All of this is illustrated in this example:

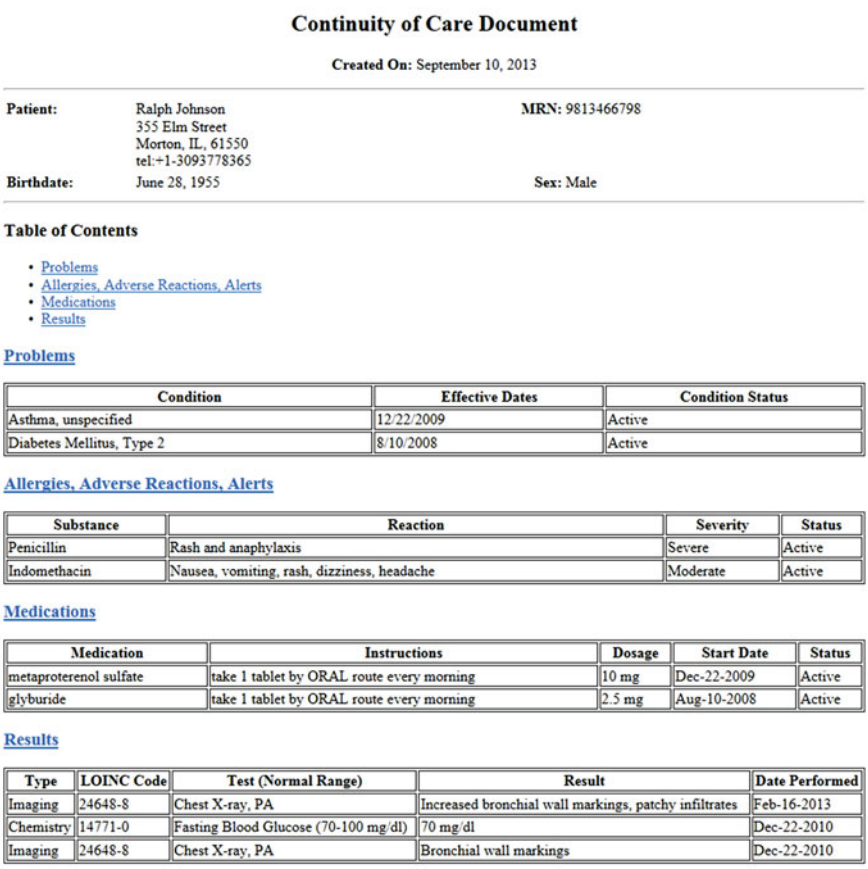
MSH|^~\&|GHH LAB|ELAB-3|GHH OE|BLDG4|200202150930||ORU^R01|CNTRL-3456|P|2.4
PID|||555-44-4444||EVERYWOMAN^EVE^E^^^^L|JONES|19620320|F|||153 FERNWOOD DR.^ ^STATESVILLE^OH^35292||(206)3345232|(206)752-121||||AC555444444||67-A4335^OH^20030520
OBR|1|845439^GHH OE|1045813^GHH LAB|15545^GLUCOSE|||200202150730|||||||| 555-55-5555^PRIMARY^PATRICIA P^^^^MD^^|||||||||F|||||444-44-4444^HIPPOCRATES^HOWARD H^^^^MD
OBX|1|SN|1554-5^GLUCOSE^POST 12H CFST:MCNC:PT:SER/PLAS:QN||^182|mg/dl|70_105|H|||F

We looked earlier only at the OBX segment, the most interesting part clinically since it contains the test results. In that discussion, we mentioned the difficulty in knowing what 1554-5 represents in the OBX line and how that illustrates a limitation of EDI/X12. It was developed when computer memory was limited and expensive so it's quite compact and therefore cryptic. It was designed to be read by computers, but not humans.

Here's a part of the same message in XML (the technology introduced into HL7 in Version 3) that we didn't look at previously:

```
<code code="1554-5" codeSystemName="LN"
codeSystem="2.16.840.1.113883.6.1"
displayName="GLUCOSE^POST 12H CFST:MCNC:PT:SER/PLAS:QN"/>
<statusCode code="completed"/>
<effectiveTime value="200202150730"/>
```

While it's not as obvious as something designed just for human reading, it is much clearer in the XML version that 1554-5 is a LOINC code (abbreviated as LN) for the lab test. 2.16.840.1.113883.6.1 is the OID (Object Identifier, a hierarchical international system for uniquely identifying objects) for LOINC that is included so that a computer can know what code this is by looking it up in an OID database.



Continuity of Care Document

Created On: September 10, 2013

Patient: Ralph Johnson
355 Elm Street
Morton, IL, 61550
tel:+1-3093778365

MRN: 9813466798

Birthdate: June 28, 1955

Sex: Male

Table of Contents

- Problems
- Allergies, Adverse Reactions, Alerts
- Medications
- Results

Problems

Condition	Effective Dates	Condition Status
Asthma, unspecified	12/22/2009	Active
Diabetes Mellitus, Type 2	8/10/2008	Active

Allergies, Adverse Reactions, Alerts

Substance	Reaction	Severity	Status
Penicillin	Rash and anaphylaxis	Severe	Active
Indomethacin	Nausea, vomiting, rash, dizziness, headache	Moderate	Active

Medications

Medication	Instructions	Dosage	Start Date	Status
metaproterenol sulfate	take 1 tablet by ORAL route every morning	10 mg	Dec-22-2009	Active
glyburide	take 1 tablet by ORAL route every morning	2.5 mg	Aug-10-2008	Active

Results

Type	LOINC Code	Test (Normal Range)	Result	Date Performed
Imaging	24648-8	Chest X-ray, PA	Increased bronchial wall markings, patchy infiltrates	Feb-16-2013
Chemistry	14771-0	Fasting Blood Glucose (70-100 mg/dl)	70 mg/dl	Dec-22-2010
Imaging	24648-8	Chest X-ray, PA	Bronchial wall markings	Dec-22-2010

Figure 1: Data from a CCD is presented in a more useful, human-readable form. This illustrates the data that can be contained within a CCD and also that XML can be visualized in a browser. (Courtesy ABEL Medical Software, All Rights Reserved)

Before moving on to document standards, it is important to understand that, over many years, HL7 has expanded greatly from its original messaging mission. Some now argue that it has become overly complex and virtually impossible to fully implement. This idea is what led to FHIR, which we'll discuss at the end of this section.

There are many other components of HL7 we won't discuss here. For example, there is now even an HL7 standard for developing HL7 standards called the **HL7 Development Framework** (HDF). For our purposes, it should be sufficient to know it exists and that HDF seeks to promote a collaborative standards development effort grounded in real-world problems.

At the beginning of this section, we discussed syntax and semantics and said that syntax is structure while semantics is meaning. The HL7 **Reference Implementation Model** (RIM) defines the semantics of a common set of administrative,

financial and clinical concepts in order to foster interoperability. It consists of five abstract concepts: Every happening is an **act**. Examples include procedures, observations and medications. Acts are connected through an **act relationship** such as composition, preconditions, revisions and support. **Participation** defines the context for an act including its author, performer, subject and location. The participants have **roles** such as patient, provider, practitioner, specimen or healthcare facility. Roles are played by **entities** such as persons, organizations, material, places or devices. Before leaving RIM and moving on to some of the things it is used for, it is important to emphasize again that the goal is interoperability and, in support of that, RIM can be used both for HL7 v3 messages and for clinical documents constructed, at least in part, from the data in them. This was a key goal of the effort. FHIR, while substantially simpler, is also based on elements taken from RIM.

The **Clinical Document Architecture** defines HL7 v3 RIM-based documents that are assembled from elements, including administrative and clinical data, for particular purposes. After a period of use, a consolidation effort removed inconsistencies that crept into the initial documents, resulting in the current version called **Consolidated CDA** (CCDA). The CCD is the key CCDA document that we're mentioned several times previously, primarily at transitions of care when, for example, patients are admitted to or discharged from a hospital or referred by a PCP to a specialist. Figure 1 is a human-readable rendering of a CCD using a browser. Since browsers "understand" XML, this can normally be done by simply opening a CCD file using one.

CCDA documents are assembled from templates, essentially reusable XML components. Templates are constructed at the document, section and data entry levels. These correspond conceptually to the parts of a paper form (such as those commonly used by physicians) in which the document consists of sections, each of which consists of fields where specific data items can be recorded. Here's the CCDA section template for objective clinical data such as the physical exam. Note, in particular, that the actual recorded clinical data can be free text, as it is here:

```
<section>
 <templateId root="2.16.840.1.113883.10.20.21.2.1"/>
 <code code="61149-1" codeSystem="2.16.840.1.113883.6.1"
 codeSystemName="LOINC"
 displayName="OBJECTIVE DATA"/>
 <title>OBJECTIVE DATA</title>
 <text>
 <list listType="ordered">
 <item>Chest: clear to ausc. No rales, normal breath sounds</item>
 <item>Heart: RR, PMI in normal location and no heave or evidence of
 cardiomegaly,normal heart sounds, no murm or gallop</item>
 </list>
 </text>
</section>
```

You should not be surprised that 61149-1 is the LOINC code for Objective Narrative so, even though a computer may not be able to do much with the free text, it will know what it is and could intelligently add it to the proper section of a chart. As we saw earlier, following the code is the OID for LOINC:

codeSystem = "2.16.840.1.113883.6.1"

So, equipped with these two numbers, a computer knows what "dictionary" to look up the code in (LOINC) and what code to look for (61149-1).

Here is a template for a specific item of clinical data, the patient's age:

```
<observation classCode="OBS" moodCode="EVN">
<templateId root="2.16.840.1.113883.10.20.22.4.31"/>
<!-- Age observation template -->
<code code="397659008" codeSystem="2.16.840.1.113883.6.96"
displayName="Age"/>
<statusCode code="completed"/>
<value xsi:type="PQ" value="57" unit="a"/>
</observation>
```

Here the code is 3976590080, a nine-digit number which we would expect from SNOMED-CT. In fact, the OID that follows ("2.16.840.1.113883.6.96") is for SNOMED.

These two examples should give you a feel for what's going on behind the scenes in a CCDA document and how it helps a computer make sense of the information contained in the document and do intelligent things with it such as filing it in the proper place in an EHR or doing the automated medication and clinical data reconciliations we saw earlier.

In the rest of this section, we'll discuss some advanced standards efforts that don't fit neatly into the data or interoperability categories. In all cases my objective is to familiarize you with them and their interesting objectives which illustrate some of the future directions of health informatics.

Standards for Clinical Decision Support: The goal of the **Arden Syntax** is describing medical logic so that it can be shared across EMRs to support CDS. CDS is the idea that computers can suggest optimal therapy based on a patient's electronic record or spot potential mistakes before they happen. To do either, they not only need the electronic record, they need to "understand" at least a limited subdomain of medical knowledge. Arden consists of Medical Logic Modules (MLMs), each of which supports one clinical decision.

Here, in Figure 2, a group of five MLMs are used to provide CDS for the use of warfarin. For nonclinicians reading this book, warfarin is a common but potentially dangerous blood thinner often given to patients who are at risk of thrombotic (clot-based) strokes or other potential blood clotting problems such as can occur after an

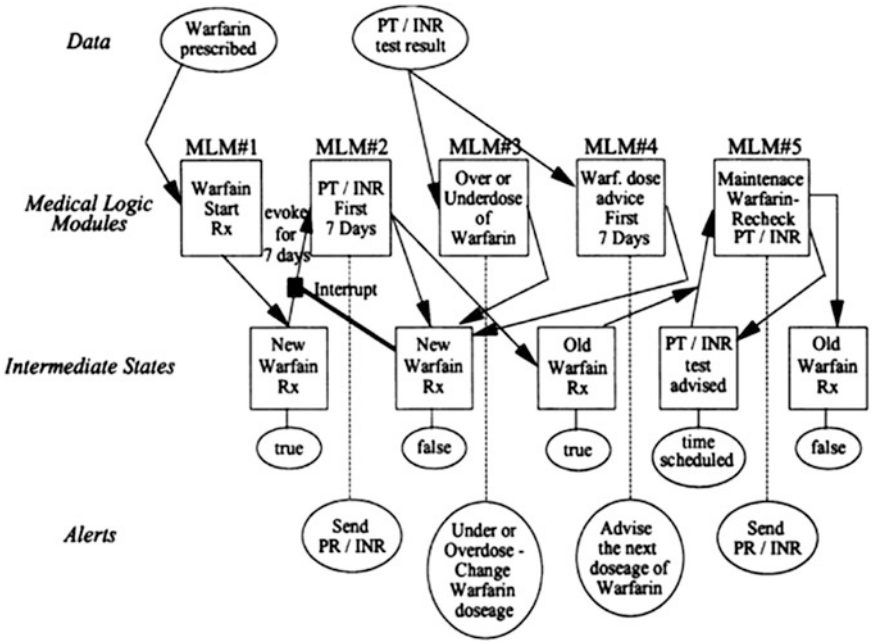


Figure 2: A group of five Arden Syntax MLMs are sufficient to provide CDS for the proper use and dosing of warfarin. (Courtesy AMIA). (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2579091/)

artificial heart valve is installed. Too little of the drug may mean another stroke; too much can cause excessive, even fatal bleeding. Warfarin control can be affected by many other drugs and even the patient's diet, so it is tricky to manage. A clinical test called a PT/INR is used to assess the degree of blood thinning.

Together, these five modules—once they are fed clinical data about the patient such as their PT/INR—can make warfarin dosage recommendations to the physician.

Arden was developed, with support from IBM, at Columbia University with the idea that centers of excellence could develop MLMs and share them with other institutions. However, there are a number of technical, essentially interoperability, challenges to actually deploying Arden. The so-called curly braces problem is particularly illustrative. Here's the applicable part of one Arden MLM:

```
creatinine := read {'dam'="PDQRES2"};
last_creat := read last {select "OBSRV_VALUE"
from "LCR" where qualifier in
("CREATININE",
"QUERY_OBSRV_ALL")};
```

So, to check for the patient's kidney function, Arden needs to know the creatinine level and, to get it, **{'dam' = "PDQRES2"};** must be interpreted by the EHR in a particular hospital to fetch that value from its proprietary database. So, even though the purpose of Arden is to share CDS, because EHRs aren't readily interoperable, we run into roadblocks like the curly braces problem which must be resolved locally by each institution. This drives up the cost and complexity of implementing Arden and that has, in part, led to low Arden adoption.

Health eDecisions was a nearly two-year-long standards effort sponsored by ONC that is now folded into a broader Clinical Quality Framework Initiative (CQF).[4] It proposes using parts of the Arden technology along with web services designed to make fetching the needed data much easier. The JASON proposal, if adopted and implemented, could potentially provide a universal solution to this problem where the needed clinical data is included in the proposed simplified standard data model. For example, assuming that FHIR is ultimately adopted to implement JASON, each EHR would convert its own data representations into the corresponding FHIR equivalents (called FHIR Resources which are actually groups of related data elements). Thus, an Arden rule could be universally written using the FHIR syntax, obviating the need for interpretation at each receiving location.

The **Clinical Information Modeling Initiative** (CIMI) is an attempt to enumerate the detailed models of hundreds to thousands of medical ideas to achieve consensus among clinicians. It could be used to define standard messages or structured documents, as components of clinical rules, and to automate or facilitate constructing data entry or reporting templates. The CIMI group lists simplification of existing models (perhaps referring to RIM) as one of its key goals.

The **Context Management Specification (CMS)** began at Duke University as the Clinical Context Object Workgroup. It should be of interest to any physician reader since it seeks to facilitate the integration of diverse applications at the point of use. It proposes to do this by managing issues such as:

- The identity of a patient whose data the user wants to view or update via the applications
- The identity of the user who wants to access the applications
- A moment in time around which temporal data displays should be centered by the applications
- A particular patient encounter that the user wants to review via the applications

The details are highly technical but, if it were implemented, applications from different vendors would all automatically be centered for the clinician around the same patient once the patient is identified in any Context Management Specification-compliant application. The clinician could move about among them as though they were a single system. This is an ambitious, but potentially very important, effort since hospital-based clinicians commonly complain about the time and effort involved to use many different systems that work in different ways to manage their

[4] http://wiki.siframework.org/Clinical+Quality+Framework+Initiative.

patients. Duke and Caradigm, a health analytics company, is partnering to offer CMS commercially. You can learn more about this interesting technology by watching a Youtube video of a Duke webinar sponsored by Caradigm.[5]

Fast Health Interoperability Resources: We've repeatedly mentioned FHIR as a potentially important new direction in standards and as a potential means of actually implementing the JASON report's call for a national interoperability framework. I cannot overemphasize the possible significance of FHIR in clinical practice and other applications of health informatics. I've alluded to that earlier when I suggested, for example, that physicians might be able to choose their own informatics tools for recording and reviewing patient data and use them across multiple hospitals and clinics irrespective of the underlying EHR in each organization. Interoperability has been the grand challenge preventing progress on many fronts in the field and FHIR may well be the long-sought solution although, as of this writing, it is too early to know for certain if this will happen.

FHIR began with an arguably seminal 2011 blog post, *The Rise and Fall of HL7*. In it, Eliot Muir, founder and CEO of Interfaceware, Inc., a major Toronto-based HL7 solutions provider, said: "Complicated standards can be pushed for a while but ultimately markets reject them."[6] He then went on to essentially say that RIM is doomed because it is too complex, no single model will work in all domains, it's internally inconsistent (which is often true of extremely complex constructs including hospital-wide enterprise information systems even from a single vendor) and ultimately, it's too expensive to implement. He then recommended a new approach based on web services (the way you already interact with web sites to look up and order products, for example). Australian standards guru, Grahame Grieve, made this idea concrete and added a simplified data model (based on HL7 RIM) to facilitate easier and more rapid adoption and implementation and named it Fast Healthcare Interoperability Resources.[7] In addition to a simplified data model (Resources) and a web service approach to accessing the data in them, FHIR also deals with relationships among the data items, again in a simpler way than is implemented by RIM.

We discussed that a similar web services approach was specified by the proposed Health eDecisions standard for CDS.[8] Through this approach "centers of excellence" could make up-to-date, evidence-based tools available so that any EHR or other clinical system could query them for advice. Such an approach has an important benefit in that only one system needs to be maintained as medical evidence changes and physicians using that system know what trusted entity stands behind the advice.

To make it clear what FHIR might facilitate in the relatively near future, Figure 3 provides a working prototype provided to me by Cerner and Polygot Systems. It shows that the Meducation SMART on FHIR application is easily launched from

[5] https://www.youtube.com/watch?v=asUUCzNG71g

[6] http://www.interfaceware.com/blog/the-rise-and-fall-of-hl7/.

[7] http://hl7.org/implement/standards/FHIR-Develop/.

[8] http://wiki.siframework.org/Health+eDecisions+Homepage.

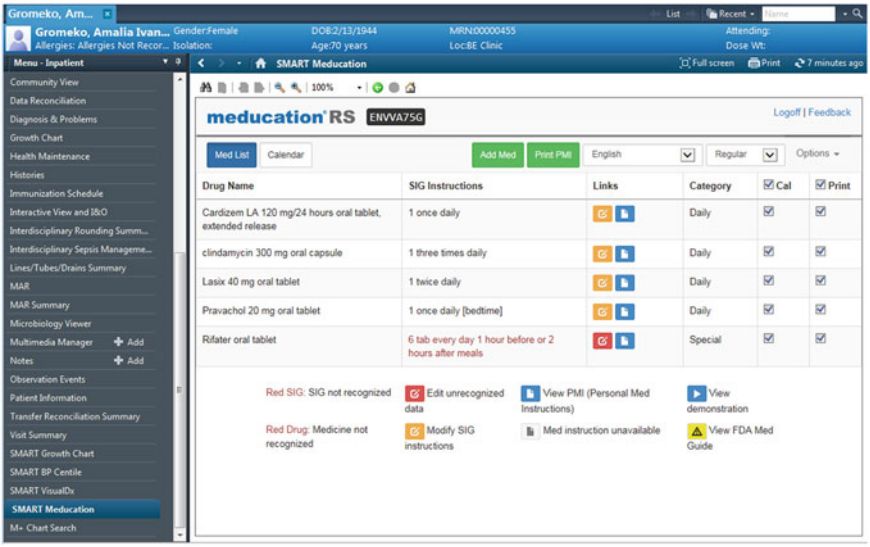


Figure 3: SMART on FHIR apps are directly accessed by physicians during charting using the same menu they might use, for example, to add notes. This is a large step forward in workflow that saves physician time, eliminates duplicate data entry and should lead to increased use of the apps. (Courtesy Polygot Systems, All Rights Reserved)

the same PowerChart Menu (it is highlighted at the bottom left of the screen) used for other charting purposes. You can verify this by reading some of the menu choices above it in the list. Meducation displays the current patient's medication list, obtained via a call to the Cerner Millennium FHIR service. This service maps clinical data from Cerner's representation to the FHIR data model and makes it available by the web services approach we've briefly discussed. Importantly, the physician does not have to do any duplicate data entry or re-identify the patient. They simply click the SMART Meducation entry in the menu on the left part of the screen they are already using for charting. Look carefully and you'll also see SMART Growth Chart listed near the bottom of the menu along with some other SMART on FHIR apps.

This highlights at least two important benefits: It saves time by incorporating new capabilities directly into the physician's workflow, thus encouraging use of new tools, such as CDS. It also avoids any potential errors that might be caused by duplicate entry of existing data. As a result of these potential benefits FHIR adoption is growing quite rapidly, according to Grahame Grieve. Organizations can post their interest on the FHIR site.[9]

Before we discuss FHIR in more detail, you may be wondering what a web service is. To illustrate this as clearly as possible in nontechnical healthcare terms, we'll use

[9] http://wiki.hl7.org/index.php?title=Organizations_interested_in_FHIR.

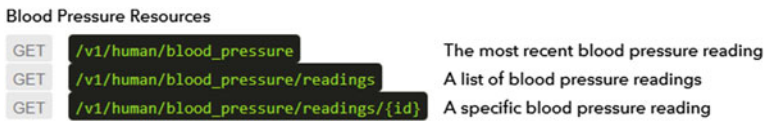


Figure 4: Human API's defined GET statements are different queries for blood pressure. Note that they are essentially human-readable. Implementation is relatively easy and fast facilitating use of the data. (Courtesy Human API, All Rights Reserved)

an interesting service provided by a company called Human API.[10] As you know, a patient's provider-recorded medical data can be stored in more than one place, particularly if the patient has multiple chronic diseases being cared for by several providers. However, there are also many other places patient data may be stored and from which it can be accessed. For example, lab test results may be directly available from the company that did the tests. Increasingly patients may be recording data using mobile, wireless or wearable devices or by using an app on their smartphones. It would obviously be desirable for providing well-integrated and coordinated care, to pull all of this data together to create a clear, quantitative picture of an individual's health. Human API does this aggregation and makes the aggregated data available via simple web services. In addition to technical issues in meshing data from disparate sources, obtaining permission to access data can be an obstacle. As we've discussed, many have proposed that an answer might be to let patients control their own data and make sharing decisions directly. That's the basic idea behind Human API.

The company aggregates data from a growing list of wearable and other devices, other web sites and health-related apps. The current data sources are listed on the Human API site which also solicits suggestions for additional data sources from its visitors.[11] They then normalize the data, a process that is conceptually similar to mapping it into a standard data model such as the one being developed by FHIR. However, their service precedes FHIR and is patient-, rather than provider-facing and currently uses its own model. Patients can then give permission for their physician or any other interested person or entity to access all or part of their data via a Representational State Transfer (REST) API. Figure 4 shows a part of how that actually works technically. Data is retrieved using a GET statement which does what you would expect—it fetches data from a remote database. You should also pretty easily be able to see what each of these three alternate GET statements for retrieving blood pressure readings do.

The following highlighted discussion of FHIR is somewhat technical in places, so skip ahead if you wish As we've said, the FHIR specification includes a simplified, standard data model, serialization formats in XML and JSON (JavaScript Object Notation) and a REST API for querying clinical data. The

[10] https://humanapi.co/.

[11] https://humanapi.co/sources.

data model is comprised of a set of modular components called "Resources." The FHIR resources are a subset of RIM: essentially the most important, commonly used data constructs that are clearly organized and categorized on a page at the FHIR web site. I suggest you visit this site if you are more technically inclined.[12] These resources can easily be assembled into working systems to solve real-world clinical and administrative problems. FHIR is designed to be suitable for use in a wide variety of contexts including mobile phone apps, cloud communications, EHR-based data sharing, server communication in large institutional healthcare organizations and other environments.

FHIR uses resource tags in the underlying data structures. Here's an example, the FHIR Resource for a patient medication allergy or intolerance:

```
<AllergyIntolerance xmlns="http://hl7.org/fhir"> doco
<!-- from Resource: extension, modifierExtension, language, text, and contained -->
 <identifier><!-- 0..* Identifier External Ids for this item --></identifier>
<criticality value="[code]"/><!-- 0..1 fatal | high | medium | low -->
<sensitivityType value="[code]"/><!-- 1..1 allergy | intolerance | unknown -->
<recordedDate value="[dateTime]"/><!-- 0..1 When recorded -->
<status value="[code]"/><!-- 1..1 suspected | confirmed | refuted | resolved -->
<subject><!-- 1..1 Resource(Patient) Who the sensitivity is for --></subject>
<recorder><!-- 0..1 Resource(Practitioner|Patient) Who recorded the sensitivity --
></recorder>
<substance><!-- 1..1 Resource(Substance) The substance that causes the
sensitivity --></substance>
<reaction><!-- 0..* Resource(AdverseReaction) Reactions associated with the
sensitivity --></reaction>
<sensitivityTest><!-- 0..* Resource(Observation)
        Observations that confirm or refute --></sensitivityTest>
</AllergyIntolerance>
```

The highlighted values should be clear to most clinicians as the attributes that one might want to document for an allergy or intolerance along with their potential values if these can be predicted. This FHIR Resource derives from the corresponding HL7 RIM Data Model which only technically inclined readers may want to review.[13]

[12] http://www.hl7.org/implement/standards/fhir/resourcelist.html.

[13] http://wiki.hl7.org/images/1/1b/Allergy_and_Intolerance_INFORM_2013_MAY.pdf.

As we've said, in FHIR the interaction between the systems and the database is via REST APIs, one of several commonly used technologies for implementing web services (SOAP and JSON are others). Any web service requires Internet-connected server(s) that are always available and will respond to requests if they are formatted in a manner specified by the technology being used. REST and SOAP format requests in XML, while JSON uses a subset of JavaScript. REST is popular with programmers because it is relatively easy and fast to implement and is more compact, facilitating rapid innovation and development. Generally there are four REST commands: GET, PUT, POST or DELETE. As we just illustrated using Human API, GET retrieves a resource (which can be virtually anything), PUT replaces an existing resource with a new or updated version (whose location must be specified), POST adds a new resource (no location need be specified) and DELETE removes a resource.

In addition to creating a universal app platform, another potential ramification of FHIR is the replacement of the CCDA with a simpler alternative for exchanging a set of clinical data for a specific use case. In December 2014, Dr. John Halamka, a well-known leader in health information technology, announced the Argonaut Project. It essentially proposes to replace CCDA through an expansion of FHIR. FHIR has already taken advantage of JSON objects (they can also be in XML) as a means of packaging what are essentially sections of documents. Why not go all the way and use sets of FHIR Resources to exchange the same data that is contained in far more complex CCDA documents?[14]

The initial Argonaut Project Use Case is the already familiar transitions of care —when a patient is referred from one institution or provider to another, such as a hospital discharge to the care of a physician or to a subacute or long-term care facility. At present, the CCDA-specified, XML-formatted CCD is the standard specified under Meaningful Use in such situations. It is also widely used to meet the VDT requirement to give patients access to their medical records.

The CCD is actually an extremely complex document. Moreover, it is rather loosely specified, so CCDs can vary substantially based on the EHR that produced them. Argonaut proposes to extend the FHIR Resources to encompass the entire Meaningful Use Common Dataset.[15] This would be done within the FHIR framework which seeks to prevent individual Resources from becoming overly complicated, so any needed new data would be spread out to appropriate existing Resources and some new Resources might be defined. FHIR is an emerging standard that encourages extensions, so this is also is in keeping with the spirit of the standard.

[14] http://geekdoctor.blogspot.com/2014/12/the-argonaut-project-charter.html.

[15] http://ccda.sitenv.org/Common+MU+Data+Set.

Argonaut also proposes to also develop a new FHIR REST GET API that could be used by organizations to request the information contained within the CCD, but in the form of a set of FHIR Resources that, in total, provide the same information that is in the current CCD specification. Currently Direct might be used to make this request but Dr. Halamka also believes that in time, Direct will prove to have been a transitional technology that would in most, if not all cases be replaced by the FHIR APIs as a simple, inexpensive means of data sharing. In this context, FHIR has the clear advantage of supporting both push (e.g. the referring physician sending a patient summary to the specialist) and pull (e.g. a patient subscribing to automatic updates to their PHR whenever they visit the physician).

Further evidence that FHIR may become the national interoperability standard was provided by the list of project collaborators that includes major enterprise software vendors Cerner, Epic, McKesson and Meditech along with major health systems including Intermountain Health, the Mayo Clinic and Partners Healthcare System.[16] In my blog post on Argonaut, Dr. Halamka explains that "the vendor community has struggled with Meaningful Use, and Stage 2 in particular. There is real concern about what Stage 3 might be like and whether there might be a government interoperability mandate in it as proposed by JASON. I find the vendor community in favor of an approach in which industry collaborates with other stakeholders. It is entirely possible that long before 2017, when Stage 3 is due, the matter of interoperability will be settled. At least I hope that's the case."[17]

A Brief Review: We've now completed our review of data and interoperability standards. This is an immense area and we've only skimmed the surface in an effort to equip you with a working knowledge of the areas being standardized, the approaches and technologies being used and the potential benefits. You should recognize that, for the most part, health informatics standards have been developed within the industry and have only recently begun to adopt approaches, such as XML and web services, that were invented elsewhere, are very widely used and, as a result, are virtually free for deployment and use in healthcare. This is an important issue for HIE which, as you also know, has historic problems developing a sustainable business model. Dramatically reducing the cost would certainly go a long way toward resolving that issue.

You should also appreciate the inevitable tension that exists between the need many feel to create very detailed standards for everything and the resulting complexity that creates. Often this complexity gets in the way of the primary objective, which is implementation and use of the standards in day-to-day patient care. The increasing recognition of that problem seems to be leading to a new willingness to look outside of healthcare and to adopt more facile and easy-to-implement approaches, such as FHIR.

---

[16] http://mycourses.med.harvard.edu/ec_res/nt/6209858F-CDDD-4518-ADF8-F94DF98B5ECF/Argonaut_Project-12_Dec_2014-v2.pdf.

[17] http://www.informationweek.com/strategic-cio/can-argonaut-project-make-exchanging-health-data-easier/a/d-id/1318774?

We've also now completed our discussion of the core technologies of health IT, so we're ready to see how it is being used in the real world. We'll begin with what is still a very challenging endeavor, even after decades of effort, and which is also an issue that I'm sure many readers have been waiting for us to discuss—the development of practical, efficient and easy-to-use electronic medical record systems.

EHR Design and Usability Challenges

In the 1950s the clinical data in medical records of patients in the United States were mostly recorded in a natural, English-language, textual form. ... Such patients' data were generally recorded by health-care professionals as hand-written notes, or as dictated reports that were then transcribed and typed on paper sheets, that were all collated in paper-based charts; and these patients' medical charts were then stored on shelves in the medical record room. The process of manually retrieving data from patients' paper-based medical charts was always cumbersome and time consuming. An additional frequent problem was when a patient was seeing more than one physician on the same day in the same medical facility; then that patient's paper-based chart was often left in the first doctor's office, and therefore was not available to the other physicians who then had to see the patient without having access to any recorded prior patient's information.

— Morris F. Collen, MD, Computer Medical Databases 2012[1]

This section provides what I feel are key insights into the current issues in using electronic health records as well as common limitations of those systems. In part, the intent is to inform readers considering the selection of an EHR about factors that I feel are commonly ignored or not given sufficient weight in the selection process by the typical practice.

Physician Attitudes about EHRs: Despite the significant EHR design and usability challenges we'll discuss in this section, most physician adopters agree that there are substantial benefits to introducing digital records into clinical practice. Recall, for example, that in the 2013 Rand/American Medical Association physician survey we discussed earlier, only 18 percent of EHR adopters would prefer to return to paper. However, physicians in that same survey complained about "poor EHR usability, time-consuming data entry, interference with face-to-face patient care, degraded clinical documentation (as a consequence of template-based notes) and inefficient and less fulfilling work content."[2] As we'll see, depending on the EHR physicians are using and, more particularly, its approach to data collection, these can certainly be valid complaints.

[1] http://link.springer.com/chapter/10.1007/978-0-85729-962-8_3#page-1.
[2] http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR439/RAND_RR439.pdf.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_9

An interesting and more recent, but more limited survey (600 physicians, though the sponsors claim a confidence level of 95 percent with a confidence interval of 4 percent) was conducted by Software Advice, a company that provides research and reviews of software applications, and Research Now, a market research company. It found that:

- Mobile EHR users reported higher levels of satisfaction and fewer challenges with their EHR than nonmobile users.
- Investing more in patient portals (web tools to communicate with patients) was a top priority, partially due to the need to improve patient engagement.
- Over half of users reported having difficulties integrating data from external systems with their EHR system.[3]

The survey also found that smaller practices (which are more likely to make an independent purchase decision than larger practices, which are more likely be aligned with or owned by a health system) were somewhat more likely to be satisfied, but in line with the 2013 Rand/AMA study mentioned earlier, some 75 percent of practices of all sizes reported that they were at least "somewhat satisfied" with their EHR.

Challenges and Opportunities: The major EHR challenges we'll discuss in this section are: 1) how to efficiently and accurately collect high-quality, comprehensive clinical data from busy and often computer-averse providers; 2) how to visualize EHR data in a way that supports a provider's mental model in order to improve the quality and efficiency of care; 3) how to incorporate efficient work processes and flow into EHRs; and 4) how to achieve meaningful interoperability, in particular to support care coordination.

First, why are EHRs a good idea? The fundamental answer is that physicians are in the data business. They collect data, analyze that data, make decisions based on it and follow up using still more data so they can make adjustments as needed. Despite this, medical records and record keeping often receive minimal attention in medical training and most medical schools offer no training at all in health informatics beyond learning how to use whatever electronic record systems are deployed in their hospitals and clinics.

The observation that medical data is not maintained as it should be may have first been made in 1863 by Florence Nightingale, the founder of modern nursing but also a social reformer and, more pertinently, a statistician:

> "In attempting to arrive at the truth, I have applied everywhere for information but scarcely an instance have I been able to obtain hospital records fit for any purpose of comparison. If they could be obtained they would enable us to decide many other questions besides the ones alluded to. They would show subscribers how their money was spent, what amount of good was really being done with it or whether their money was not doing mischief rather than good."[4]

---

[3] http://www.softwareadvice.com/medical/userview/ehr-report-2014/.

[4] Nightingale F (1863) *Notes on Hospitals*, 3rd Edition (Longman, Green, Longman, Roberts & Green, London), p. 175.

Dr. Lawrence L. Weed is best known for his problem-oriented medical record, but he may also have been as responsible as anyone for reintroducing this problem in the mid-1960s when he wrote that:

> "Medical schools should hold their graduates responsible for their medical records no matter where they may be. Indeed if they were to survey the performance of their graduates through their records they would help community hospitals with their standards and begin to get some feed-back on all the medical education programmes underway. In addition, records on a large scale would become available for computer analysis."[5]

Dr. Weed also introduced a computerized version of the POMR at the Medical Center Hospital of Vermont, as part of the Problem-Oriented Medical Information System project. Given the technology of the day, the system required a mini-computer at each bedside but provided a touch-screen interface, a very innovative technology for the late 1960s.

Even prior to this, in 1961, IBM and Akron Children's Hospital began planning a system and announced the initiative in early 1962. Patient data was entered by a nurse or assistant into an IBM 305 RAMAC, the first commercial computer that used a magnetic hard drive for storage.[6] In a statement that might well still be made today, Roger Sherman, the hospital's administrator, said: "If we can mechanize much of this routine clerical work, our doctors and nurses will be able to spend more of their time using their professional training to give more direct and attentive care to patients."[7] It is worth noting that today, over 50 years later, a common complaint from physicians is that the time spent doing electronic charting detracts from patient interaction.

By the late 1970s there were several operational ambulatory electronic medical record systems. A study of them by the Office of Technology Assessment concluded that "Medical information systems can be used to educate and assist medical professionals during clinical care, reducing the need to rely on memory. Potentially, they can increase efficiency and reduce or contain institutional costs. They can provide a way to monitor and evaluate the quality of medical services. They can eliminate duplication of data collection and can provide accurate, accessible data for evaluating and planning medical care services. Finally, they can be used to supply data that have previously been unavailable to researchers and policy makers."[8] Once again, these arguments for digital records are just as relevant today.

However, despite the long recognition of their potential value, the adoption of clinically meaningful EMR systems was scant up through 2008–09 when a pair of studies published in the *New England Journal of Medicine* showed that only around

[5]Medical Records, Patient Care, and Medical Education, Irish J Med Sci. 462:271-82.

[6]http://www-03.ibm.com/ibm/history/exhibits/storage/storage_PH0305.html.

[7]http://bits.blogs.nytimes.com/2012/02/17/the-miracle-of-digital-health-records-50-years-ago/.

[8]http://ota.fas.org/reports/7708.pdf.

4 percent of practicing physicians[9] and 1.5 percent of non-federal hospitals[10] were adopters at a level that is similar to what we would today call Meaningful Use.

These low adoption levels persisted despite research that has consistently shown that medicine, as it is currently practiced, is error prone and patients are often damaged as a result. So, another aspect of the data problem is looking at why physicians make errors. The University of Washington's School of Medicine lists three principle reasons and data management is central to all of them:

1) An **incomplete knowledge base:** The physician can't know everything given the rapidly increasing storehouse of medical research and knowledge.
2) An **error in perception or judgment:** Physicians are human and are subject to human error.
3) A **lapse in attention:** Physicians are very busy and may not focus sufficiently on each case.[11]

This is not a new observation. In its landmark 1999 publication, *To Err Is Human: Building a Safer Health System,* the IOM famously said that from 44,000–98,000 Americans are killed each year in U.S. hospitals due to largely preventable reasons. Many of those listed—such as adverse drug events, improper transfusions, wrong-site surgery and mistaken patient identity—have to do with mishandling of data. This leads to one of the key arguments for adoption of electronic medical records.

Physician adopters tend to agree. Figure 1 presents some key results from the 2013 Rand/AMA's *Survey of Factors Affecting Physician Professional Satisfaction*. This survey sought to determine the factors leading to professional satisfaction and involved 30 practices, 28 of which had EHRs. Initially physicians weren't specifically surveyed about their EHRs, but it soon became apparent that they are a determinant of professional satisfaction so questions were added to explore that. The results represent the sentiments of 22 surveyed practices that used EHRs and were interviewed after these questions were added. It found that 61 percent of responding physicians felt that their EHR improved care quality. However, that same percentage felt that the EHR required them to perform tasks that other staff members could perform.[12] The study found a high correlation between EHR satisfaction and overall professional satisfaction. Those physicians who felt the EHRs disrupted workflow or patient interaction had lower overall job satisfaction, while those who felt EHRs improved the quality of care had higher overall professional job satisfaction.

[9] http://www.nejm.org/doi/full/10.1056/NEJMsa0802005.
[10] http://www.nejm.org/doi/pdf/10.1056/NEJMsa0900592.
[11] http://depts.washington.edu/bioethx/topics/mistks.html.
[12] http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR439/RAND_RR439.pdf.

% Strongly Agree or Agree	Total	PCP	Surgical Specialist	Nonsurgical specialist	Other
Administrative/Financial Benefits					
Faster and more accurate billing for services	74	80	67	77	72
Time savings through e-prescribing	67	78	64	64	71
Savings from not managing/storing paper	59	66	51	58	66
Practice/worksite efficiency increase	53	61	50	51	60
Clinical Benefits					
Improved communication and care coordination	67	56	64	70	76
Immediate data availability	59	63	53	56	77
Guideline prompts and timely lab results	56	64	49	55	63
Patient submission of information	41	43	42	38	51

Figure 1: Key results from the 2013 Rand/AMA survey of physicians shows that the majority note both administrative/financial as well as clinical benefits from the use of their EHR. The data suggests that specialists may see the benefits of care coordination to a greater degree than the PCPs who refer patients to them. Much earlier, we saw that specialists often see patients with incomplete or no data on prior care. (Adapted from the 2013 Rand/AMA Physician EHR Survey)

The *Deloitte Center for Health Solutions' 2013 Survey of U.S. Physicians* looked more specifically at the issue of how an EHR could improve care quality. It provides recent data from 613 randomly selected physicians and found (with a margin of error of 3.89 percent at the 95 percent confidence level) that two-thirds of the surveyed physicians cited improved communication and care coordination as a clinical benefit of their EHR.[13]

The Data Collection Challenge: All of these surveys indicate physician dissatisfaction with spending time on menial tasks and most of this is in data entry. In addition to problems with inefficient data entry and loss of nuances in notes, electronic records can introduce new sources of error. The root cause for much of this is the magnitude of the data collection challenge. Recall how large and complex the data standards are. That isn't arbitrary—human anatomy, physiology and disease is that complex. It is also, of course, not as quantifiable or well understood as the data used in other fields, such as engineering, which involve objects and systems designed and built by humans. To accurately reflect patients—their diseases

[13]http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/Health%20Care%20Provider/us_dchs_2013PhysicianSurveyHIT_051313%20(2).pdf.

and their personalities as well as their family and life situations—an electronic record must be able to capture nuances that are not easily represented by a coding system. Yet, to be maximally useful for many secondary purposes such as reporting and research, the EHR would ideally collect most of the record as structured, coded information. How do you do that? How can you provide busy physicians with a way to accurately and efficiently collect such a complex dataset in a form that is maximally useful for a variety of purposes in addition to direct patient care? That's the EHR data collection challenge—and it has been with us for decades. The following paragraphs explore different structural approaches and some of the causes for error in data collection.

The most common approach is a template—essentially an electronic version of the paper forms physicians have used for years in their practices. Another, less commonly used approach, is charting against a structured "tree." In this approach, once a patient problem or symptom is identified, the likely data that would be collected about it (such as associated symptoms and physical exam findings) can be documented by selecting the correct values from lists of choices that are offered based on their relevancy to the problem or symptom already indicated. This is a form of adaptive charting and it does have the potential to offer a richer set of structured choices than most templates, but it can lead to excessive clicking in order to navigate down multiple trees to find and select the appropriate values.

Templates are the most popular option in large part because they are familiar. However, once again, physicians complain of excessive clicking to enter information. Moreover, where a list of possible choices is long, it is far too easy to pick the wrong one. This can be the case particularly if the EHR does not "adapt" to the clinical situation by limiting the choices to those most relevant to the current patient. This is termed an "active error"—the entry of inaccurate information. It's one of the seven common sources of EHR data entry errors cited in 2011 in a panel discussion of Usability in Health Information Technology at the Fifth Annual MIT Information Quality Industry Symposium.[14]

Another source of error identified at this MIT symposium was incorrect system defaults. This can occur in at least two ways. An EHR could have been set up to default to a particular value which the provider may overlook that while charting, thereby accepting a default value that is inappropriate for the current patient. Another far more insidious feature is when the EHR "remembers" the values from a patient's prior visit to save the physician time in charting the current visit. The obvious potential error is to document what may have been correct previously but is wrong now. Another source of error derives from the desire to save the provider time navigating among screens by fitting as much as possible onto each data entry screen. This can lead to a cluttered and/or illogical layout that can cause information to be overlooked.

The next source of error cited at this seminar—inconsistent data across locations—should not come as a surprise. This is the interoperability issue we've been discussing

[14] http://mitiq.mit.edu/IQIS/Documents/CDOIQS_201177/Papers/00_01_Proceedings_Intro.pdf.

throughout the book. If multiple physicians are involved in a patient's care, aggregating all the data is a potential challenge. However, if it is aggregated, accuracy can be compromised by varieties of ways different clinicians document the same clinical issues, particularly where text entry is involved. For example, they might use different terms for the same thing or nonstandard abbreviations or acronyms. Similarly, physicians may not record all the needed data leading to potential misinterpretations of the past history or even the current treatments. Finally, where multiple locations are involved in a patient's care, it may not be clear which of several observations/events is definitive or even the most up-to-date.

The data entry challenge should now be clear. Physicians are busy. Their work depends on data taken from a huge and complex set of possible values. They have little, if any, informatics training. EHRs could be better able to help with this. It should be possible to innovate and offer approaches to the computer-physician interface that are more imaginative and ultimately more successful than mimicking the paper tools that have been used for decades. We turn to that now through a series of case studies. Before we do that, though, I need to make it clear that I have no financial or other relationship with any of these companies. I found them by actively seeking innovative approaches to charting. There may well be other, and possibly even better, companies or approaches of which I am unaware. Finally, I am not recommending any specific EHR system for any particular practice.

Case Study: M*Modal. Back in the 1970s when I was involved in developing an early EMR, the data collection problem was already clear. In fact, for the most part our physicians dictated notes which were transcribed into the EMR. I used to tell visitors that "someday" voice recognition would make this far more efficient and accurate. Has that day arrived? To answer that question we'll look at M*Modal, one of the most advanced medically specific "speech understanding" technologies. Originally the company was MultiModal Technologies, founded in 2001 by recent Carnegie-Mellon Computer Science PhDs. It changed its name to M*Modal and, iIn 2011, it merged with MedQuist and the merged company adopted its name. As shown in Figure 2, the technology not only translates voice to text but it recognizes clinical concepts in that text and codes them into SNOMED-CT. This is very impressive and offers an interesting approach to the data entry challenge since it allows physicians to continue to express things in familiar and efficient ways while producing coded information as a byproduct. However, it is not yet perfect. The specific accuracy levels are highly dependent on a) noise levels, b) medical specialty, c) amount of similar data previously seen, and d) the subset of SNOMED-CT codes of interest (e.g. allergies, diseases, symptoms, procedures). Very specific clinical concepts, such as smoking history, that are normally expressed in one of a few possible ways can be more accurately identified than concepts that can have a large number of possible expressions, such as the patient's problem. Under the right circumstances, high levels of accuracy are possible. For example, the company says it has 97–100 percent speech understanding accuracy for radiologists dictating on high quality microphones in a quiet image reading room.

As impressive as it is, M*Modal is a tool; it isn't an EHR. The company spent years offering the technology only to medical transcription companies. Millions of

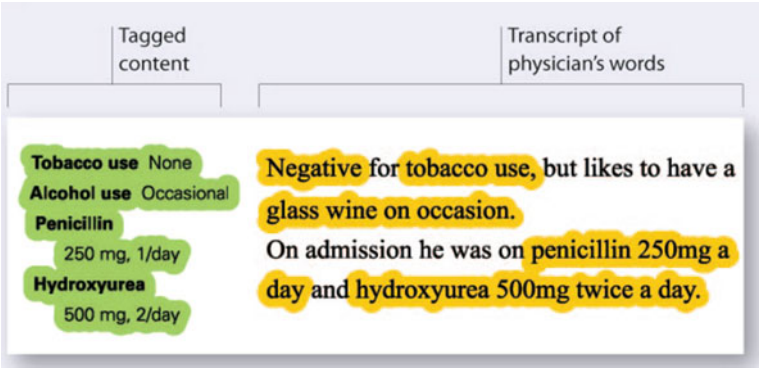


Figure 2: M*Modal's voice engine converts dictation to text, finds clinical concept in that text and encodes them into data standards such as SNOMED-CT. (Courtesy M*Modal, All Rights Reserved)

physicians' notes were transcribed by the system which also indicated things it couldn't deal with. These were fixed by a human expert. All of this activity was performed on the company's servers (the system is cloud-based) so, in addition to fixing the transcriptions, the human employees were also training the system to get better. Once the company felt it was ready, they offered to integrate it into commercial EHRs. The first to do that was Greenway Medical Technologies which offers it as part of its PrimeSuite EHR. The voice recognition is built into the software and is an option for data entry in a number of clinical contexts. However, the recognition of clinical concepts with encoding to SNOMED-CT, while planned, is not yet a part of PrimeSuite.

Case Study: Medical Informatics Engineering (MIE). MIE was founded in 1995 by Doug Horner, a Purdue engineering graduate, as a health information exchange including a clinical repository that was mostly for laboratory test data. This was at the very beginning of the Internet so it is particularly remarkable (and a testament to Doug's foresight) that the company was web-based from the outset and even provided the needed network connectivity to their clients' offices. Over a series of years, the company's offering evolved into an EHR starting with high-speed digital storage and management of transcribed documents. The next phase was to replace paper charts using scanning integrated with bar code technology so the system would understand the contents of each scanned document, technology Doug had previously developed for banking. They began with a dermatology practice where the providers were already utilizing transcribed dictation for their notes. Because of its image management, the service allowed them to more compactly have their dictation and diagrams in one place. MIE claims to have eliminated paper charts in this dermatology practice while reducing the number of transcriptionists from six to two by increasing their productivity.

In 1998, MIE introduced WebChart (which may well have been the first web-based EMR). They introduced their second generation "Minimally Invasive EMR"

around 2000. Doug feels that physicians are and should be "cognitive not clerical." To help with this, the system "learns" each doctor's practice patterns and, over time, can predict what they are going to input into specific fields and fill it in for them, a concept that is now widely used in the system. These learned concepts are not problem specific. Instead, the system answers questions based on the patient's history and the most common questions for each chief complaint. Another capability of the system is to anticipate what the physician will do. For example, in a case typically requiring antibiotic therapy, the drug the physician has most often used in the past for similar cases would appear first on their list of choices. These preferences (e.g. each physician's approach) are built into a "medical library" allowing nurses to switch libraries as they rotate among the physicians in their practice. MIE provides some interesting case studies on its web site.[15]

The company also offers a patient portal and the NoMoreClipboard personal health record that is designed to work with any EHR system. Partially as a result of offering both capabilities, the company is installed extensively in occupational health clinics for major employers who typically use both the EHR and PHR. To give one example of the integration between the two systems when used together, the EHR evaluates patients' risks of developing particular chronic diseases, such as diabetes or hypertension, and tells the PHR what questions each patient should answer to further evaluate risk and monitor progression toward that disease or achieve success in avoiding it.

Case Study: Praxis. Praxis was founded by Richard Low, MD, a UCLA/Yale-educated physician, who first trained in surgery and practiced emergency medicine before changing to internal medicine. He recalls attending a seminar given by a physician/lawyer and first realizing how important medical documentation is. He says he later found out that the "average physician spends 2.5 hours per day doing paperwork, that's 8.5 years of his career." Based on his examination of clinical records, he determined that "no two doctors chart the same." His goal was to develop something that would save doctors time while allowing them to maintain their individual approach to charting. To accomplish this, he started Praxis in his native Argentina in 1989 with $15,000 of his own money which, he says, went a great deal further there than it would have in the U.S. The Praxis EMR was introduced in 1993 and, since its second product release in 1998, the profitable company, now based in Marina Del Rey, CA, has grown with no outside financing.

Ironically, the way Praxis makes documentation easier and more efficient for physicians is, in itself, not easy to understand. At its core is the notion of a "clinical concept." Praxis does not define these concepts in a standardized way as SNOMED does. Rather, it considers a concept to be an indivisible clinical unit in the view of each individual physician. In essence, a concept is a basic element of the way each physician thinks about medicine.

Based on this idea, the company developed what it calls a "concept processor." The computer science term for this is an artificial neural network (ANN). An ANN

[15] http://www.mieweb.com/solutions/webchart-ehr-case-studies/.

is an advanced computing approach used to model complex relationships between inputs and outputs or to find patterns in data. It is being used widely in many industries. A bank, for example, might use an ANN to find people who are more likely to be submitting fraudulent credit card charges based on patterns that might not be obvious at all to a person looking at the data, in part because of its massive scale. If your bank has ever called you to verify a credit card charge that call may well have been triggered by their ANN. The technology is also being applied in healthcare to look for fraudulent claims.

The basic idea behind Praxis is that physicians develop a method of going from inputs—such as the history, physical exam and lab results—to outputs, such as a diagnosis and treatment plan. Praxis says that its concept processor "learns" how each individual physician does this for the problems they commonly see and uses that knowledge to save the physician time on subsequent visits by anticipating what they will likely do and document. This prediction is based on the *most similar* past patient. The system starts as a "blank slate" in each practice but, over time, it gets better and better at finding the closest encounter to the current one. After around 50 iterations of a particular problem the system is well trained and, according to the company, can accurately find the closest matching or even "identical" prior encounter. Based on that, it brings up the clinical concepts the physician has used in the past for similar patients, as illustrated in Figure 3. This saves the physician time and also serves to provide clinical reminders, reducing the chance that something will be overlooked or forgotten. Praxis has (laudably, in my view) avoided the temptation to automatically document for the physician who must specifically

Figure 3: Based on the physician's entry of the most specific clinical finding, Praxis presents a likely note based on the most similar past patient(s). To avoid charting unintended "default values" the physician must click on each clinical concept to confirm and/or edit it. (Courtesy Praxis, All Rights Reserved)

decide what to chart for each patient from the list of concepts presented by clicking on each of them. The physician can, of course, add more or edit as appropriate for each case.

Documentation flow in Praxis is quite different from traditional charting. It normally doesn't begin with the chief complaint—the patient's own, typically vague, statement of their problem—as physicians are trained to do. This is because the chief complaint will usually not provide sufficient specificity for the concept processor to find the best prior case. Physicians using Praxis are trained to first enter the most clinically specific "input" they can (e.g. a clinical finding). This allows the concept processor to more accurately find the identical or best match to the patient being seen. Physicians can group subsets of an overall problem, such as acute pharyngitis (sore throat), according to whatever clinical divisions make sense to them. This further assists in getting to the best possible matching case and its associated clinical concepts. Physicians can post their own personal approach to clinical concepts so, for example, a family physician can import the approach used by an expert neurologist to evaluate headache.

Real-Time Clinical Decision Support: We've seen how Praxis uses machine learning to help physicians document more efficiently. In yet another example of its potential, analytics can be used to point out to a provider potential documentation errors *as they are being made*.

Despite their obvious importance, patient problems lists are often inaccurate, incomplete or out of date. Research has shown that problem list completeness can range as low as 4.7 percent for complex patients to around 50–60 percent for common chronic disease patients to a maximum of 78.5 percent for breast cancer.[16] Meaningful Use requires that providers must "maintain an up-to-date problem list of current and active diagnoses" so new tools are needed to help providers meet this requirement.

Automated methods to accomplish this often use either problem inference (or proxy) rules or natural language processing (NLP) techniques. Problem inference techniques use related clinical information such as laboratory tests, medications and billing codes to infer problems. For example, a patient receiving metformin who has had multiple abnormal HbA1c tests is likely to have diabetes. NLP is used to analyze and code information from free-text entries such as progress notes. We just saw an example of this in our M*Modal case study.

An interesting 2012 paper in the *Journal of the American Medical Informatics Association* reported on a randomized controlled study done by researchers at the Brigham and Women's Hospital in Boston.[17] A clinical alerting system, as shown in Figure 4, uses inference rules to detect and notify providers of missing documentation of any of 17 clinical study problems (asthma/COPD, breast cancer, coronary artery disease, congenital coagulopathy, CHF, diabetes mellitus, glaucoma, hypertension,

[16] http://jamia.bmj.com/content/18/6/859.abstract?ijkey=bfc5d354051e747061432b084e8a7a102e991435&keytype2=tf_ipsecsha.

[17] http://jamia.bmj.com/content/early/2012/01/03/amiajnl-2011-000521.full#ref-6.

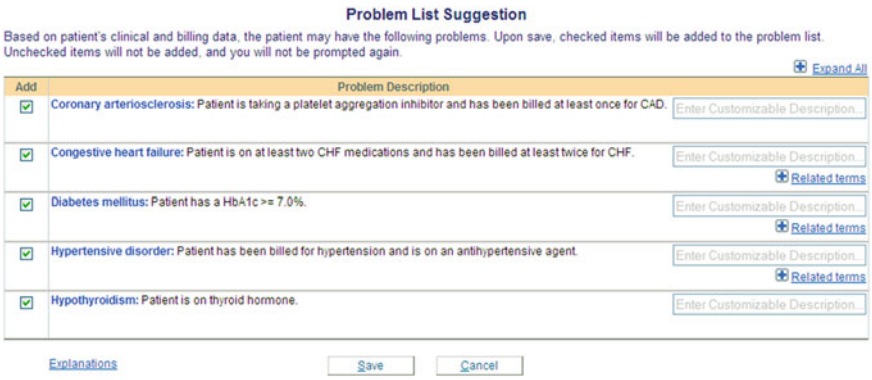


Figure 4: Inference rules use clinical data such as medications and lab tests as well as billing to infer the presence of clinical problems and prompt physicians to add missing problems to a patient's list. (Courtesy Adam Wright, Partners HealthCare)

hyperthyroidism, hypothyroidism, myasthenia gravis, osteoporosis/osteopenia, rheumatoid arthritis, renal failure/insufficiency, sickle cell disease and stroke). Overall, 14 of the 17 study problems were approximately three times more likely to be documented when alerts were shown (the statistically significant differences ranged from 1.54 to 6.89).

At least one early stage company, Diameter Health,[18] is developing algorithms to do similar error checking in clinical documentation. However, they say their approach is different from the Brigham work which is focused more directly on hospital-specific documentation. Rather, they use CCD documents as input, suggesting an ambulatory care focus. As of this writing, the company says they are in the rapidly developing prototype stage incorporating feedback from early clients.

Physician Preferences: Do physicians actually prefer EMRs with innovative approaches to data collection? What follows is largely my own opinion, but it is based on survey data. I leave it to you to explore the latest (2012) EHR survey data from the AAFP.[19] It groups physicians by practice size and presents their most common EHR systems and satisfaction levels with those systems. With the increasing tendency of health systems to employ physicians, particularly those in larger practices located nearby, it is increasingly likely that a large practice will implement the EHR provided by the enterprise software vendor selected by the health system. Based on my own admittedly subjective, but direct, observations of several (but not all) of the top rated systems, the physicians that are in smaller

[18] http://diameterhealth.com/.

[19] http://www.aafp.org/fpm/2012/1100/p23.html.

practices and more likely chose EHRs on their own, do in fact, pick more innovative or well-designed solutions to data collection such as those we've discussed. The 2014 Medscape EHR report finds similar results.[20]

The Data Visualization Challenge: Another major EMR challenge appears after the data is collected. How to best present the data to physicians in a way that supports what *Computational Technology for Effective Health Care*: *Immediate Steps and Strategic Directions*, an important National Research Committee report, refers to as "their mental model"? The report describes this by saying that clinicians "spend a great deal of time sifting through large amounts of raw data, when ideally, IT systems would place raw data into context with current medical knowledge to provide clinicians with computer models that depict the health status of the patient."[21]

Earlier, in the discussion of data segmentation, we looked at a design proposal from Dr. Jonathan Nebeker to illustrate the added power that comes when software understands clinical relationships. This is also a key objective of the DS2 project at the University of Illinois' ONC-funded SHARP project, which was also aimed at a more automated approach to data segmentation. An understanding of clinical relationships would also be a key technical capability for putting raw clinical data into the right context for each patient. Of course, as you know, SNOMED-CT provides a matrix of relationships that could assist where clinical data from an EHR is coded into it.

To my knowledge, this capability doesn't yet exist in a real-world EHR, so I don't have a case study. We'll have to speculate as to how an EHR with this capability would support the physician's mental model. We'll do this by discussing what I consider a major deficiency in most, if not all, EHRs—they don't distinguish among fairly obvious care scenarios in which they are used.

Consider once again a multichronic disease patient's chart. It was actually charts like this that got me interested in electronic records as a medical student working in an outpatient diabetes clinic. The diabetic patients' charts were very large and complex. In fact, as a result of their heft, we often used red wagons to cart them around! Given such a chart, a PCP would need a quick overview to adequately plan the visit and provide comprehensive care. The electronic equivalent of the thick chart on a red wagon typically doesn't facilitate gaining that quick overview to the extent it could. However, that same PCP seeing an otherwise healthy patient visiting because of a sprained knee is concerned with little else other than possibly any routine preventive care that might be due. Two different patients and two different clinical situations would each best be served by an EHR that recognized those differences and presented the right data, in the right format, at the right time. Yet most, if not all, commercial EHRs can't make this distinction, at least not automatically.

[20] http://www.medscape.com/features/slideshow/public/ehr2014#1.

[21] http://www.nap.edu/catalog/12572.html.

On the other hand, a specialist seeing that same multichronic disease patient will be focused mainly—if not entirely—on the organ or body system for which that physician cares. As a result, different specialists will probably be interested in different data from the same patient's very large chart. Each of them will prescribe medications and other treatments that are specific to the disease being managed. Current EHRs generally don't understand this and don't automatically make these distinctions, although both MIE and Praxis might suggest, or at least prioritize in list form, the most likely treatments for a condition that each specialist sees regularly.

We'll see later on how the analysis of large clinical datasets derived in part from EHRs is providing new knowledge and insights. We've seen some early attempts to introduce these same concepts into EHRs themselves, such as the incorporation of SMART on FHIR into Cerner's PowerChart EHR. It seems inevitable that, as time goes on, EHRs will become more "clinically adaptive" to both facilitate the capture of data and to present it in the most useful way for a specific physician seeing a specific patient for a specific reason. Any provider considering the purchase of an EHR would be wise to look for vendors that are at least starting down that road. A good alternative would be vendors that embrace concepts such as FHIR that enable third-party developers to do innovative things with the data in their EHR.

Process and Workflow: Integration into and actually improving workflow and process are important goals if information technology is to be maximally effective in improving both care quality and practice efficiency. This has consistently proven harder to achieve than goals such as improved care coordination. For example, the *Deloitte Center for Health Solutions' 2013 Survey of U.S. Physicians* we cited earlier showed that, while two-thirds of physicians whose practice has an EHR meeting Meaningful Use Stage 1 requirements report "improved communication and care coordination capabilities", only around half of those physicians report a "practice or worksite efficiency increase."[22] The 2012 AAFP survey we also cited earlier reported that the "areas of lowest satisfaction are EHRs' effects on productivity, their effects on the physician's ability to focus on patient care and vendor support (only 16 percent, 24 percent and 36 percent positive responses, respectively)."[23]

It is arguably too early to fully assess the situation. A 2011 review of research on the benefits of EHR adoption included studies that assessed the impact on efficiency and seems to find mixed results. For example, "a study that assessed the efficiency effects of a health IT implementation could find that it both decreases transcription costs yet increases the time physicians spend performing administrative functions related to the electronic health record."[24] A 2010 paper in the *Annals of Family Medicine* studied the gradual implementation of an EHR in a residency-based

[22] http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/us_chs_2013SurveyofUSPhysicians_031813.pdf.

[23] http://www.aafp.org/fpm/2012/1100/p23.html.

[24] http://content.healthaffairs.org/content/30/3/464.full.

family medicine outpatient clinic and concludes that "work flow improved with EHR implementation."[25] This same theme of gradual implementation and incremental change is described in an interesting video on the ONC web site.[26]

As we'll now discuss, I speculate that at least some of the differences found by research into the impact of EHRs on physician efficiency relate to how well they have been implemented within the context of each practice's workflows and processes. A commonly made mistake is to keep essentially the same workflows and processes rather than to carefully re-engineer them to take maximal advantage of the capabilities of the new information system.

It is probably useful to think about this challenge from two perspectives: the characteristics of the EHR that support efficient, and even improved, workflow and process and the impact the individual practice implementation has. MIE's WebChart provides an interesting example of workflow-oriented EHR design. To facilitate a gradual transition, as suggested by the paper and the video cited in the previous paragraph, WebChart supports multiple forms of data entry to document patient encounters, orders and other components of the patient visit. As we mentioned earlier, paper documents can be bar coded and stored in the EHR which can then treat them more intelligently. For example, it is possible that a patient's lab profile for a particular test might include a proper chronological sequence consisting of both digital results received electronically and results on paper. The EHR knows the results belong to that profile and knows where to place them within it because of the bar coded information.

No matter what the EHR, the way it is implemented matters. Ideally, a practice would analyze and understand its manual workflow and processes and decide in advance of automation what their goals are for new, improved and more efficient approaches. Without this, it is far too easy to introduce automation into a business and end up with the same approaches that were appropriate in a manual world but which do not take advantage of the capabilities of the new electronic system. It is even possible that the practice will end up simultaneously supporting both the old manual process and a new and different electronic version. This is a recipe for reduced productivity in the face of "automation." Such an analysis and plan can also inform a more educated and appropriate system selection. Without it, systems may be selected, as they are far too often, based on a lists of features, each of which might represent the desires of a specific person or functional area, but with no real thought as to how the entire system will function to improve the workflow and processes of the practice. I cannot overemphasize how often this last, deeply flawed scenario plays out when health providers conduct a system search. It is also important that physicians who will be users of these systems make a serious effort to understand and participate in the initial process of defining the goals of their practice's EHR selection and implementation.

---

[25] http://www.annfammed.org/content/8/4/316.full.

[26] http://www.healthit.gov/providers-professionals/video/dr-scott-webb-s-ehr-story-efficiency-and-patient-interaction.

There is at least one interesting 2008 study from the *Journal of General Internal Medicine* that seems to confirm that experienced EHR users do put a greater emphasis on workflow and process. It compared the perceived advantages of EHRs as seen by paper-based practices embarking on their first EHR implementation to practices planning to replace an existing EHR. The paper-based practices seemingly focused more on functionality and listed among their priorities ability to have remote access of patient information, improved revenues and improved communication among and between practitioners and patients. In addition to a better system, practices with an existing EHR listed enhanced streamlining of workflow processes as a significant priority from a new system.[27]

The Interoperability "Opportunity": You may be surprised to learn that there was pushback within the health informatics community when the Obama administration proposed HITECH. The often-expressed concern was that it would lead to the widespread deployment of EHRs that weren't interoperable and that had many of the usability issues we've just discussed. Of course, these problems had existed for decades and it wasn't clear how they could be resolved without strong action by the government. Moreover, EHR adoption levels had remained very low over the same long time period and, arguably, would never improve substantially without new incentives to encourage adoption. It is likely (but I cannot verify this) that the architects of HITECH knew going in that the interoperability problem would become problematic and would have to be addressed, but they chose to deploy now and fix it later.

One commonly expressed idea for a solution is to "wrap" these systems in a layer that provides interoperability and can support independently developed apps. Apps are typically smaller, more specialized programs used for a specific function. As with the now ubiquitous smartphone apps, they are developed by many different organizations but all run within a common platform provided, for example, by the phone or, in this case, a software layer that interacts with the enterprise software.

A 2013 conference at Arizona State University and organized by Dr. Robert Greenes focused on this idea of a "universal app platform" as an interoperability solution. The app platform would sit above an interoperability layer that would convert EHR data into a standard model that the apps could interact with. Only the interoperability layer would need to interact with each EHR through its own, proprietary data model. This is similar in many respects to what a translator does. The development of apps would be greatly facilitated by eliminating the need to deal with each EHR's particular, and often proprietary, data representations. If this were done correctly, an app could work with any underlying EHR. We've already encountered this concept: this is what SMART on FHIR is designed to accomplish. We've even seen an example of an app using the standard FHIR data model translated for it by an FHIR adapter developed by the enterprise EHR vendor, Cerner.

[27] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2517887/.

In that earlier discussion, we considered apps built into a patient's EHR that could access the required data for that patient and offer CDS, education or some other service. Not mentioned earlier were some particularly intriguing possible uses of this app platform that are directly responsive to the EHR challenges we've been discussing. For example, apps for specialist physicians could work equally well and the same way no matter what hospital that physician is visiting or what EHR it uses. These same apps might be "adaptive" in the ways we've suggested. In fact, it is likely that the existence of such a platform would greatly lower the barrier to market entry, thereby encouraging entrepreneurs and innovators to try new approaches. This also applies to individual physicians who could use whatever app best suits their preferences for data entry and retrieval and not be limited to the tools provided by their particular EHR vendor.

In fact, this sort of technological thinking and approach is already widespread on the Internet where interoperability has long been recognized as being in everyone's best interest. The proposed next generation of the Internet is termed the "semantic web" and it would be able to "understand" the content on web pages through a standard system of tags called the Resource Description Framework (RDF). I attended a 2013 Dagstuhl workshop in Germany where there was a great deal of discussion about bringing healthcare into the RDF as a way of establishing interoperability.[28]

However, as we've seen for HIE, everyone's interests are not necessarily aligned toward interoperability. A health system that has spent many millions of dollars to implement an integrated information system that extends out to its referral and patient network isn't necessarily interested in a "level playing field" with other health systems in its area that would, at least to some degree, negate the competitive advantage it is trying to create.

Despite this, there is clear evidence that the vendor community is feeling market (and governmental) pressure to come up with interoperability solutions. CommonWell Health Alliance™ was a surprise announcement at HIMSS 2013. It currently consists of 14 total members including seven founding members and seven additional contributing or general members. Membership currently represents acute and ambulatory care EHR suppliers, as well as laboratory, retail pharmacy, perinatal care and long-term care health IT systems. CommonWell members seek to define and promote a national infrastructure with common standards and policies, and build interoperability into their software so that providers can use them seamlessly within their existing workflow.

The group says that "we believe that in order for health IT to work, it must be inherently interoperable. As such, we are working hard to define and promote a national infrastructure with common standards and policies, as well as offer services that are embedded natively within vendors' own software to help solve many of the challenges associated with interoperability."[29]

[28] http://www.dagstuhl.de/de/programm/kalender/semhp/?semnr=13342.

[29] http://www.commonwellalliance.org/.

RelayHealth, part of McKesson Provider Technologies, the largest health IT vendor by revenue, is the initial provider to deliver the necessary collaboration services including:

- **Patient identification and linking:** Assist health IT suppliers to more quickly and accurately identify patients as they transition through care facilities.
- **Record locator and retrieval:** Help providers locate and access their patients' records, regardless of where the encounter occurred, by providing a "virtual table of contents" that documents available data from each encounter location.
- **Patient access, privacy and consent management:** Deliver a patient-authorized means to simplify management of data sharing consents and authorizations.
- **Trusted data access:** Provide certification, authentication and auditing services that facilitate consistent trusted data sharing among member systems.[30]

The trusted data access services are of particular interest for our purposes. Through them, CommonWell supports sharing of clinical data that is encoded into SNOMED-CT, LOINC or other formal ontologies that have an OID designation. Interestingly and significantly, CommonWell also supports FHIR, although at present, this is limited to administrative purposes.[31] The service was initially launched in four geographies (Chicago, IL; Elkin and Henderson, NC; and Columbia, SC) at more than 12 provider sites. In November 2014 CommonWell announced the transition from initial launch to nationwide expansion of its services. At the time of the announcement, athenahealth, Cerner, CPSI, Greenway Health and McKesson had signed member service agreements with CommonWell in order to offer its services to their clients nationwide.

It is often noted that two of the enterprise software vendors with the largest market share (Epic and MEDITECH) and other major EHR vendors are not members of the alliance. It is possible that pressure from their clients may bring more vendors into the alliance. For example, the widely respected Cleveland Clinic has developed and maintains its own web services APIs (developed before FHIR, but it is being considered for the next version) to provide more facile access to the data in its Epic EMR and it also creates apps that use them. Significantly, this is bidirectional technology, so the apps don't just utilize data from the EMR but also store the data they collect back into it.

Examples of these include:

- An app through which rounding physicians can pull up patient lists on their phones to review a key subset of clinical data from the EMR. Patient lists are stratified based on a risk score.
- A mobile "vital signs" entry application for nurses. The application also issues alerts if there is a significant percent change since the last entry. An interesting

[30] http://www.commonwellalliance.org/faqs/.

[31] http://www.commonwellalliance.org/wp-content/uploads/2014/10/CommonWell-Services-Specification.pdf.

feature is that nurses can customize the user interface to their preferences in order to better adapt to their workflow. It is claimed to reduce errors and missing documentation.

- A patient-facing questionnaire app that uses web services to access the EMR and leave a note if the patient has not answered a specified list of questions, dynamically produced based on their record in the EMR. The questionnaires also adjust based on what each patient has already entered. The results appear in flow sheets the physician can review from within the EMR.[32]

At present, Epic dominates the large health system enterprise software marketplace and, largely as a result of that, has the largest installed bae of EHRs with 11.6% of the market and 53.7% of practices with over 40 providers.[33] As a result, their position on interoperability is of particular interest. In 2013, at its annual user group meeting, the company released a concept called Open Epic which supports a number of protocols and standards (including FHIR[34]). However, the clinical use cases provided on the Open Epic site seem to be far more focused on importing data into the Epic EHR than obtaining data from it. There does not appear to be a significant capability to retrieve individual data elements, as would be required for some of the functionality implemented by the Cleveland Clinic, but it is certainly possible that this might come in later releases, particularly as their support for FHIR expands.[35]

Epic's founder and CEO, Judy Faulkner, is questioned about their interoperability views in an interesting 2013 *Forbes* interview.[36] In it she was specifically asked about the concept of an app platform and in response, she mentioned the work at the Cleveland Clinic we just discussed and went on to express a view somewhat different from the more open app platform concept other vendors have implemented:

> "Developers have to work through a customer. We don't let anyone write on top of our platform, come read our code and study our software. I worry about intellectual property at that point. With our customers, we make sure we have signed agreements. They know they have to respect our software. Customers can do it in a controlled environment, but not the whole world. You'll see us do more and more of that."

However, her thinking may be evolving because, just as I was completing this book, EPIC announced that they would be opening an app store to be called App Exchange. Details are not available as I write this.

Absent widespread industry agreement, the interoperability challenge may only be solved through action by the federal government. The issue and the possibility of

---

[32] http://www.radiologybusiness.com/topics/imaging-informatics/siim-experiment-web-technologies-points-future-health-it?nopaging=1.

[33] http://www.skainfo.com/health_care_market_reports/EMR_Electronic_Medical_Records.pdf.

[34] http://open.epic.com/Interface/FHIR.

[35] http://open.epic.com/Clinical/HL7v2.

[36] http://www.forbes.com/sites/zinamoukheiber/2013/05/15/a-chat-with-epic-systems-ceo-judy-faulkner/.

a government approach is getting increased attention. H.R. 4157 (109th), the Health Information Technology Promotion Act of 2006, is a proposed law to do just that and it passed the House, but not the Senate. H.R. 4015: SGR Repeal and Medicare Provider Payment Modernization Act of 2014 also passed the House. Among other things, this proposed law would have required EHRs to be "interoperable by 2017" and would "prohibit providers from deliberately blocking information sharing with other EHR vendors."[37] The difficulty the VA and DOD have had in making their separate EHRs interoperable has brought more attention to the issue in Washington and interoperability was a key theme at the 2014 HIMSS conference.

How would it be accomplished? Earlier we mentioned the JASON report *A Robust Health Data Infrastructure*[38] that proposes something quite similar to the interoperability layer, using open, but mandatory (for EHR vendors) APIs to create an app platform through which providers, patients and payers would access data from any and all certified EHRs. We also mentioned that the JASON task force was convened by CMS and ONC to recommend how to implement the JASON proposals and that it both largely supported them and specifically recommended an effort to finish FHIR (it is currently a draft standard) in time for the 2017 launch of Meaningful Use Stage 3.

Some argue that actually implementing the JASON proposal after such a mandate might take several years or even a decade. The Argonaut project suggests that the vendor community is prepared to move faster than that, in part because of pressure from major health system clients and in part by the desire to avoid a government mandate. Just as I was completing this book, ONC released *Connecting Health and Care for the Nation A Shared Nationwide Interoperability Roadmap DRAFT Version 1.0*. Since it won't be finalized until this book is in print, I won't comment on it in much detail beyond listing ONC's four key objectives:

1) Establish a coordinated governance framework and process for nationwide health IT interoperability
2) Improve technical standards and implementation guidance for sharing and using a common clinical data set
3) Enhance incentives for sharing electronic health information according to common technical standards, starting with a common clinical data set
4) Clarify privacy and security requirements that enable interoperability

In this draft, ONC seems to be taking a neutral stance toward what the available, improved technical standards for sharing and using a common clinical data set might be. It discusses Direct, SOAP-based web services and FHIR as "several transport techniques widely adopted today." However, the Standards for Interoperability section of its 10-year "Timeline of Select High-Level Critical Actions for Near-Term Wins" calls for implementing a FHIR specification for what it calls a

[37] http://www.californiahealthline.org/articles/2014/2/7/lawmakers-announce-bipartisan-deal-to-repeal-replace-sgr.

[38] http://healthit.gov/sites/default/files/ptp13-700hhs_white.pdf.

"common clinical dataset" by the end of 2016 (just in time for Meaningful Use Stage 3 as the JASON Report suggested). The only other specific technology in that section of the timeline is the release of implementation guidance for CCDA 2.0 by around the same time. Both the timeline and an illustration of the dataset are found in the reference cited here.[39]

The IOM's Vision: The Institute of Medicine of the National Academy of Sciences envisions a "learning health system designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety and value in health care."[40] Achieving this vision requires more than solving the data collection and visualization challenges we've identified and the IOM provides specific guidance we'll turn to next. However, a careful reading of the vision should convince anyone that achieving it will depend on obtaining higher-quality clinical data, analyzing it and presenting the results in useful ways for healthcare providers and their patients.

The specific capabilities and care delivery innovations for which the IOM calls align well with what we've discussed so far and will be discussing later on. Specifically, EHRs should store **comprehensive patient data** and use it to provide **cognitive support for providers and patients**. Aspects of this are **integrating patient-specific data** (where possible) in order to personalize care, integrating **evidence-based guidelines** to improve outcomes, reducing costs and enhancing the safety of care and **rapidly integrating new knowledge and technology** into clinical practice.

The IOM also goes on to advocate for both **practice-wide** and **population-level** care management, something that significantly scales up the data to be managed and a clear opportunity for innovative visualizations of the kind we'll see in later sections.

The IOM also specifically calls for recognizing the **home as a new care setting**, which aligns with more use of home-based technologies for monitoring and treatment, and the use of the information technologies such as personal health records to promote engagement and communication with health professionals.[41] The IOM also clearly envisions a healthcare system in which patients are far more engaged and involved than has traditionally been the case. This is increasingly becoming both practical and affordable through the patient-facing technologies and systems we'll discuss in the next section.

---

[39] http://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf.

[40] http://iom.edu/Reports/2012/Best-Care-at-Lower-Cost-The-Path-to-Continuously-Learning-Health-Carein-America.aspx.

[41] *Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions*, (Stead and Lin 2009), 165–166.

Patient Engagement and Empowerment

We believe that all individuals should be able to readily access, understand and use their personal health information.

—*Johns Hopkins Web Site*[1]

Years ago, my professor told me that, as a practicing physician, I would occasionally see chronic disease patients for a few minutes and think that what I did during their visit would make all the difference. In fact, he said, it was what *those patients did between the visits* that would make that difference. Patients at home have historically been largely disconnected from what we think of as the healthcare system. Today, we have an increasingly rich set of technologies and tools to empower patients to be more involved in maintaining their wellness, preventing disease and managing disease more effectively, should they develop it. These tools can also connect patients to their providers, to interested family members and friends and to others with similar conditions or health issues. I encourage physicians, who have not done so, to create a PHR of their own and then encourage their patients to do the same using any of the publicly available tools that are listed on ONC's site.[2] The process is generally straightforward and not particularly time consuming and the systems generally provide good online help. For those needing more help, the American Health Information Management Association provides an online guide to setting up a PHR.[3]

Patients increasingly want to engage digitally with their health providers and their medical records and feel this would lead to more participation in their own care. The proposed rules for meaningful use stage 3 recognize the value of patient-generated data by requiring that at least 15 % of an eligible provider's records contain it. It is also clear that physician encouragement and education of their patients to use these tools is key. A survey of 2017 U.S. adults conducted online in

[1] http://www.hopkinsmedicine.org/howard_county_general_hospital/patient_visitor/medical_records/my_personal_health_record.html.
[2] http://healthit.gov/patients-families/maintain-your-medical-record.
[3] http://www.myphr.com/.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_10

September 2014 found a majority (64 percent) do not currently use online patient portals. Of these, 35 percent did not know a portal was available and 31 percent said their physician had never spoken to them about portals. However, 57 percent of those who don't use a portal say they would be much more interested and proactive in their personal healthcare if they had online access to their medical records. Importantly, of those Americans who do use online patient portals, 59 percent say they have been much more interested and proactive in their personal healthcare since they received access. This interesting survey also looked at generational differences in attitudes about access to electronic records and the preferred platforms for accessing them.[4] A September 2014 ONC white paper provides detailed analysis of another consumer survey on attitudes about the use of online patient records. It found that: "A majority of individuals who accessed their online medical record found the information very useful. Seven in 10 individuals nationwide reported that having an online medical record was very or somewhat important. Among individuals who did access their online medical record over 90 percent perceived having access as very or somewhat important, and notably, a majority (62 percent) of individuals who did not access their online record considered online access to be somewhat or very important." Based on these results, it concluded that "even in the early stages of online medical records, individuals do value and find use from basic capabilities to access their online medical record."[5] A nationwide survey of 406 U.S. patients who had recently seen their PCP (or had an appointment scheduled) suggests that the availability of digital tools influences patient selection of a physician.[6]

Of course, patients can use the Internet to search for health related information but, even if they are able to properly phrase a search, the results can still be problematic given the many biased or even bogus health-related web sites. To help with this, Google recently announced that their widely used search engine would respond to queries about common health conditions with vetted, relevant medical facts.[7] Beyond patient initiated search, you're already familiar with what may well be the key patient-facing (e.g. for direct use by patients) technology – the personal health record. While the exact form of a PHR is evolving, the central ideas are clear and have been so for some time. A PHR is **patient controlled**; **patients can record data** in it and that data can become part of their electronic health record. **Patients control access** to their PHR data; which can and ideally does become a **lifelong health record.**

We'll use Microsoft's HealthVault as the example, but as we just discussed, there are several widely used PHRs listed on an ONC web page and they are largely functionally similar. PHRs are definitely not just for consumers who already have medical problems. For quite some time now, PHR users have been able to record their allergies, medical problems, medications, lab test results and other health-related information. They have also been able to record data they may have

[4] http://news.xerox.com/news/Xerox-EHR-survey-finds-Americans-open-to-online-records.

[5] http://www.healthit.gov/sites/default/files/consumeraccessdatabrief_9_10_14.pdf.

[6] http://technologyadvice.com/medical/blog/study-patients-value-digital-services/.

[7] http://googleblog.blogspot.com/2015/02/health-info-knowledge-graph.html.

measured themselves at home, such as weight, blood pressure or blood glucose. While it is desirable that PHRs record such a rich array of information, historically many people have found the work involved more than they want to take on. People are also challenged by unfamiliar terminology. To help with this, HealthVault offers smart lists that will focus on and display the likely correct terms as people type in a few letters.

More recently, to help patients access their medical data, Meaningful Use has dictated the export of standardized clinical summaries (such as the CCD), largely alleviating the burden on the patient to record that data. As we discussed earlier, Stage 1 requires that more than 50 percent of patients seen by an eligible provider in each reporting period have access to an electronic clinical summary within three business days of their visit. Stage 2 requires that "more than 5 percent of all unique patients seen by each eligible provider during the EHR reporting period (or their authorized representatives) view, download or transmit to a third party their health information." Clearly, patients downloading their record into a PHR would meet that requirement. This also means that patients using a PHR could import these summaries from all providers caring for them, potentially making the patient the owner of a comprehensive health record, as called for by the IOM. To support this, HealthVault provides each user with a Direct e–mail address to which their providers can send a CCD. HealthVault can extract clinical information from the CCD and use it to populate the appropriate fields in the PHR substantially reducing the data entry burden for the patient.

With the growth of the wireless Internet and ubiquitous computing, PHRs are now accessible anytime/anywhere, typically via a smartphone app. They have also become app platforms from which independent developers can write programs that can access patient-stored data in the PHR. As of this writing, HealthVault hosts 117 apps for purposes such as accessing and storing lab test results or imaging studies, importing prescription information from a major pharmacy chain connecting with health systems or engaging with physicians for the management of heart disease or diabetes. These last two apps are provided by the American Heart Association and the American Diabetes Association and have both patient-facing and physician-facing components. The physician view provides a convenient dashboard through which all patients using the app can be monitored in one place.

As we said earlier, HealthVault and other PHRs can upload data from devices the patient may use at home. As of this writing, HealthVault can communicate with 230 devices that capture physiologic and/or fitness data. I've intentionally not accompanied this PHR discussion with any graphics. This is perhaps the most accessible technology we'll discuss so, if you don't already have one, I encourage you to create a HealthVault or an alternate PHR account and explore its capabilities. If your healthcare provider can provide one, try to get your CCD and upload it into your PHR. You don't need a Direct address to do this—a USB stick will suffice.

Finally, to further facilitate the collection of personal data, PHRs are increasingly being combined with wearable devices such as special watches or wrist bands that are able to support people interested in maintaining their health by automatically recording traditional physiological data (such as heart rate) along with other

information typically not found in a medical record, such as daily activity level, sleep patterns and even level of exposure to ultraviolet radiation.[8,9,10]

OpenNotes: OpenNotes goes beyond giving patients access to clinical summaries and offers them the opportunity to read and correct their actual medical records, including visit notes written by their doctors, nurses or other clinicians. It began in 2010, with support from the Robert Wood Johnson Foundation, when more than 100 primary care doctors from three medical institutions began sharing notes online with their patients.[11] There are now at least 20 participating institutions, including leading organizations such as the MD Anderson Cancer Center, Beth Israel Deaconess Medical Center and the Geisinger Health System.[12]

I attended a talk about the experience at Geisinger, one of the three pilot sites, in which the speaker explained that provider attitudes about OpenNotes improved with experience using the technology. This is reinforced by an early study of the first three sites in the *Annals of Internal Medicine* which reported that: "Overall, 69–81 percent of participating PCPs across the three sites and 92–97 percent of patients thought open visit notes were a good idea, compared with 16–33 percent of nonparticipating PCPs."[13]

A 2013 paper looked in more detail at patient response to the technology and reported three main benefits:

> First, patients reported that seeing their records had a positive effect on care communication between visits as well as during encounters. Second, access was felt to improve patients' knowledge about their own health and prompted greater desire for self-care. Third, patients reported that health record access improved participation in their care in a variety of ways.

Interestingly, in view of the emphasis Meaningful Use places on patients having access to and being able to share their health records, the study also found that:

> More than half (55.43 percent, 2503/4516) of the participants who reported viewing at least one visit note would like the option of letting family members or friends have their own Web access to their visit notes and 21.70 percent (980/4516) reported sharing their visit notes with someone during the study year.[14]

Geisinger has now significantly expanded its OpenNotes project to 1100 providers and approximately 170,000 patients. In an article in *Modern Health Care*, Geisinger CEO, Dr. Glen Steele, is reported to have said in an October 2014 keynote talk at the College of Healthcare Information Management Executives that

---

[8] http://www.microsoft.com/microsoft-band/en-us.

[9] https://www.apple.com/watch/.

[10] http://www.samsung.com/global/microsite/gear/gearfit_features.html.

[11] http://www.rwjf.org/content/dam/web-assets/2010/07/open-notes.

[12] http://www.myopennotes.org/who-is-sharing-notes/.

[13] http://annals.org/article.aspx?articleid=1033220.

[14] http://www.jmir.org/2013/3/e65/.

"more than 80 percent of patients opened the doctor's notes included in their record … and the majority reported feeling more in control of their care, better prepared for their visits and more likely to take medications as prescribed."[15]

He went on to discuss one of the early goals of OpenNotes: patients finding and correcting errors in their records. Often-cited concerns by physicians not using the technology are fear of malpractice suits or loss of patient respect and trust. These fears have not been supported at Geisinger where Steele said that, as a result of patients accessing care and their records online, "the amount of error correction is significant" but that despite this, "providers have been 'minimally aggravated' by the changes in the system." This is a fascinating new technology and quite a number of articles are listed on the OpenNotes site, a good place for readers interested in learning more about it.[16]

Case Study: PatientsLikeMe: PatientsLikeMe was founded in 2004 by three MIT engineers, Benjamin and James Heywood (brothers) and their longtime friend, Jeff Cole. Five years earlier, the Heywoods' brother, Stephen, had been diagnosed with amyotrophic lateral sclerosis (ALS, also known as Lou Gehrig's disease) at the age of 29. The family soon began searching the world for ideas that would extend and improve Stephen's life. Based on this experience, they saw the need for an environment to share and collect data, typically on innovative treatments for incurable disease. To accomplish this, they built a unique social networking system over a research platform, since getting patients engaged in aggregated clinical research was their primary mission.

The site now encompasses well over 1000 conditions, even though prior to April 2011, there were only 20. It is free to patients and accepts no advertising, but it is a for-profit business. The objective is to gather data from patients about their illness experience and make it available in aggregated form to organizations that are interested in particular patient cohorts. Examples might be pharmaceutical companies or companies with other early-stage medical products that want to learn from patients having the condition they seek to treat. For example, a pharmaceutical company might partner with the site to create a portal (a site within the site) for engaging organ transplant recipients where it can talk to and learn from them, while at the same time considering their aggregated data. The site recently launched a suite of three services designed to facilitate participation by its members in clinical trials:

- **Trial Access** for pharmaceutical companies to develop and deploy custom research programs and an expanding repository of patient opinions and attitudes about participating in clinical trials.

[15] http://www.modernhealthcare.com/article/20141030/NEWS/310309945/geisinger-using-technology-to-increase-patient-engagement-ceo-says.

[16] http://www.myopennotes.org/research/publications/.

- **Community Access** a collaboration tool for researchers and PatientsLikeMe experts to gain insights from patient-reported data.
- **Access Services** which allows companies to collect and analyze real-world data from PatientsLikeMe participants.[17]

To create a clinically relevant research platform, PatientsLikeMe uses structured surveys to collect patient-reported data. Novel treatment, symptom and condition data populate the "User Voice dashboard," where it is reviewed and curated to assure data integrity. The company receives around 75 user voice entries per day. Some may be duplicative. For example, there could be a spelling difference or the patient could have entered two concepts together, such as "pain and depression." The spelling error would be recognized to avoid duplicate concepts and the combined concepts would be split so the patient can monitor each separately and each can be aggregated for research purposes. All clinical data is coded in the background using standardized terminologies. Symptoms and side effects are coded into SNOMED-CT and the Medical Dictionary for Regulatory Activities, a medical terminology used to classify adverse events associated with the use of biopharmaceuticals and other medical products. Diagnoses are coded into ICD-10.

Despite this high degree of coding, as much as possible of the "patient voice" is maintained.

PatientsLikeMe points out that the patients self-manage most of their care. As shown in Figure 1, the site helps patients put their conditions in context; organize the status of symptoms, treatments and side effects; and prepare themselves for a clinician encounter through the use of a profile they create. They try to help patients answer the question "given my status, what is the best outcome I can hope to achieve and how do I get there?" To help them understand the answer, the site offers patients connections to other similar patients and patient communities.

PatientsLikeMe is perhaps best known for a dramatic research study initiated by ALS patients themselves. A report from Italy suggested that lithium might slow the progression of their disease.[18] To validate this study, a group of ALS patients decided to experiment on their own with the treatment. Patients using lithium asked the company for support to find the effects, if any. After 12 months of using the tools on the site, patients showed that lithium had no effect on their disease progression.[19] A similar traditional clinical trial would, of course, have taken significantly longer. This nontraditional study dramatically illustrates the potential for patient empowerment to even reach the level of accelerated clinical research by making health records more accessible and easier to share, aggregate and analyze.

Patient Portals: Before PHRs and sites like PatientsLikeMe, hospitals and physicians were able to establish "portals," essentially web sites where their patients

[17] http://news.patientslikeme.com/services.

[18] http://www.pnas.org/content/105/6/2052.long.

[19] Wick, P., T.E. Vaughan, M.P. Massagli, and J. Heywood. 2011. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol* 29(5):411–414.

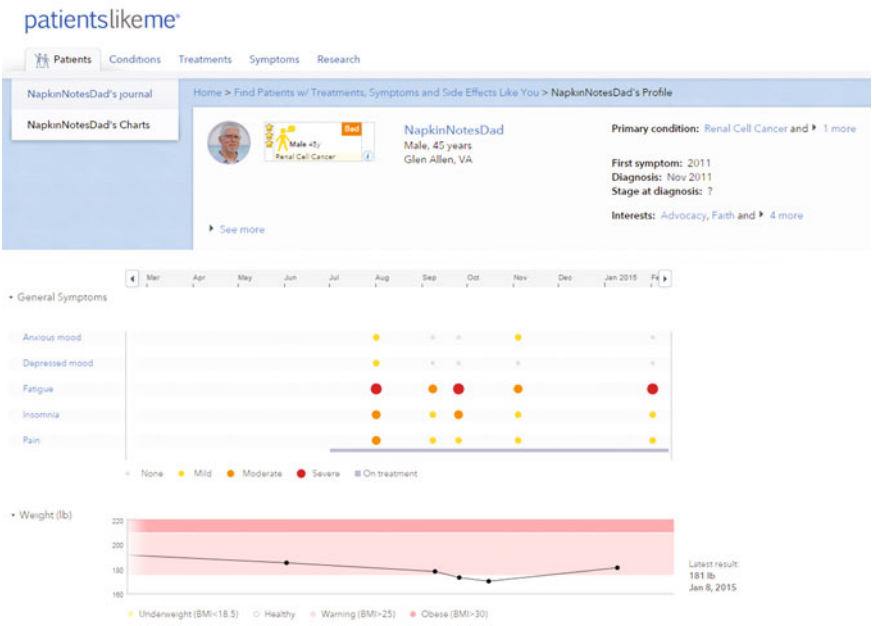


Figure 1: PatientsLikeMe provides patients with an integrated view of their health to organize and track their symptoms, treatments, side effects labs and tests and other data, so they can see how they are progressing and prepare for clinician encounters. Here are two selected sections of a profile from a member, a 45 year old cancer patient, showing what he has reported as general symptoms and weight in the past year. (http://www.patientslikeme.com/patients/view/231000?utm_source=plm_blog&utm_medium=blog&utm_campaign=garth_callaghan_december). (Courtesy PatientsLikeMe, All Rights Reserved)

could do some of the things we have previously discussed under personal health records. Often patients also have access to other functions when the portal is more integrated than the typical PHR with the information systems of the healthcare organization providing their care. The first patient portal was developed by a company that was acquired by RelayHealth, the same company we mentioned as the initial provider of the CommonWell Alliance's interoperability services. As shown in Figure 2, patients using the RelayHealth portal can take advantage of integration with the healthcare organization's systems to get their lab test results, request or even make appointments and securely communicate with their physicians, even using Direct. Except for lab test results, these are not services typically available in a PHR. Patients can also pay their bills, something also not typically offered by PHR sites. These additional services emphasize a key difference between portals and PHRs. Portals are provided to the patient, typically by their provider, while PHRs are generally patient-initiated and operated. Providers use portals to improve their service to their patients by making it more convenient to use them and their practice, in part to insure that they remain loyal. It can also be the case that the

Figure 2: Patients can obtain lab test results, send messages to their physicians, make appointments or pay bills using the RelayHealth portal. They can also export or download their health data through its support of Direct and the Blue Button. (Courtesy RelayHealth, All Rights Reserved)

"self-service" features of a portal offload work from the provider's staff, making the office more efficient. Importantly, patients can download their health data from portals, something that will probably cause them to be more widely deployed and promoted to patients as a result of the Meaningful Use VDT mandate. Patients might obtain their records by downloading a CCD from the portal or via a Direct e–mail. It might also be done using Blue Button, the patient-facing technology we turn to next.

Blue Button: Before discussing Blue Button it is important to note that the largest health system in the U.S., the VA, was an early and successful EHR developer. Given its size, the VA's Veterans Health Information Systems and Technology Architecture system is the country's most widely used EHR. My students are always fascinated when a physician from the Atlanta VA shows them how any data for any VA patient can be accessed and viewed together with all similar data (such as images) with just a few clicks, no matter where in the huge VA

system of 163 hospitals, some 800 clinics and 135 nursing homes the data was collected. The VA was also early to emphasize patient engagement. This led to its early adoption of the Blue Button initiative that began at a Markle Foundation Work Group on Consumer Engagement meeting in early 2010.[20] The original idea was to make it easy for VA patients to access, download and share their health data. Blue Button was a simple human-readable text file when it was adopted by ONC in 2012.[21] Their standards and interoperability workgroup expanded the concept, added more structure and dealt with Direct integration to download the patient's record from a portal and move it to the patient's PHR. As a result, a patient seeing multiple providers could bring data from all of them into one patient-managed technology, such as a PHR.

In the latest version, Blue Button+, the data is an XML document formatted according to CCDA. You should recall that CCDA documents are constructed from templates for the document, its sections and the individual data entries. Figure 3 provides a list of some of the sections of the Blue Button+ CCDA to give you a more detailed idea of the data that it can contain.

Even more recently, the Blue Button REST API effort offers an approach to using Blue Button+ for both its traditional role, push—where providers send the files to patients or their designated PHR/app as well as using it for pull—where patients can, in effect, subscribe and receive updates as new data comes into their provider's EHR. We earlier used the Growth-tastic! app as a proof-of-concept of the Blue Button+ pull approach to accessing data. As new data becomes available, the app can retrieve it and automatically update the growth chart.[22]

Home Telehealth: In the HealthVault discussion we saw that patients could upload data from a large number of devices. This is, of course, increasingly possible, inexpensive and convenient with embedded motion and other sensors in smartphones being combined with sensor-containing smart watches or wrist bands designed to work with them. Companies are building more and more sensors into these devices. A recent entrant, the Microsoft Band, is claimed to track heart rate, steps walked, calories burned and sleep quality.[23] It's clear these devices are primarily aimed at maintaining fitness, but they can be equally applicable to managing disease, once it develops. Physiological sensing for consumers and patients is growing more sophisticated. Wrist pulse oximetry (measuring the oxygen content of blood) is already a reality. More advanced devices, such as noninvasive glucometers, are undoubtedly coming in the future. The continued microminiaturization of electronics probably means that a single future wearable device will essentially be a virtually complete clinical and fitness monitoring station.

You may be surprised to learn that collecting data in the home to help with care is actually quite an old idea. In the mid-1990s Steve Kaufman, one of the earliest

[20] http://en.wikipedia.org/wiki/Blue_Button_(health_maintenance).

[21] http://www.healthit.gov/patients-families/join-blue-button-movement.

[22] http://growth-pull.bluebuttonpl.us/.

[23] http://www.microsoft.com/microsoft-band/en-us.

Section	Description
Header	Patient information and demographics
Allergies, Adverse Reactions, and Alerts	Includes status and severity of each.
Encounters	Surgeries, ED visits, etc.
Immunizations	Immunizations and vaccines
Medications	As prescribed by the provider
Care Plan	Planned activities and encounters
Discharge Medications	Part of hospital discharge summary
Reason for Referral	Written reason for referral
Problem List	Concerns, complaints, and observations
Procedures	History of procedures
Functional and Cognitive Status	List of impairments
Results	Includes laboratory tests
Social History	Observations like smoking, drinking, etc.
Vital Signs	Includes height, weight, blood pressure, etc.
Discharge Instructions	Written discharge instructions

Figure 3: Sections of the Blue Button+ XML according to CCDA. (Source ONC)

innovators in this space, introduced his Home Assisted Nursing Care (HANC) robot that offered a wide variety of voice-controlled nursing services to patients at home.[24] Among other things, it dispensed medications at the proper time and assisted in taking physiological measurements using the relatively large and non-intelligent, patient-operated devices of the day.

The traditional assumption was a more passive role for the patient than is appropriate and technologically possible today. We've seen that, with HealthVault, patients can initiate their own data collection program and send the results as a CCD to their physician. Medicine has largely not yet caught up with this reality and, as we saw from survey data, patients are often unaware of it. But this will change,

[24] http://www.homemods.org/resources/life-span/high.shtml.

particularly as incentives foster an alignment between the increasing independent capabilities of patients and the need of providers to deliver more cost-effective care. This, of course, closely parallels what has already happened in other domains, such as financial services, travel and shopping, where consumers have assumed many of the roles previously played by service or product providers. As we discussed earlier, patient portals have already introduced this self-service concept to healthcare.

Since HANC, numerous commercially packaged home telehealth systems (including one that I was involved in developing starting in the late 1990s) have sought to assist patients in recording **subjective data** about aspects of care including **symptoms**, **activity level** and **compliance** with treatment plans, as well as **objective data** such as **weight**, **blood pressure**, **pulse**, **temperature**, **blood oxygen levels** and even heart activity through an **electrocardiogram**.

Despite a great deal of research and commercial activity, another key piece of objective data—**medication compliance**—has remained very difficult to obtain in an accurate and cost effective manner. Devices and software are available to remind patients to take their medications, to monitor when they open a pill container and to even dispense the proper medications at the right time. However, there are continuing challenges with updating the more advanced devices when medication orders change. Filling medication dispensing systems can also be challenging for patients and their caregivers. The gold standard would be having objective knowledge that patients actually consumed their medications at the proper time and in the proper dosages.

The leader in that space may well be Proteus® Digital Health. To measure ingestion, the company offers an FDA-cleared technology that includes three components: an ingestible sensor that is taken as a part of (if it is incorporated into the medication by the manufacturer—something the company says it is working on) or along with the medication (at patient initiation); a wearable patch that receives a signal emitted by the sensor when it is activated in the stomach and also captures data such as heart rate and patient activity; and a Bluetooth-enabled device (such as a smartphone) that can receive data from the patch for use by the company's Helius software system for managing and monitoring compliance by patients, their families and their providers.[25]

Mobile Devices, Sensors and Apps: As we've seen, a growing variety of objective data are increasingly available from low cost, wearable devices that are increasingly grouped together in smartphones and in associated smartwatches, wrist bands or other devices that can collect data and communicate it to a smartphone. When combined with apps on the phone, they can create a self-management system that often includes the ability for family members or caregivers to monitor the status of an elderly patient. This trend is accelerating as mobile device manufacturers increasingly target health as a market. The iPhone 5 introduced integrated activity monitoring. This data could become more useful through Apple's recent introduction of HealthKit—an aggregation point for health data from multiple sources

[25] http://www.proteus.com/.

and through which apps on its smartphone or tablet can access that data.[26] Other major smartphone companies are likely to introduce similar capabilities. At the same time that it introduced its Band wearable device as an open technology that would work with any major smartphone, Microsoft announced its Health Intelligence Engine. The company claims it will provide "actionable insights" based on data gathered from a variety of devices, including those provided by other organizations.[27] Interoperability seems to be gathering momentum in this new space as companies recognize that, in the long term, analytic services may be of greater commercial value than the devices that collect the data used by those services. We'll look at efforts toward interoperability of mobile health technologies later on in this section.

As mobile devices and the apps that can run on them become more sophisticated the FDA is grappling with the degree to which they should be regulated as medical devices and how it should regulate the software apps that increasingly manipulate the data from those devices.[28] AliveCor's smartphone case that can monitor a patient's heart beat is a good example of advanced clinical functionality that has been through the FDA process. The company claims that its software can "immediately detect atrial fibrillation in an ECG and track trends"; make the results available to the users of the devices; and give them the ability to share those results with their physicians.[29] This does seem to have clinical value, as reported by an article in the respected journal *Circulation*, in which researchers found they could successfully use the device to screen for atrial fibrillation.[30] It seems clear, based on results like this, that mobile devices and apps are becoming serious medical tools and that FDA regulation may be appropriate, particularly when devices or associated software apps manipulate or visualize the data they collect in order to give advice or influence the behavior or decisions of either patients or physicians. On the other hand, devices and apps increasingly target prevention and wellness, areas that would presumably not be subject to FDA oversight.

To clarify this, in February 2015 the FDA issued a pair of final guidance documents. With respect to software it said it intends to apply its regulatory oversight to "only those mobile apps that are medical devices and whose functionality could pose a risk to a patient's safety if the mobile app were to not function as intended".[31] It also said it does not intend to enforce compliance with the regulatory controls that apply to Medical Device Data Systems (MDDS) that include medical devices, medical image storage devices, and medical image communications

[26] https://developer.apple.com/healthkit/.

[27] http://www.microsoft.com/microsoft-health/en-us.

[28] http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ConnectedHealth/MobileMedicalApplications/ucm255978.htm.

[29] http://www.alivecor.com/.

[30] http://circ.ahajournals.org/cgi/content/meeting_abstract/126/21_MeetingAbstracts/A16810?sid=eccbd879-a917-4c97-95a5-3cd91cf91327.

[31] http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf?source=govdelivery&utm_medium=email&utm_source=govdelivery.

devices that it defines as "those that do not modify the data and do not control the functions or parameters of any connected medical device." An MDDS does not include devices intended for active patient monitoring and these would be subject to FDA oversight.[32]

No matter their goal, PHR apps may have an advantage in providing patients with advice because they have access to a wide range of information about the patient, including their medical records. The AHRQ maintains a database of innovative activities using various forms of telehealth for a wide variety of purposes.[33]

It may also be increasingly possible to monitor and understand patient behavior at home, a critical part of changing it to better manage chronic disease, or to make earlier and more accurate diagnoses of subtle behavioral conditions, such as autism. This field is called Computational Behavior Science or "behavior imaging" and is an active research area at Georgia Tech[34] where it is the basis for at least one commercial spin out.[35]

Ginger.io Case Study. Ginger.io is in a closely related space and uses smartphones to improve the delivery of mental healthcare. The company uses simple surveys, passive data collected from patients' smartphone sensors and deep analytics models developed at MIT. Its mobile application works in conjunction with a provider-facing, web-based dashboard to identify patterns in patient behavior and reported mental state that may impact health and wellbeing.

The information collected from the patient-facing app is presented in the dashboard to identify patients in need, in order to proactively deliver the right care to the right people at the right time. Related research at MIT and Harvard shows that, by reaching patients in their time of need, these targeted interventions can improve care efficiency, reduce readmissions and hospitalizations and strengthen the connection between patients and providers.[36]

Virtual visits is a final application of technology in the home. Inexpensive video conferencing is now widely available on virtually any computer device, smartphone, tablet and even some televisions that provide app platforms and integrated web cams. As a result, direct interaction between a patient at home and a professional or other caregiver can be inexpensive and simple to achieve. This application is not without controversy. These services may not be reimbursed by insurance. Further, some state medical boards take a dim view of physicians treating patients over the Internet if they have not previously actually seen that patient in person.

[32] http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM401996.pdf?source=govdelivery&utm_medium=email&utm_source=govdelivery.

[33] https://innovations.ahrq.gov/taxonomy-terms/telehealth.

[34] http://www.cbs.gatech.edu/.

[35] http://behaviorimaging.com/.

[36] http://web.media.mit.edu/~anmol/fp-325madan.pdf.

Despite this, direct provision of virtual medical care is accelerating. One interesting example is conducting virtual appointments with patients who cannot travel to the Courage Kenny Advanced Primary Care Clinic (a primary care medical home).[37] Entrepreneurs recognize this opportunity and have started companies, such as HealthByConnect, that provide technologies that allow physicians to provide Internet-based care to their patients.[38] The potential for this approach is illustrated by its adoption by the largest U.S. healthcare insurance company, WellPoint. It now offers video visits to its members and other consumers, where this form of care is legal, under its trade name LiveHealth Online. Sessions last around 10 min and cost $49. Physicians can prescribe some medications, if appropriate, where that is legal. The technology is provided by Boston-based American Well which offers a similar service to consumers via its own network of primary care providers.[39]

Challenges: The challenges in using technology for patient engagement and empowerment are similar in many respects to those with EHRs. While efficiency may not be as important, usability is arguably even more important and difficult to achieve. Patients, including the elderly who may live alone or not have support at home, will often need to use these technologies without assistance. The technologies are obviously capable of generating vast amounts of data from millions of patients. No one has time to look at it all, so analytic and visualization tools that minimize false positives and create actionable information that triggers timely, clinically relevant alerts are of critical importance. This is yet another area where machine learning and other tools of artificial intelligence are being applied in healthcare. RIMIDI, an Atlanta start-up, exemplifies this by using predictive analytics to support more proactive diabetes care by physicians based on glucose readings taken by patients at home.

Another key challenge is data privacy and security. You may recall that in the section on privacy, security and trust we discussed studies that show that a substantial majority of patients are concerned about the misuse of their digital data and that these concerns are a major reason for the still-low adoption of patient-facing technologies despite a high degree of satisfaction among patients who are using them. The frequent press accounts of data breeches in many industries almost certainly add to these concerns. Up until now I've not mentioned one aspect of the JASON report—its call for new approaches to securing health data. This is a highly technical subject that is well beyond the scope of this book but it is increasingly clear that new approaches are needed and that one promising alternative is to secure the data itself—both at rest in a database and in transit among care providers, patients and other valid users—such that those illicitly acquiring it would still have no access to it for unintended or illegal purposes.

[37] https://innovations.ahrq.gov/profiles/medical-home-patients-disabilities-and-chronic-conditions-improves-access-and-self.

[38] https://www.healthbyconnect.com/.

[39] https://www.americanwell.com/.

As we mentioned earlier, interoperability is also an issue for devices in the home. Vendors have typically wanted a revenue stream based on data collected by their device(s) and have usually offered a web portal to access that data, often in vendor-specific, proprietary formats. This creates the obvious problem of "data siloes," further complicating the already-complex problem of interoperability and retarding the creation of an integrated, comprehensive view of patients and their care. We've discussed the role that PHRs or tools like Human API could play in bringing together data from diverse sources, aggregating it and making it available under each patient's control. The JASON report includes "self-reported data from embedded and wireless sensors" in its list of data sources. Implementation of JASON might largely solve this problem.

However, the problem is so obvious, that other solutions have been proposed or are being implemented. The Continua Health Alliance is a non-profit, open collaboration of healthcare and technology companies with over 200 member companies around the world. It is dedicated to establishing a system of interoperable, personal connected health and fitness devices and associated software. The alliance says that it is not a standards-creation body. It is currently selecting initial connectivity standards and "is working to identify and resolve gaps in some standards bodies so that personal telehealth solutions are interoperable and contribute toward improved health management." Additionally, the alliance is writing guidelines and publishing white papers on specifically how to use standards to achieve interoperability across many companies and devices.[40]

Case Study: Open mHealth. This non-profit collaboration is also seeking to address the interoperability problem created by what it says are more than 10,000 patient-facing apps and devices. It is doing this by working with multiple stakeholders in the mobile health space to design a common language for health data, an open developer platform and proof-of-concept patient-facing tools using the platform. The effort must deal with many of the data access issues we discussed with respect to Human API and it uses the same JSON notation used by FHIR to represent its data model. However, the models are different because FHIR, at least at present, is provider-facing, while Open mHealth (along with Human API) is patient-facing.

So, while FHIR accesses generally more robust and detailed data from EHRs, Open mHealth is concerned with the typically limited data available from patient-facing technologies and devices. However, in both approaches, data conversion must be performed. Data from EHRs is converted to the FHIR data model by software often called an "adapter." Open mHealth uses the term "shim" to refer to similar software that might, for example, convert data accessed via the HealthVault API to its Open mHealth equivalent. The glucose shim for HealthVault is called "omh:blood-glucose:1.0" and it would convert HealthVault's representation of a plasma glucose level shown here:

[40] http://www.continuaalliance.org/.

```
<?xml version="1.0" encoding="utf-16"?>
<blood-glucose>
 <when>
  <date>
   <y>2006</y>
   <m>1</m>
   <d>1</d>
  </date>
  <time>
   <h>9</h>
   <m>30</m>
   <s>0</s>
   <f>0</f>
  </time>
 </when>
 <value>
  <mmolPerL>7.444444</mmolPerL>
  <display units="mmolPerL">7.444444</display>
 </value>
 <glucose-measurement-type>
  <text>Whole blood</text>
  <code>
   <value>wb</value>
   <family>wc</family>
   <type>glucose-measurement-type</type>
   <version>1</version>
  </code>
 </glucose-measurement-type>
 <outside-operating-temp>true</outside-operating-temp>
 <is-control-test>true</is-control-test>
 <normalcy>1</normalcy>
 <measurement-context>
  <text>Before meal</text>
  <code>
   <value>BeforeMeal</value>
   <family>wc</family>
   <type>glucose-measurement-context</type>
   <version>1</version>
  </code>
 </measurement-context>
</blood-glucose>
```

into the corresponding Open mHealth compliant data schema (e.g. the Open mHealth JSON object) shown here:

```
{
  "blood_glucose": {
    "unit": "mg/dL",
    "value": 120
  },
  "effective_time_frame": {
    "time_interval": {
      "start_date_time": "2013-02-05T07:25:00Z",
      "end_date_time": "2013-06-05T07:25:00Z"
    }
  },
  "blood_specimen_type": "plasma",
  "temporal_relationship_to_meal": "fasting",
  "temporal_relationship_to_sleep": "on waking",
  "descriptive_statistic": "minimum",
  "user_notes": "feeling fine"
}
```

However, some fields would not be completed if HealthVault did not have the corresponding data or does not provide it in their API properly associated with the glucose reading. Unlike Human API, all data is not centrally stored, so devices and apps can directly offer data to interested users via the Open mHealth API. Currently Linq, an initial proof-of-concept, is being piloted at Stanford's Preventive Cardiology Clinic.[41]

Despite, and to some degree because of these challenges, technology to empower and remotely monitor, educate and even treat patients is one of the most dynamic, innovative and rapidly growing domains within health informatics. As a result, no book can be reliably up-to-date in this area, so I encourage you to take advantage of the many resources on the Internet to keep up with it.

Summary: We've reviewed the many ways that technologies are empowering patients to become more active participants in their healthcare team and the healthcare decisions that will affect them personally. We've seen the progression from portals, which are provider office-based tools, to PHRs that patients can use independently of their providers. We've seen that patients increasingly can measure even relatively sophisticated parameters about their health and wellness at home.

[41] http://www.linqhealth.co/.

Patients can now bring together data from many sources, including their medical record, and manipulate it using apps. Physician visits are even being delivered virtually to patients using devices such as smartphones and tablets.

However, for the most part with respect to when these patient-facing technologies intersect with medical practice, we've still had a "one patient at a time" model in mind. We turn now to a topic that is also rapidly gaining visibility and attracting new technologies—helping physicians proactively, effectively and efficiently manage their entire practice population on a more continuous basis.

Population and Public Health

When I sit in a Ugandan village looking into the eyes of a man suffering from tuberculosis and AIDS, I cannot be distracted by any other issue. I am there to understand his medical situation and do what I can. I am a doctor, practicing medicine, focused on one patient's needs. But when I lean to the side and see a long line of other patients waiting to see me, when I wonder why so many people are sick with these same two diseases and how we could improve prevention, then I am doing public health.

— Chris Whalen, MD Epidemiologist, UGA College of Public Health[1]

So far in this book we've been looking at technologies and issues almost exclusively through the lens of the traditional one-patient-at-a-time approach to care delivery. It is likely that this will always be the primary focus of physicians in practice. However, a different approach is needed for patients who have multiple chronic diseases and for understanding the source and possible approaches to preventing disease across entire populations. Success in managing chronic disease requires engagement on a far more continuous basis than is possible under the traditional care model. Public health, by definition, is taking a broad look, as eloquently stated in the quote that introduces this section, at why people get sick and what might be done to reduce the incidence of disease in a population. In both of these approaches, informatics has a key and essential role to play by aggregating and reporting data in ways not done in classic medical practice and that was usually not anticipated, and is therefore not supported, in traditional EHRs. Next we discuss **popHealth**, a system that *was* designed to support data aggregation and reporting for population and public health.

popHealth: This is an open source Quality Measure Reference Implementation supported by ONC. Figure 1 presents the basic system architecture. The key design characteristic is that queries are run against data in each provider's EHR, as shown in the center of the diagram, without that data ever leaving the provider's control. The advantage of this approach is that providers are more likely to participate because of lessened concerns about loss of control of their data and, in particular, how it might be used and how that analysis might portray them. Moreover, data security is simplified and concerns about it are reduced since protected health

[1]Unpublished communication.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_11

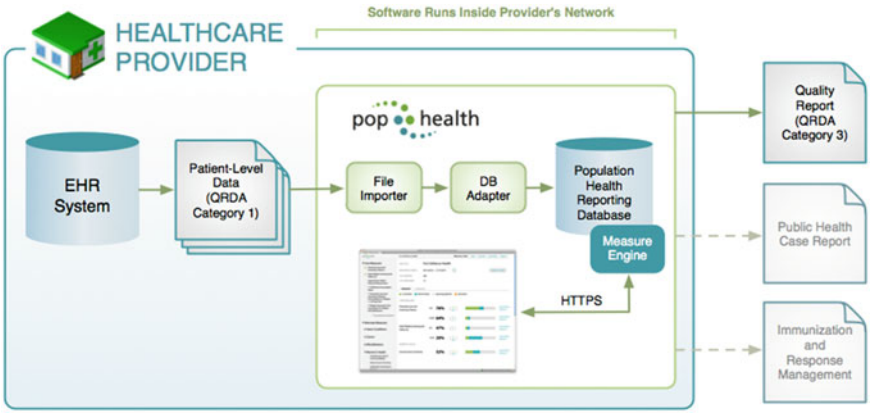


Figure 1: This illustrates the basic popHealth system architecture. The key design characteristic is that queries are run within the provider's firewall so protected health information (PHI) never leaves their control. (Source ONC)

information need not be transmitted from the provider to some central system. A major disadvantage of this "federated" approach is the need to deal with the specifics of each EHR (essentially the curly braces problem we encountered earlier with the Arden syntax) using software installed on site and the resulting need to update that software should the EHR design change.

"Data lockers" is an alternative approach we'll look at in more detail later that avoids some of these issues. Here data is stored centrally, but providers still control access to their own data, decides what queries warrant response, and may even be able to review specific query results based on their data before release. This has many of the same advantages with respect to encouraging provider participation. The review of query results may seem strange at first, but individual providers and health systems often worry that they will look bad if their data is used for comparative purposes. However, even though this may lead to certain entities opting out of specific queries, some participation is better than none.

As shown in the upper rightmost box in Figure 1, popHealth query results may be reported using HL7's Quality Reporting Document Architecture (QRDA) standard for quality reporting. QRDA reports are not patient-specific, protected clinical data but, rather are de-identified statistics at one of three levels of detail, called QRDA categories:

- **Category I (Patient-level) Reports**: Data for *one patient* for one or more CQMs
- **Category II (Patient-list) Reports**: Data for a *set of patients* for one or more clinical quality measures
- **Category III (Aggregate-level) Reports**: Aggregate data for *one provider* for one or more CQMs

popHealth aggregates and summarizes quality metrics from individual QRDA files. Figure 2 summarizes 500 patients from a 10-provider practice (in which individual physicians could be using different EHRs). This practice might be under an outcomes-based contract where revenue is tied to meeting goals for these quality metrics. Here, they're doing well on smoking screening, but not so well on weight screening and follow-up in the 18–64 year old subpopulation. The selection of a subpopulation illustrates the need to identify the right group of patients for each quality metric if the resulting statistics are to be clinically meaningful. For example, you would not normally perform screening mammograms for breast cancer on males and the recommendations for this screening are increasingly pointing to a specific age range for women. Note again in Figure 2 that, to highlight this key issue, the numerator—typically the number of patients whose care met criteria—is in green for each metric, while the denominator—the applicable target patient group for the metric—is in blue.

Where a practice-level quality metric is low, the next step would typically be to drill down to the individual providers to see who might be the source of the problem, so popHealth also provides data at the provider level. Here, comparative performance among providers for a given metric can be a powerful tool to incent

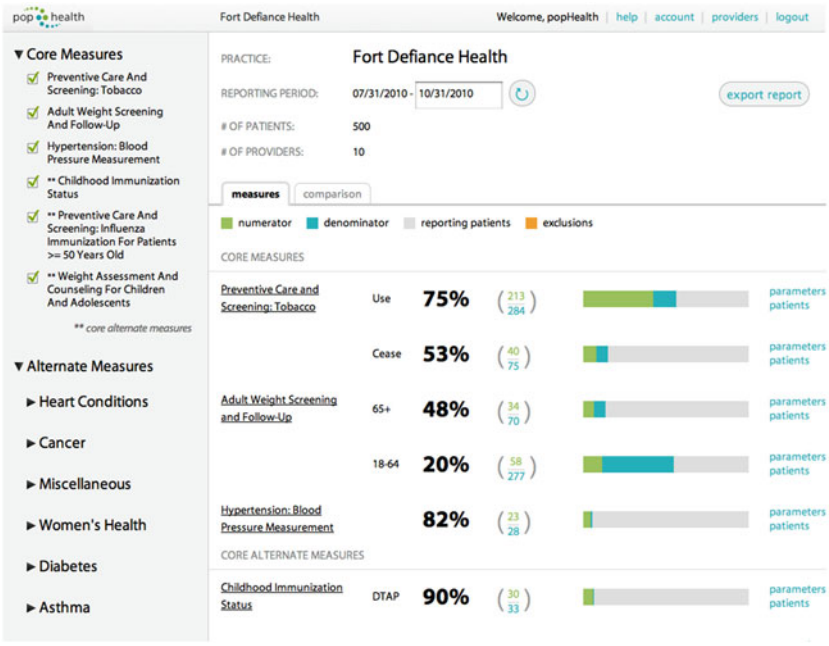


Figure 2: popHealth aggregates and summarizes quality metrics for 500 patients from a 10-provider practice (in which individual physicians could be using different EHRs). The numerator (the number of patients meeting each criterion) is clearly identified using the color green while the denominator (the target subpopulation for this metric) is in blue. (Source ONC)

improvement among those whose performance lags. Even though the data is collected at the population level, the results can be also be aggregated and reported for individual patients. Such a report might be reviewed as part of a patient visit and makes the patient's status with respect to recommended screening and preventive care easy to encompass at a glance. Care outside of guidelines is flagged to make that harder to overlook.

Quality Health First (QHF): Earlier we discussed the Indiana Health Information Exchange as the premier example of centralized HIE in the U.S. popHealth is designed to deal with the reality that, in most other places, there may be no effective HIE in place. You'd expect that IHIE could do more with their relatively facile access to a rich storehouse of data, and they do.

QHF, IHIE's population health service, can support more sophisticated searches than would normally be found in a distributed query framework using a data model simpler than IHIE's. As a result, a wider variety of patient subsets can be analyzed across a larger clinical dataset for all participating practice groups.

QHF also provides a dramatic example of the potential to add transparency to healthcare through public reporting of quality metrics. We've previously discussed the difference between process and outcomes measures, using HbA1C as the exemplar. The process measure is whether this test is done periodically according to guidelines and its actual value is a measure of the outcome of diabetes care. Patients with diabetes may well want to know which providers care for it properly, but how can they find that information? In general, there is no good answer but in Indiana, many providers voluntarily allow posting, at the practice level, of both HbA1C-based process and outcome measures. Each practice is also compared to state and regional averages. Some "practices" are solo so that data is actually at the physician level.

Figure 3 shows the QHF *outcome* metric report based on an HbA1C of 9 percent or less. It reveals that seven practices are below the state average (red line) and five are below their regional average (green line). This is only a part of this report and only one of a large number of quality metrics that are posted on the QHF site.[2]

Of course, simple statistics alone can be misleading or can present an incomplete picture. The commonly cited example of this is that the outcome of diabetes care can be affected by the mix of patients an individual practice sees. The term for this is risk adjustment. Social, economic and even population racial, ethnic and other factors might influence the severity of the disease providers see and might affect their success in managing it. To help with this, QHF breaks down their reports by regions of the state and patients can compare individual practices in their region or area where, presumably, these practice-level differences in patient mix are less pronounced.

An interesting new approach to population health and even more traditional care delivery is offered by the addition of a rich set of geocoded data to the EHR itself. Duke University has emerged as a leader in this area.[3] The new data can be

[2] http://www.ihie.org/public-reporting.

[3] https://www.dtmi.duke.edu/news-publications/is-geomedicine-reaching-a-tipping-point.

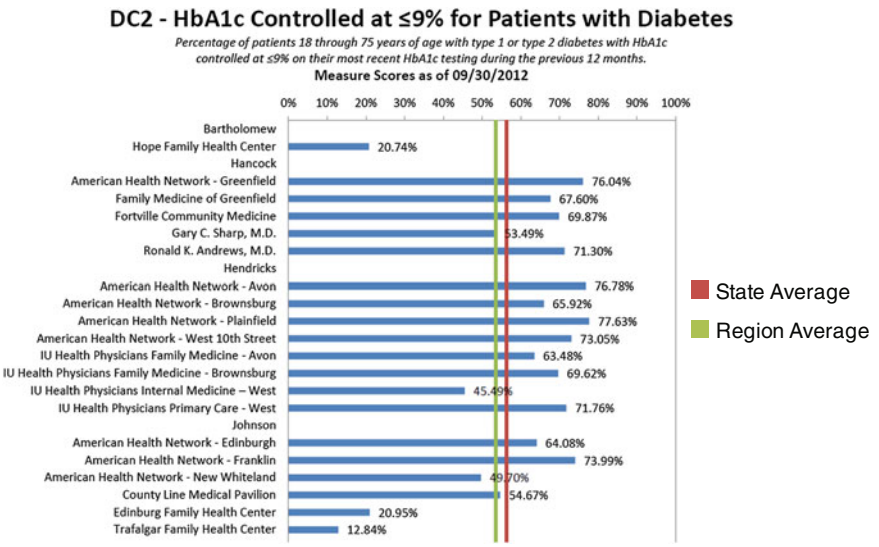


Figure 3: Public QHF reporting of diabetes care outcomes based on HbA1C level shows seven practices are below the state average (red line) while five are below their regional average (green line). (Courtesy © IHIE)

obtained from a variety of public sources (the United States Postal Service and the United States Census Bureau) and private sources (global positioning system companies). Once linked accurately to each patient's physical address, this data can provide a rich set of indicators for many purposes, including risk-adjusting quality and outcome metrics.[4] It also offers the potential for identifying patient cohorts at increased risk for disease or cohorts which might present increased management challenges based on their social, economic and environmental factors. For example, living near a pollution source such as a major highway may exacerbate asthma and complicate its management in children.

Case Study: Wellcentive: With the growth in outcome-based contracting by insurance companies, Medicare and major employers, commercial population health management tools and systems are available for use by physician practices and health systems. The typical application is to measure and monitor the performance of a group of physicians who have contracted collectively to cost effectively provide an acceptable level of defined quality metrics. Wellcentive is one of the larger providers of these services, but certainly not the only one. Rather than repeat what we've already discussed, we'll consider some features of their reporting that we've not yet seen. For example, a Wellcentive report of overall performance for a list of practice-defined alerts could include metrics specific to particular pay-for-performance or outcome-based contracts. These would require the collection and aggregation of data not specified under Meaningful Use.

[4] http://www.ncbi.nlm.nih.gov/pubmed/22160817.

Figure 4 is an interesting example of Wellcentive's use of visual analytics and illustrates the use of a broader dataset than is traditional in quality reporting. The tool is written in JavaScript so it can run on a variety of computers, including mobile and tablet devices. It can be configured to show many metrics in multiple visualizations on the same dashboard or report. When this is done, they are "connected" so the user can highlight portions of one visualization and see the corresponding portion in each visualization, but with different metrics or dimensions.

In this case, we're seeing only one graphic that shows providers (each box is a provider) based on the percentage of their diabetic patients that are out of control (using the general standard of an HbA1c of greater than 9 percent) and by the percentage of diabetic patients in their practice but not seen in the prior year. The shading of each box correlates with the percentage of patients out of control, with darker indicating poorer performance. The size of each box correlates with the percentage of diabetic patients not seen in the past year, with a larger size indicating fewer patients with annual visits. Thus, the providers in the upper left corner who have the smallest and lightest boxes are, in the traditional sense, the top performers.

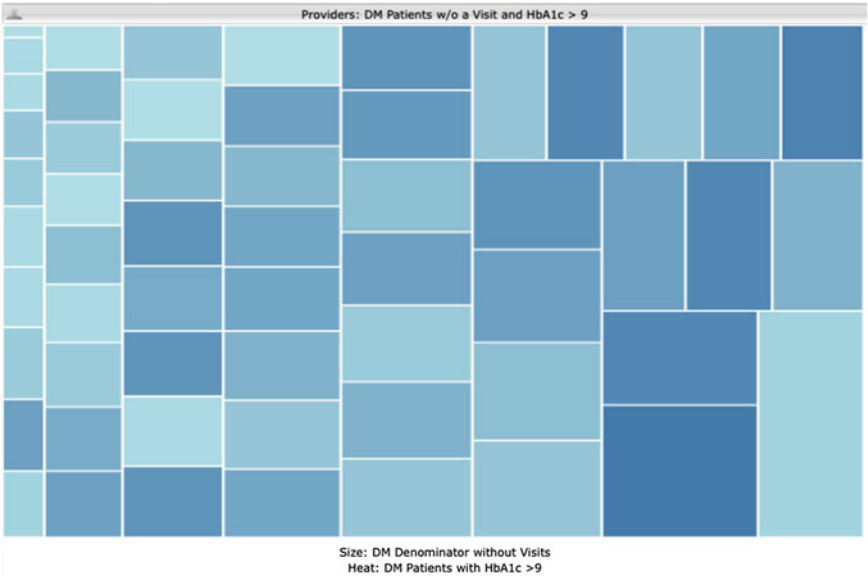


Figure 4: A Wellcentive visual analytics report (in a format called a "treemap") helps find the most cost-effective providers of diabetes care. Each provider is represented by a box with a darker box representing poorer control of that provider's diabetes population and a smaller box representing a higher percentage of annual visits for those patients. The providers at the upper left have the best control and also see their patients at least annually. The provider at the lower right has a virtually equal degree of control with far less frequent visits and may deliver the most cost-effective care. A further analysis of this provider's care process for diabetes would be worth doing to find approaches other, less cost-effective, providers might adopt. (Courtesy Wellcentive, All Rights Reserved)

They see their diabetic patients annually and have them under good control. However, the provider in the lower right corner is very interesting. This practice has a well-controlled diabetes patient population despite not regularly seeing their diabetic patients in the office. It would be interesting to compare this practice's care processes with those in the upper left corner. Perhaps, for example, the practice at the lower right is using remote monitoring and only sees patients who are getting out of control while not wasting time and money seeing those it knows to be in good control. This provider may, in fact, be the most cost effective and, hence, the most profitable practice under an outcome-based reimbursement model and their diabetic care process might be something the others should adopt. An analysis of this kind presents a more complex data aggregation and curation problem than is normally the case in quality reporting (which is typically based on a small and well-defined set of quality metrics) since it requires aggregating, normalizing and standardizing claims, EHR *and* clinical lab data.

Public Health: The use of healthcare data for public health research could easily be the subject of an entire book, but, from an informatics perspective, it resembles population health because it requires similar data aggregation and analysis technologies and approaches. However, public health takes a broad view of the factors determining health and disease, so it uses a more disparate set of data sources to understand the impact of more factors and determinants of health and disease.

We will next be looking in more detail at the technologies for aggregating data from diverse EHRs, but the Biosense 2.0 system is an interesting example that was more specifically designed by the CDC for public health queries.[5] An often-cited example of its use is the Tarrant County Public Health Department (TCPH) in Fort Worth, TX, which has used Biosense 2.0 to collect and analyze data from 60 regional hospitals for syndromic surveillance, monitoring for disease outbreaks, which is a key mission for public health.[6] At TCPH, the system can provide a time-ordered display of the number of visits (typically these would be emergency department visits) for a particular clinical problem. The ability to produce this in a timely manner across an entire region is a powerful surveillance tool that can, at least in theory, be near real time, and therefore facilitate much faster public health response to a potential problem. Achieving this ability for a possible bioterrorism attack or an infectious disease outbreak, where timely response would be particularly important, is a key goal of the Biosense 2.0 effort.

In the final section we'll discuss the role that combining geocoded data (tied to a specific location) and clinical data could play in providing new insights about disease causation and management. Public health is another domain in which this is increasingly important. One example is the Public Health Disparities Geocoding Project at the Harvard School of Public Health that is geocoding public health surveillance data and using census-derived, area-based socioeconomic measures to

[5] http://www.cdc.gov/biosense/biosense20.html.

[6] https://sites.google.com/site/biosenseredesign/community-forum/biosense20intexasastatusreport.

monitor socioeconomic inequalities and their contribution to racial/ethnic and gender inequalities in health with the goal of increased visibility of these issues.[7]

We've now seen some of the potential uses of data when it is aggregated across EHRs (and from other sources) to manage entire patient populations against defined quality, population health management or public health reporting objectives. We now turn to the technical details of how this is actually done.

[7] http://www.hsph.harvard.edu/thegeocodingproject/.

Aggregating Data

In many cases, aggregating individual data sets into big-data algorithms is the best source for evidence as nuances in subpopulations (such as the presence of patients with gluten allergies) may be rare enough that individual smaller data sets do not provide enough evidence to determine that statistical differences are present.

— The "Big Data" Revolution in Healthcare, McKinsey, 2013[1]

Here in the U.S., physicians who aren't employed by or affiliated with a health system that makes the choice for them, can select any EHR—and there are literally hundreds from which to choose. Unless something like the JASON recommendations are mandated or an alternate interoperability framework such as CommonWell is universally implemented, these EHRs will not be sufficiently interoperable. This is problematic for aggregating data from them for care coordination and for other so-called "secondary uses" of EHR data, including research and population and public health.

For now, solutions have been proposed and implemented for the aggregation of a more limited set of quality metrics and some other well-defined datasets. The technical description for the data aggregation problem in the environment we currently have here in the U.S. is "distributed query in a federated environment." This means securely obtaining useful data from diverse EHR systems and other sources without some national interoperability framework in place and while the source data typically remains at those sources. To explore this we'll look at three current open-source, distributed query technologies. Those designed for population and public health are similar and, for the most part, they simplify things by focusing on a small and well-standardized set of quality metrics. As we saw with the Wellcentive Case Study, having access to more extensive data can lead to other insights such as the most cost-effective approach to care. This more advanced analysis usually requires clinical detail that is not provided by quality metrics.

[1] http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/healthcare%20systems%20and%20services/pdfs/the_big_data_revolution_in_healthcare.ashx.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_12

It may also require that this clinical data is aggregated and combined with other data, such as the cost of care that is typically derived from claims.

The distributed query effort is now under the ONC supported Query Health Initiative (QHI).[2] The technological approach is evolving so the details you find, should you visit the QHI site, may be different from what is discussed here; however, these examples should provide a good basic understanding of the issues.

hQuery: The primary goal for this system is simplicity which is achieved, in part, by working with a limited data set that is well-standardized because of its key role in quality measurement and other common clinical activities. As shown in Figure 1, hQuery provides an attractive, modern point-and-click interface (called the hQuery composer) to the query builder who might even be a nontechnical healthcare provider. At the same time, the queries that can be formulated are somewhat limited. In addition to the query builder, hQuery uses a simple patient information model to further facilitate query building by nontechnical users (and to simplify the data aggregation task). Data for possible reporting are forwarded from the source system to an hQuery Gateway, software that is often co-located with each source system. The gateway also receives queries in a standard format and forwards them to another piece of software called the adapter that knows how to translate them for the source system. Once again, this is an interoperability solution that deals, in part, with the curly braces problem of mapping a standard specification of a query into the specific data model of a particular EHR. The query results from the EHR might be a CCDA document that the adapter can convert to the hQuery standard data model for transmission back. I provide a reference to a useful video that illustrates more of how this is actually done.[3] The final, aggregated results are attractively presented and can include frequency, time and geographic distributions.

i2b2: Informatics for Integrating Biology and the Bedside (i2b2) is an NIH-funded effort based at Partners HealthCare System in Boston. Its mission is to enable clinical investigators, even if they are located at multiple sites, to conduct research using state-of-the-art genomics and biomedical informatics. i2b2 is necessarily far more sophisticated and complex than hQuery to support the much larger set of questions that could be posed in research environments. However, it still provides an easily understood database schema that can be used to create a data warehouse where large clinical abstracts from an enterprise EHR can be stored for future analysis. This is important because the database schema of commercial enterprise EHRs can be dauntingly complex, proprietary and not necessarily designed to support ad hoc queries.

i2b2 implementations consist of modular software elements called “cells” that communicate via web services and, as a group, are called a “hive.” Some cells are required while others are optional. It is also possible to develop and connect custom cells to the hive. The role of most cells and their associated web services is clear

[2] http://wiki.siframework.org/Query+Health+-+Project+Charter.

[3] http://wiki.siframework.org/Query+Health+-+Summer+Concert+Series#x-hQuery.

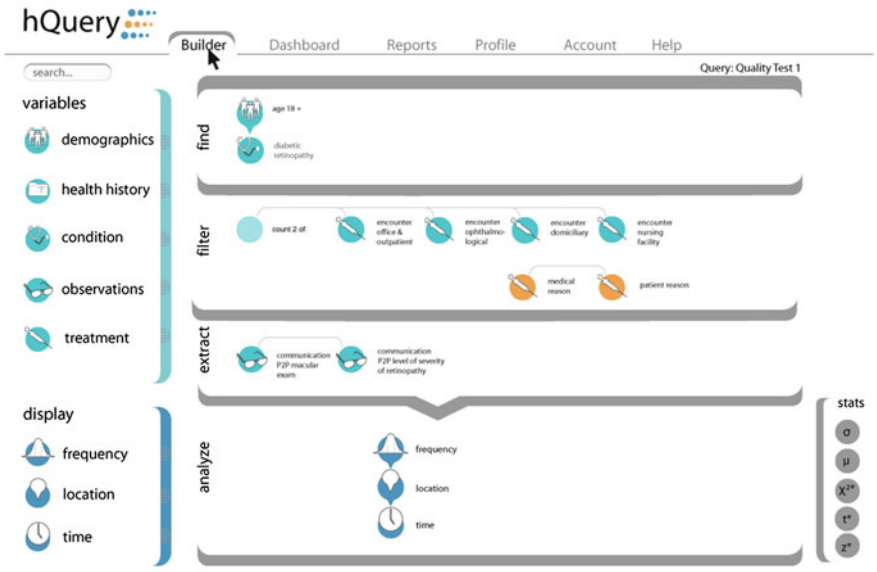


Figure 1: hQuery provides a simple point-and-click user interface that is designed so that nontechnical users, such as providers, can initiate queries. (Courtesy Gregorowicz and Hadley, ONC) (http://wiki.siframework.org/file/view/hQuery+Summer+Concert+Presentation.pdf)

from their name. The specifics probably aren't important to most readers, but if they are to you, there is an interactive page on the i2b2 site you can visit to further explore the cells.[4] i2b2 can be used to weld together a multi-institutional research collaborative. A good example is the University of California Research eXchange involving five medical schools in the statewide university system.[5] There is an organized i2b2 users' group to share information and applications and to create the potential for translational research through collaboration and federated queries across institutions.

PopMedNet: It is often desirable to bring together data from many sources and providers, but those data sources may be cautious about participating because of concerns about how their data will be used and whether their institution might be portrayed negatively in some comparative analyses. Earlier, we mentioned the concept of "data lockers" as a potential solution. PopMedNet implements this architecture and allows each source to maintain control of its data, with the capability of accepting and responding to queries that might go to many data sources. It is intended to support medical product safety analysis based on aggregated reporting of complications from multiple institutions as well as the comparative effectiveness of alternate treatments, again resting on data from multiple sources

[4] https://www.i2b2.org/resrcs/hive.html.

[5] http://www.slideshare.net/CTSIatUCSF/uc-braid-ucrex.

and other studies. We discussed the CDC's Biosense 2.0 system to show how data lockers have been effective in making public health departments from many states comfortable with contributing their data to the project. By all accounts, this strategy has been effective.

Query Standards: We mentioned Query Health earlier as an umbrella effort organized by ONC. It has proposed an architecture that utilizes and brings together all three of the technologies we've just reviewed into one query framework in order to facilitate cross-platform queries. The technical vehicle for that is distributed query standards.

There are four kinds of query standards, each of which specifies a different key element of distributed query:

- **Envelope** standards define the packaging for sending/receiving queries.
- **Format** standards provide a *Declarative* specification of the query (we'll explain what that means next).
- **Results** standards, such as QRDA that we discussed earlier, specify the format and packaging of the query results.
- **Data Model** standards, such as the Clinical Element Data Dictionary (CEDD), specify the common model that will link data from the contributing systems to a standard (much as FHIR is trying to do).

Before discussing the standards, we need to differentiate between a declarative and a procedural specification using baking a cake as an example. The declarative statement (e.g. specifying the object: "a baked cake") expresses the desired result but does not explain how to achieve it. The procedural statement (e.g. the recipe) describes how to achieve the goal. Think back to the curly braces problem with the Arden syntax: it's really the same thing. Arden provides a declarative standard, but the procedure will be different depending on the system that is the data source. Similarly, the hQuery Gateway receives a standard query on one side and, on the other side knows how to execute it against a system's internal data model. The "recipe" will depend on the specific EHR from which the data is to be retrieved.

The Query Envelope standard serves to provide identification that is unique within the network: the information requestor identification including name, e–mail and organization; the purpose and priority of the query using one of seven purpose codes (e.g. TREAT); its **priority** from 1–5 (1 highest); its **type** (1–20 characters); its **PHI level** (aggregated, limited, de-identified, PHI); and its **timing**, **submission date/time** and **optional execution date/time**.

Queries are specified in the Health Quality Measure Format, an HL7 standard for the content and structure of a quality measure in an XML document based on the HL7 RIM. It can consist of three levels of detail:

- **Metadata** such as who wrote the measure, the dates over which it is valid, who validated it and other details about how the measure works or is used
- **Human narrative** including measure description, data criteria, measure population and measure observations

- **Computer instructions** about how to count and compute the results of the measure

We introduced QRDA earlier—the standard for query results that can be reported at three levels of detail: patient-level, patient-list level or aggregate level.

CEDD is a tool implementers can use to set up their source data in support of distributed queries within a larger Query Health solution. This is not intended as a new standards development effort and began with the elements already specified in the CCD and which most EMRs already support, since they were required for Stage 1 certification.

Any set of distributed query tools or frameworks that implements these standards could, in theory, engage in the aggregation of data taken from EHRs connected to another cooperating framework. That's the basic goal of the Query Health framework.

Summary: In this section, we examined how data from multiple EHR systems can be queried and aggregated for diverse purposes from quality reporting to advanced clinical research. All of the technologies to do this serve the essential role of providing a framework over the many noninteroperable EHR systems deployed in our current healthcare system. Taken together, they illustrate different technological solutions, each of which is optimized for a specific problem in a specific cross-institutional context. In theory, should the JASON report be fully implemented, a nationwide interoperability framework based on web services might become a more universal approach to query and data aggregation, but that is likely years off, if it ever happens.

We've now covered the traditional spectrum of health informatics systems and tools. They range from electronic record systems for providers and patients to systems for sharing the data those records contain. They also now include systems for querying those systems and aggregating data for far-reaching purposes. In the final section, we'll see just how interesting the results can be when enough data from multiple sources is brought together for analysis in what is arguably the most dynamic area of health informatics today—big data and analytics.

Health Big Data and Analytics

During the past half-century, there has been an explosion of biomedical and clinical knowledge, with even more dazzling clinical capabilities just over the horizon. However, the systems by which health care providers are trained, deployed, paid and updated cannot usefully digest this deluge of information.

— Best Care at Lower Cost: The Path to Continuously Learning Health Care in America, Institute of Medicine, 2013[1]

As the quote from the IOM conveys, our ability to generate scientific knowledge has outstripped the ability of healthcare providers to absorb, integrate and use that knowledge in daily patient care. The ability of information systems to analyze huge amounts of clinical data and provide useful insights and guidance is a major focus for research and commercial development now that the substantial majority of healthcare providers are using electronic records that have the potential to provide the data for that analysis. This feedback loop is a key part of the IOM's vision of a learning healthcare system, as indicated by the title of the book from which the quote is taken. In this section, we'll explore early results from these exciting and potentially transformational analytic technologies.

First, relax. This section does not seek to teach you how to do analytics. Even though it's a highly mathematical process, you won't run into any equations. Instead, we will 1) develop a working definition of big data; 2) look at how exploring this data through analytics opens new ways of asking and answering healthcare questions; and then 3) take a futuristic tour of the many ways that models and simulations will be used to better understand and improve health quality, outcomes and efficiency.

Although there are many potentially relevant data sources (such as the geocoded data and data from personal sensors and apps we discussed earlier) we will primarily consider the end results that are being achieved by aggregating and then analyzing data that derives, at least in part, from digital health records. We'll see that we're clearly heading into a new era of medicine and healthcare delivery that

[1] http://www.nap.edu/catalog/13444/best-care-at-lower-cost-the-path-to-continuously-learning.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_13

will increasingly be powered by analyzing big data in new ways and for new purposes. Many things that are now being done by humans, including physicians, will in the future (and not necessarily the distant future) be done by, or at least assisted by, computers. This may sound scary, but it shouldn't be. This section is intended to give you at least a feel for what that future might look like.

What Is Big Data? Almost all devices now have a digital "brain" and they are increasingly connected to the Internet. This is termed the "Internet of Things" (IoT) and is one of the hottest contemporary topics in technology. As a result of the growth in digital devices and their connectivity which makes obtaining data much easier and virtually free, the world is awash in data—and it's growing at exponential rates. People have coined the term "big data" to refer to this phenomenon but can't quite agree on what that means. What separates big data from something at lesser scale? César A. Hidalgo of MIT's Media Lab says that, to qualify for the big data distinction, it must be big in **size**, **resolution** and **scope**.[2]

Let's reframe this idea in a way that is more directly relevant to transforming healthcare delivery systems. To accomplish that, the data must represent many patients and providers, must do so in detail and must be combined with other data to give the context within which care is delivered and the external rules and policies within which that delivery system must operate. Looked at a bit differently, the data must be sufficient to represent the behavior of the *complex adaptive system* of healthcare. A complex adaptive system can be briefly described as consisting of multiple independent agents, each of which is acting in its own perceived self-interest and where no entity is ultimately in charge. Moreover, if conditions (e.g. financial incentives for example) change, the individual entities will adjust their behavior (adapt) to maximize that perceived self-interest. Often-cited examples include education and healthcare.[3]

An interesting example of big data comes from work being done at Duke University to build an integrated clinical data warehouse by combining millions of patient records from their EHR with a very large set of geographic information system data.[4] Figure 1 illustrates some of the capabilities and potential insights that can be obtained by doing this. Once the patients' addresses are normalized (standardized) using a service from the United States Postal Service, their clinical data can be correlated with ethnic, demographic and economic factors, such as access to fast-food restaurants or parks. This is in keeping with the IOM's recent report, *Capturing Social and Behavioral Domains in Electronic Health Records: Phase 1*

[2]http://enterprise.huawei.com/en/about/e-journal/ict/detail/hw-195167.htm.

[3]https://www.nae.edu/File.aspx?id=7417.

[4]Rusincovitch SA, et al. Design and Implementation of an Automated Geocoding Infrastructure for the Duke Medicine Enterprise Data Warehouse. AMIA 2014 Joint Summits on Translational Science. April 7-11, 2014, Abstract: Podium presentation.

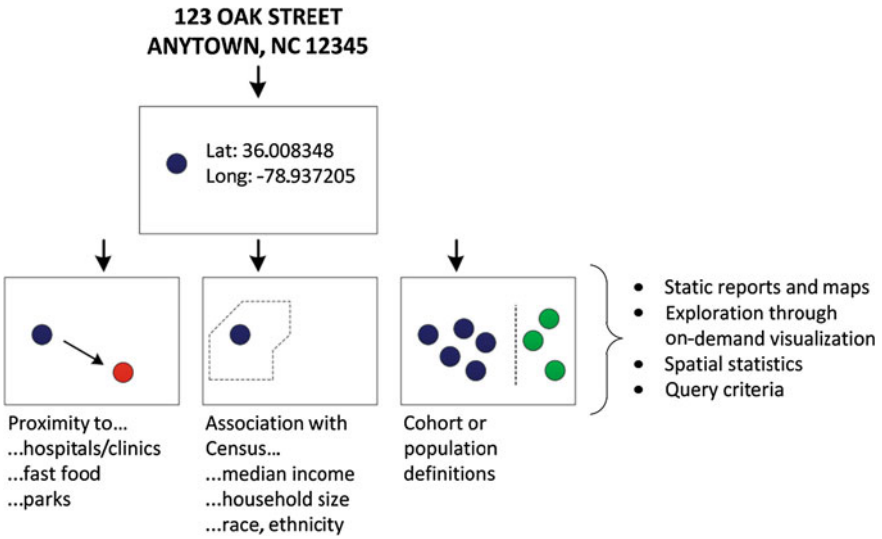


Figure 1: The Duke University effort to create "big data" by combining clinical EHR data with Geographic Information System data has the potential to reveal the social determinants of health. (Courtesy Shelley Rusincovitch, Duke University)

which says: "A number of domains of special relevance to the social determinants of health can be characterized by use of the patient's residential address." [5]

From Data to Insights: The controlled clinical trial is the historic gold standard for developing new medical knowledge. In this type of study, two randomly selected groups of patients are provided alternate therapies (or a new therapy and a placebo) and researchers use statistical analysis to see which group, on average, does better over time as defined by clinical outcomes, cost or, increasingly, both. While controlled clinical trials are familiar and effective, they are also difficult and expensive to accomplish. Finding and recruiting the right patients is hard. The trials take a long time and require expert curation. However, if the question is which of two or more treatments works best in actual human patients, there is currently not a good alternative approach.

However, let's now consider a different question: What is the *optimal* treatment strategy among already known and available options? Or, if a particular chronic care model were used in a specific context, what would the outcomes be on clinical quality and cost? As opposed to the classic research questions we illustrated in the discussion of clinical trials, determining optimal treatment would require many alternative experiments, that is, trying every possible treatment strategy on many different cohorts of similar patients. For all the reasons we just listed, this would generally be prohibitively time consuming and costly. To determine the cost effectiveness of alternative chronic disease care models, an experiment could be

[5] http://www.iom.edu/Reports/2014/Capturing-Social-and-Behavioral-Domains-in-Electronic-Health-Records-Phase-1.aspx.

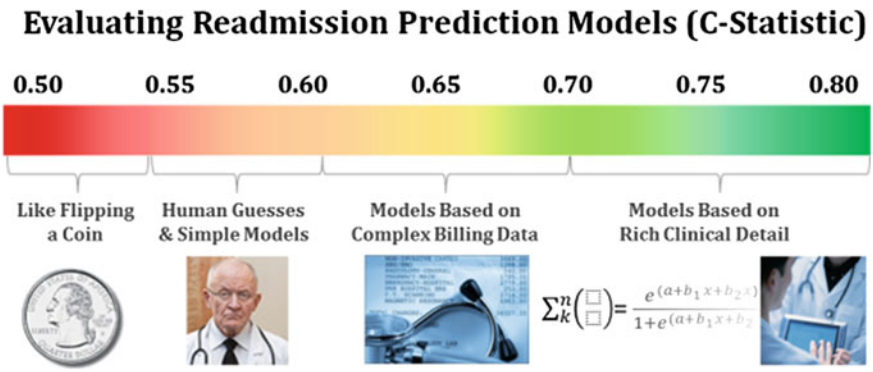


Figure 2: Research shows that, when faced with a complex multidimensional predictive decision, such as which patients will be readmitted, even skilled clinicians, case managers and simple models on average do little better than a coin flip. Models based on rich variety of clinical data often do much better. (Courtesy John D'Amore, Diameter Health, All Rights Reserved)

conducted by changing the clinical process in half of a clinic and comparing it to the other half that is operated as before. However, this would often also be too costly and complex to do. These are the types of questions best answered through modeling and simulation techniques that are increasingly possible because of the growing availability of digital health data. There are a number of techniques for modeling and simulation. Since this is not an analytics text, we'll briefly describe them, but not go into the details of how they are performed.

First, you might be wondering what role is played by models and simulations can play in making the routine delivery of healthcare better. This question was directly addressed by John D'Amore, President and Chief Technology Officer of Diameter Health, a Boston-based provider of software services including data aggregation and analysis for managing population health and outcomes-based contracts. We mentioned this company earlier in our discussion of using analytics to suggest missing entries in a patient's problem list.

In an interesting blog post, he used the prediction of hospital readmissions as an example to give some specific reasons for why analytics can improve healthcare by doing things that are difficult for clinicians:

> First, clinicians don't have easy access to the data needed to effectively predict readmissions. This requires a multivariate combination of laboratory, medication, past admission, vital sign, demographic and diagnosis data to do it well. Second, people are good at holding a few variables in working memory, like a telephone number, not the dozens of factors identified in most readmission models. Finally, who has the time? Most hospital nurses and physicians are already busy and don't have the spare time to start doing logistic regression every day for every patient discharge.[6]

[6] https://medtechboston.medstro.com/medtech-profiles-diameter-health-using-big-data-to-predict-readmission-risk/.

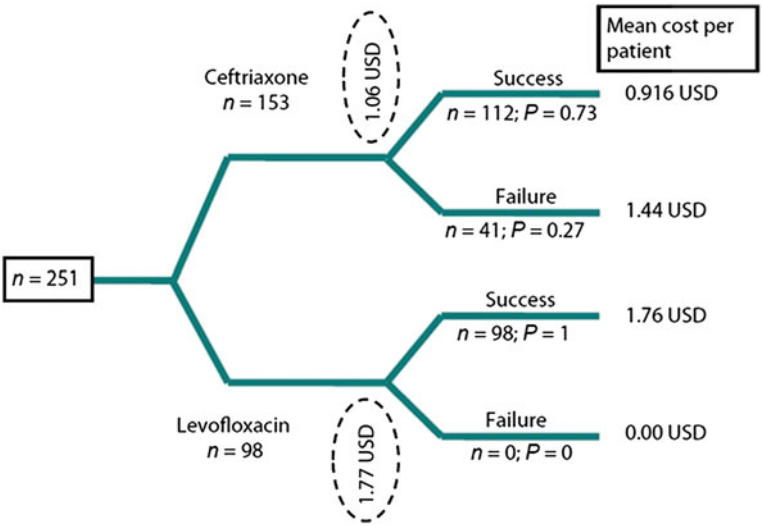


Figure 3: A decision tree for the most cost-effective choice of antibiotics illustrates that there is no recursion (going back to repeat a step) and time is not explicitly represented. (Courtesy S. Sriram, MPharm, PhD)

He further illustrates these points with the graphic shown in Figure 2—research shows that, when faced with a complex multidimensional predictive decision, clinicians, case managers and even simple models, on average, do little better than a coin flip.[7] You may recall that the ACOVE study showed that, in many instances, the care of the vulnerable elderly was optimal only around half of the time. This sheds some further insight onto that finding. As we'll see, models based on rich, multivariate clinical data often do much better.

Decision trees are directed graphs. There is no recursion; so travel through the graph is one-way. Time is not represented. Figure 3 is a generic clinical decision tree example where different antibiotic treatment choices yield different outcomes at different costs. Obviously, in such a case, there is no going back. Once a patient is given Drug A that decision can't be remade (although a later decision to switch to Drug B could be made). Also, there is no specification of time. Events happen when they occur and the model is indifferent to that, although it does show the sequence of events.[8]

Figure 4 shows a Markov Model (actually a semi Markov Model, but the distinction is unimportant for our purposes) used to compare the quality-adjusted survival and cost of three alternative drugs (ximelagatran, warfarin, and aspirin) to prevent clots in a hypothetical group of 70 year old atrial fibrillation patients with

[7] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3138589/.

[8] http://www.jrpp.net/article.asp?issn=2279-042X;year=2013;volume=2;issue=2;spage=70;epage=74;aulast=Sriram.

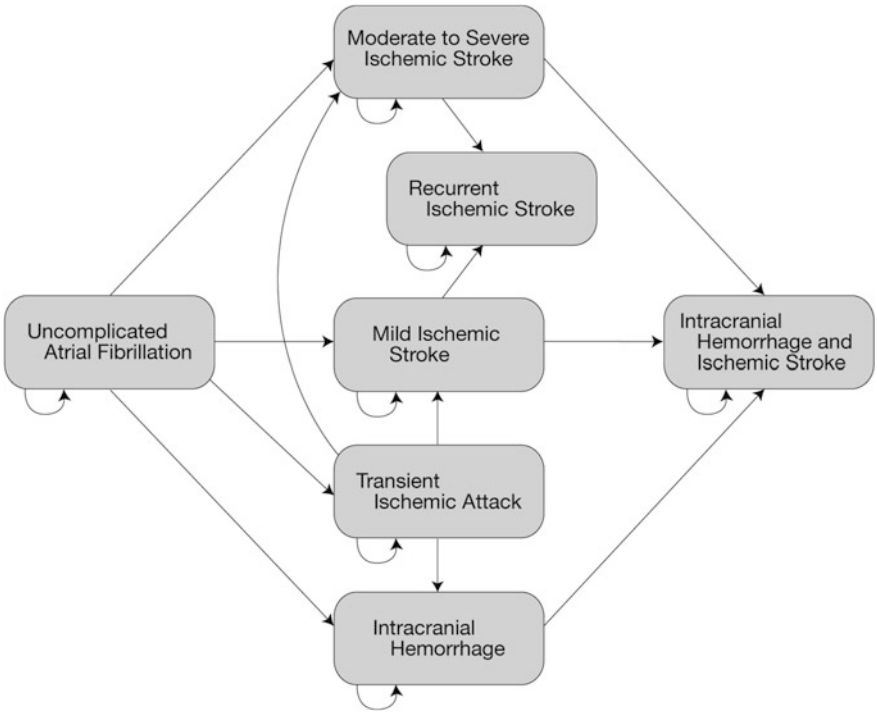


Figure 4: A Markov Model is used to compare quality-adjusted survival and cost of three alternative drugs (ximelagatran, warfarin, and aspirin) to prevent blood clots in a hypothetical group of 70 year old atrial fibrillation patients with varying risk of stroke, and no contraindications to anti-coagulation therapy. (Courtesy Dr. Cara O'Brien, Washington University School of Medicine)

varying risk of stroke, and no contraindications to anti-coagulation therapy. Atrial fibrillation is the most common cardiac arrhythmia (irregular heart beat) and it can lead to clots forming in the heart which can migrate to the brain and cause a stroke. The model starts out with uncomplicated atrial fibrillation, then cycles between health states until death occurs or a 20-year period ends. The health states are equivalent for each treatment but the probabilities, costs, and quality of life vary. Note that, unlike a decision tree, there are many paths through the model and it is possible to go back to a prior point depending on actions and probabilities.[9]

Discrete event simulation (DES) represents entities and their attributes as well as queues where they wait for something, typically because of finite resources. As shown in Figure 5, a classic healthcare example is an emergency department.[10] Anyone who has ever sought treatment in one probably knows all about queues!

[9] http://jama.jamanetwork.com/article.aspx?articleid=200332.

[10] http://www.anylogic.com/consulting/healthcare-and-pharmaceuticals.

Figure 5: A discrete event simulation of an emergency department represents the entities involved and their attributes. It models resource utilization and also has queues where agents wait for finite resources. (Courtesy AnyLogic, All Rights Reserved)

Agent-based simulation is richer in attributes than discrete event simulation, in part because agents can interact with each other and with their environment. A simulation of the Center for Health Discovery and Well Being (CHDWB), a novel clinic whose objective was to identify and mediate risks of future chronic disease, is an example done at Georgia Tech's Tennenbaum Institute.

Figure 6 is a diagrammatic representation of CHDWB including physical areas, such as the education and consultation rooms where trained coaches interact with clinic enrollees. When the simulation is actually run, all the clinic personnel and the enrollees being served move along pathways governed by the time it takes to do each element of the clinic's work, the available resources and other factors. The simulation was successful in helping identify a set of processes and a revenue model that would make the clinic self-sustaining with a return on investment to its parent organization.

Optimal Treatment of Depression: Analytics has the potential to help determine optimal treatment, which now typically means the best clinical outcomes at the lowest cost. These are, of course, not precisely definable terms. For example, there is a serious debate about issues such as the value of spending substantial sums on treatments that only prolong the life of terminal patients by small increments of time.

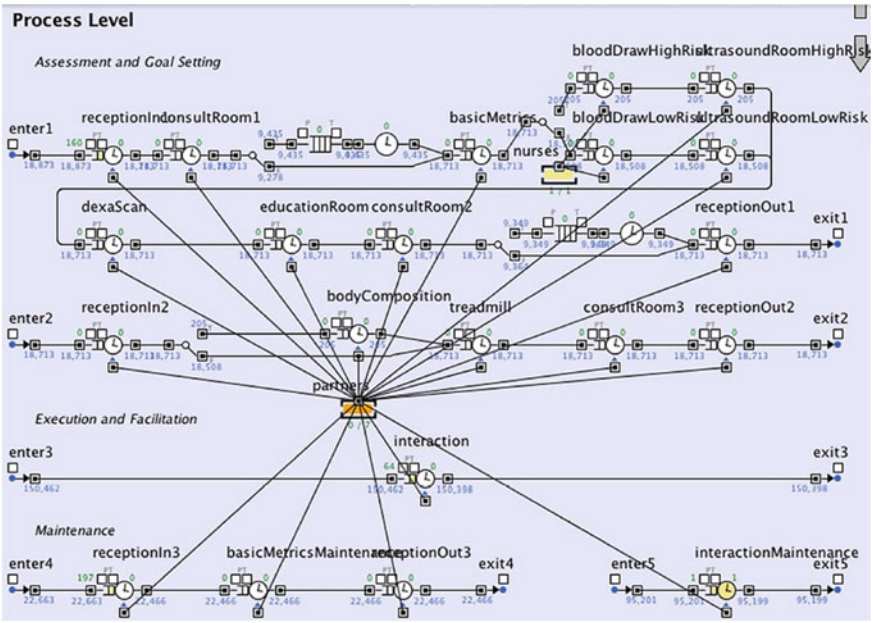


Figure 6: An agent-based model provides a detailed visualization of the operational processes at the CHDWB, a model clinic for wellness and disease prevention. (Courtesy Rahul Basole, Tennenbaum Institute, Georgia Tech)

Research at Indiana University illustrates the potential to help in less controversial clinical contexts.[11] A partially observable Markov Decision Process—a variation of the Markov Model that introduces memory—was developed and its recommendations were compared to the actual treatment of nearly 6,000 depressed patients abstracted from their electronic records. Under certain assumptions about treatment (and implicitly about reimbursement), the model delivered improvements that were better than or almost as good as real physicians *but at a better cost per unit of improvement*. As we seek to re-engineer the healthcare system, technologies like this may become a routine part of CDS for physicians. Some would even speculate that, at least in certain circumstances, they may replace physicians as the decision makers, but I leave it to you to consider that.

Early Diagnosis of Congestive Heart Failure: Earlier in this section we said that models based on rich, multivariate clinical data often do much better than clinicians who are using a more limited set of factors. We'll now discuss a specific example based on congestive heart failure (CHF) which is the single most expensive ICD-9 code. As a result, there is great interest in improving its management. Early on in CHF, the symptoms can be subtle, so it may not be diagnosed for some time even after they begin. However, since early treatment can forestall

[11] http://www.aiimjournal.com/article/S0933-3657(12)00151-0/pdf.

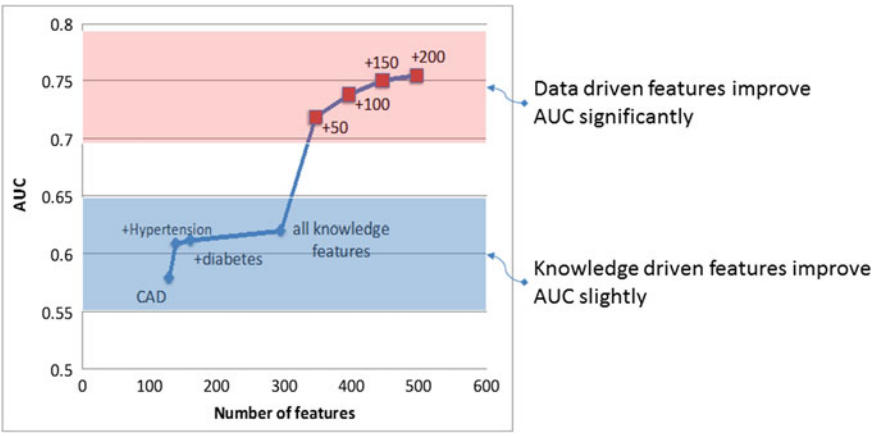


Figure 7: Adding data-driven clinical factors derived using analytics results in earlier diagnosis of CHF than using only knowledge-driven factors based on the data typically collected by physicians. Area under the curve (AUC) is a measure of how well a model will rank a randomly chosen positive instance (in this case the patient has CHF) higher than a randomly chosen negative example (the patient does not have CHF). In this example, the higher the AUC, the more accurate the model's clinical predictions. (Courtesy Jimeng Sun, Georgia Tech)

expensive complications, earlier diagnosis is preferable both from clinical and economic perspectives. To develop a model to help with earlier diagnosis, researchers first analyzed electronic patient records, including the parts that were free text such as patient-reported symptoms. They developed extraction algorithms that used the structured and text components to develop a set of up to 100,000 clinical "features." These were classified using logistic regression (a statistical technique commonly used to predict whether a patient has a condition based on characteristics of the patient) and random forest (a method that uses multiple methods to classify objects) to determine which features were predictive of CHF. As shown in Figure 7, the resulting model substantially improves the diagnosis of CHF, based solely on medical literature, even with as few as 50 added *data-driven features* with high predictive power. The figure also shows that the addition of most *knowledge-driven factors* (e.g. factors based on physician-collected data) only slightly improves the timeliness of diagnosis.[12]

Case Study: Jvion. As illustrated by the preceding example of earlier detection of CHF, it is possible to mine healthcare data and develop models that can predict clinical events with higher levels of accuracy and reliability than traditional techniques. As we also discussed earlier, a major reason for this is that these predictions can be based on far more extensive data than can usually be derived solely from an

[12]Sun, J et al., 2012. "Combining Knowledge and Data Driven Insights for Identifying Risk Factors Using Electronic Health Records." AMIA 901–10.

EHR, including a virtually unlimited number of variables related to individual patients, their communities and the general population. Earlier, we mentioned potential uses for geocoded data about a patient's environment, and this is another one.

In fact, according to *Big Data Analytics in Healthcare: Promise and Potential*,

> "Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data, such as length of stay (LOS); patients who will choose elective surgery; patients who likely will not benefit from surgery; complications; patients at risk for medical complications; patients at risk for sepsis, MRSA, C. difficile, or other hospital-acquired illness; illness/disease progression; patients at risk for advancement in disease states; causal factors of illness/disease progression; and possible co-morbid conditions."[13]

Jvion, a commercial supplier of clinical analytics to over 270 U.S. hospitals, has developed a predictive analytics platform with three core capabilities:

- **Clinical rules** are derived from evidence-based studies and are used to identify risk at the cohort level. Clinical rules are commonly included within EHRs and are effective at identifying textbook cases of illness or risk. However, they are limited in that the only way to update a clinical rule is by adding new rules and flags.
- **Statistical algorithms** are also built from evidence-based studies and used to flag potential risks within a patient population. LACE scores for 30-day readmissions[14] and Braden scores for pressure ulcers[15] are examples. The problem with these algorithms is that they never change. The same data will always produce the same results, regardless of variations in context or setting.
- **Machine learning** identifies patient risk through continuously learning models that become more accurate with each new data element fed into the system. When used alone to predict illness or adverse events, deep machine learning lacks the clinical context needed to account for evidence-based outcomes.

Jvion says that it uses all three of these capabilities to create a patient phenotype model to cluster patients in order to:

- Understand the inherent, imperceptible characteristics of a patient
- Appropriately apply clinical intelligence to support physician decision making
- Deliver a more accurate, scalable and flexible risk profile than is possible with any one predictive approach

Phenotyping is performed using a sophisticated mathematical approach to map patients into an abstract space (called an Eigenspace) in which each patient characteristic is represented, as shown in Figure 8. Jvion uses this representation to group clinically significant patient cohorts, relate them to outcomes and use those relationships to identify individual patient-level risk for multiple illnesses and

[13] http://www.hissjournal.com/content/2/1/3.

[14] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2845681/.

[15] http://www.bradenscale.com/.

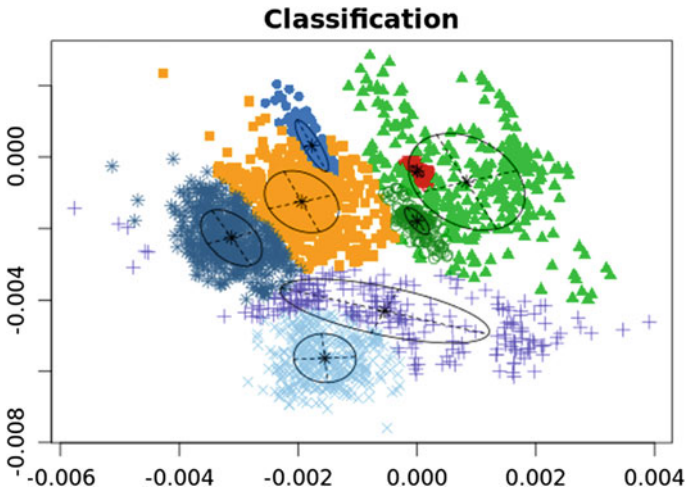


Figure 8: Jvion maps patients into an Eigenspace representation in order to identify clinically significant cohorts. Since the Eigenspace has as many dimensions as the number of clinical factors, it can't be represented graphically. This figure illustrates a two-dimensional "slice" of the patient mappings within an Eigenspace. The highlighted areas represent patient clusters. (Courtesy Jvion, All Rights Reserved)

conditions. They can also aggregate individual risk levels to determine population health trends and potential opportunities for intervention.

Jvion claims to achieve accuracy levels of between 75–90 percent depending on the particular clinical problem, case mix and population. As we said, the Braden scale is commonly used to predict the development of pressure ulcers, a problem which can significantly lengthen hospital stay. Jvion feels that the scale of the model, as with virtually all manual scoring systems for use by care providers, is somewhat limited, since it is based on only six clinical features. Using these, it maps patients into one of 19 risk levels and then to one of five intervention levels. Studies have found that the Braden Scale score is predictive of pressure ulcer development but, because of these limitations, it does not help clinicians develop an individualized prevention plan for each patient. Jvion says that the introduction of additional clinical factors (this is similar in many respects to the CHF early diagnosis example) can help focus care on each patient's specific risk factors.[16] The company's tool considers over 45 demographic factors, hundreds of patient-specific factors and thousands of other clinical factors to achieve a prediction accuracy which, it says, is some three to four times more accurate than the Braden scale.

The prediction of venous thromboembolism (VTE), which includes deep vein thrombosis (DVT) and pulmonary embolism (PE), is another example. DVT clots typically form in a deep vein, such as in the leg or pelvis. A dislodged clot can

[16] http://www.ncbi.nlm.nih.gov/pubmed/22552104.

travel to the lungs and become a PE, a condition that kills more people in the United States and the United Kingdom than AIDS, breast cancer, prostate cancer and traffic combined.[17]

DVT risk factors include inherited blood-clotting disorders, prolonged bed rest or lack of movement, hormone replacement therapy, smoking, cancer, family history and age. Some diagnostic tests for DVT (e.g. contrast venography) are expensive and accurate while other tests (e.g. observation and assessment) are inexpensive, but relatively inaccurate.[18] Among patients who survive a DVT episode, half will have long-term complications and about one-third will have a recurrence within 10 years.[19]

The Wells score to predict the probability of DVT calculates a total from 10 variables and is often used in conjunction with blood tests to determine a possible DVT risk level.[20] Like all scales designed to be used by clinicians, this approach has the advantage of being relatively simple. However, to achieve that simplicity, it does not account for the clinical context and all possible variables. For example, all cancers contribute equally to the Wells score. Yet some cancers—such as ovarian—have a higher impact on DVT risk. Other factors from a patient's health history are also not considered including age, gender, race, socioeconomic status and recent travel history.

Jvion says it trained a baseline model to a 72 percent accuracy rate at target coverage levels by using data from millions of retrospective patient encounters obtained from their client hospitals and publically available health data. Coverage represents the proportion of a data set on which predictions are made.[21] The company then applied the baseline model in live provider settings to further tune it. This process really never ends; the longer the model is used and the more patients it sees, the more accurate it gets. We saw an example earlier when we discussed how M*Modal trained its voice-understanding engine. This ability to continuously improve is one of the biggest advantages that machine learning-based technologies have over traditional CDS systems. As we discussed earlier, the ability to put these models in the cloud so that many centers can use them through web services, allows everyone to participate in improving the model and, as a result, benefits everyone.

As shown in Figure 9 for several actual patients (all identifying data has been removed), the results of Jvion's VTE prediction are presented to clinicians at the point of care as a risk score between 0 and 1. Importantly, contributing risk factors are also provided to suggest the most appropriate clinical interventions for each patient.

[17] http://www.hopkinsmedicine.org/innovation_quality_patient_care/areas_expertise/infections_complications/dvt/what_is_dvt_vte.html.

[18] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3183832/.

[19] http://www.cdc.gov/ncbddd/dvt/data.html.

[20] http://www.ncbi.nlm.nih.gov/pubmed/7752753.

[21] http://robotics.stanford.edu/~ronnyk/glossary.html.

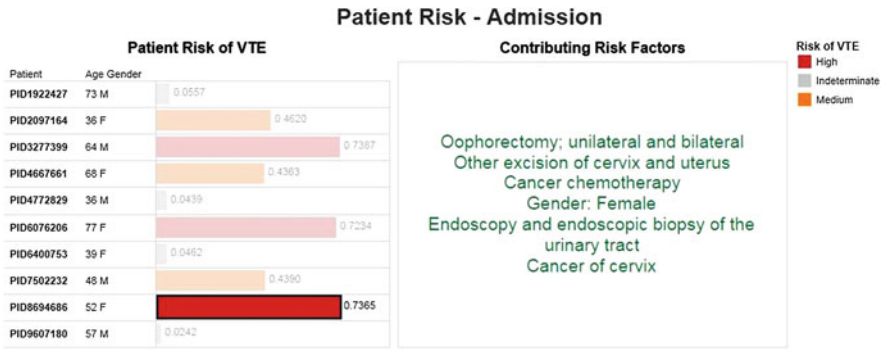


Figure 9: Jvion's model presents patient risk of VTE at time of admission on a scale of 0 to 1 along with the contributing risk factors to aid in making appropriate, personalized preventive clinical interventions. (Courtesy Jvion, All Rights Reserved)

According to Jvion, over a period of 12 months, some 85,000 patients were evaluated by the model which made 440 predictions of VTE. Again according to the company, appropriate clinical interventions are estimated to have produced $18.38 million in avoided cost and complications—some $41,770 per individual patient.

Personalized Cancer Care: We now understand that cancer is a more complex family of diseases than previously thought and that the results of treatment in any individual patient depend on the patient's particular genomic, proteomic and metabolic pathways and other personal factors, as well as on similar factors in the cancer. Moreover, given the high mutation rate in cancers, there is usually more than one cell type—essentially more than one cancer in each patient. Traditional chemotherapy drugs are metabolic poisons that selectively target rapidly growing cells and work because cancer cells divide more rapidly than most normal cells. However some cells, such as bone marrow, the cells that line the gastrointestinal tract and hair cells, do rapidly divide leading to undesirable and even dangerous side effects often seen in cancer patients. More recently developed mechanistic drugs selectively attack metabolic pathways in the cancer *if* it has the specific pathway structure that the treatment targets. In addition to the selectivity of these drugs, other important considerations in their use is that they are very expensive treatments and precious time is lost if they don't work.

Given the complexity of the biochemistry of cancer, without some analytic tool, it is hard to know which patients will benefit from a proposed therapy. Research at the Max Planck Institute in Germany was aimed at developing a predictive model (called a "Virtual Patient") to assist physicians in selecting the right mechanistic drug for each patient.[22] The very complex model contains all known cancer biochemical pathways. The researchers report that when the model is given genomic

[22]http://link.springer.com/article/10.1007/s00003-011-0755-7/fulltext.html.

data from a patient and from the primary cell types in their cancer (including if possible, cancer stem cells), it has successfully predicted the efficacy of a mechanistic drug treatment. In one particularly interesting patient, while the model's prediction of success was confirmed, metastases eventually developed and were immune to the drug. When the genomic data from these metastases was fed into the model, it correctly predicted the drug would not be effective in treating them. The model is now being offered commercially.[23]

Evidence-based Design: Research performed at Georgia Tech's SimTigrate Lab[24] considered various configurations of the beds in a hospital's surgical suite. This particular suite previously had three bed types: preoperative holding area, post anesthesia care unit and Level 2 recovery. The researchers developed a model using medBPM, a healthcare-specific modeling tool,[25] and showed that an alternate approach using universal beds that could be reconfigured for all three purposes could serve the same number of cases with approximately a third fewer beds. This is an interesting example of the use of modeling to inform more cost-effective design of a physical care delivery space, another new field that is being impacted by the availability of digital data and analytic technologies.

Process Mining: Using data from their now digital records, could we infer the underlying care processes that patients receive as they traverse the complex service areas of a hospital? More importantly, could we identify differences in these processes and the impact these have on outcomes and costs? Similar analysis is common in other industries, such as manufacturing, and is referred to as process mining. However, those other industries typically have standardized and often mechanized processes in which highly accurate sensors routinely provide digital, time-stamped data. This is not yet the norm in healthcare, still a largely manual industry where care steps are customized to each patient and, although they may be documented digitally, the data is often incomplete or inaccurate. It may also be recorded after the fact so the date/time stamp may not be accurate, making reconstruction of the time sequence of a process more difficult. Nevertheless, there are situations where process mining appears to be feasible. Research from the Netherlands—where process mining originated in the late 1990s—compared two hospitals for their care of ischemic stroke patients (patients with a clot in the arteries of their brain). The results, as shown in Figure 10, clearly indicate significant differences, with the hospital on the left using a more state-of-the-art neuro-protective approach.

Interactive Process Visualization: So far, the examples we've shown of visualizing data after some analytic process were essentially static. If the dataset changes or the analytics are run using different parameters, the visualization would

[23] http://www.alacris.de/.

[24] http://www.simtigrate.gatech.edu/.

[25] http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4114518&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D4114518.

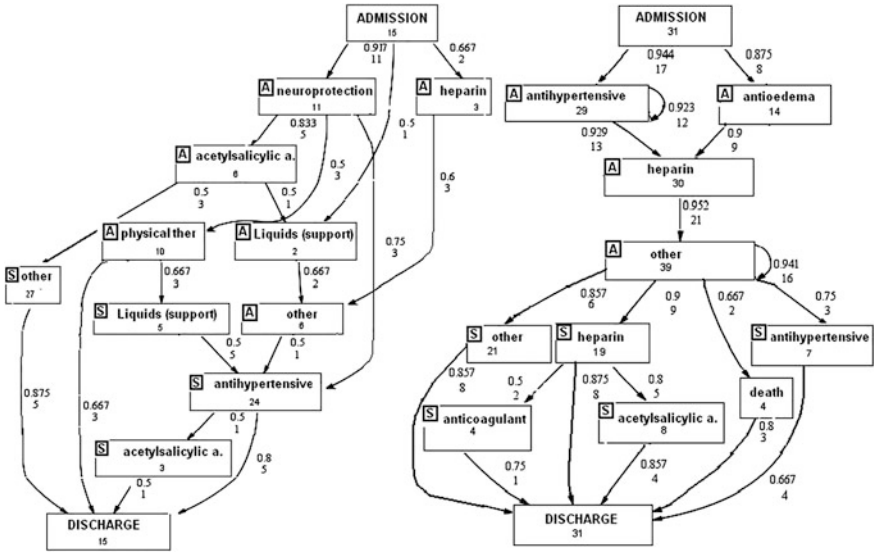


Figure 10: Process mining of emergency department data for ischemic stroke patients reveals a significant difference in the care offered at two Dutch hospitals. (Courtesy Ronny Mans) (http://ebooks.iospress.nl/publication/11642)

likely change. Suppose, however, that the visualization itself was the tool to explore the data. That's the goal of a research project at Georgia Tech that explores the care processes of a cohort of patients that went through a common care venue.

The interactive visualization example, shown in Figure 11, presents the emergency department care received by a cohort of pediatric asthma patients, derived from their EHR data. Each lab test, medication order and other care element is described as an activity. As would be expected children with a suspected infectious etiology for their exacerbation (shown in blue) receive different care than those suspected of having an allergic etiology.

Further iterations of research techniques such as this may someday inform the development of more cost-effective care pathways and also serve as tools to measure their use by clinicians in actual practice. Variations in the care provided by physicians and the impact that has on outcomes and cost are another potential use case for this technology.

Final Thoughts: I often point out to my students that health informatics is best appreciated by finding and considering the many connections that exist within the subdomains of the field. If you are a healthcare provider and actually use electronic health records, I hope that this final section provides a connection between the quality and detail of your day-to-day charting and the potential to use the digital health data you generate in many new, exciting and practical ways. Moreover, as we also solve the interoperability challenge, it will be easier for these new tools to

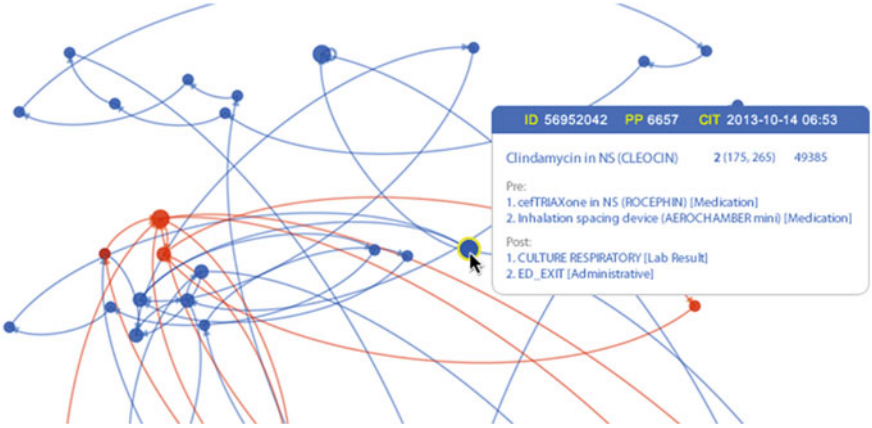


Figure 11: The visualization of clinical activities can reveal care patterns in a large cohort of similar patients. This process diagram depicts care of pediatric asthma patients in the emergency department. Patients in blue are suspected of having an infections problem while those in red are thought to have an allergic etiology for this asthma exacerbation. (Courtesy Rahul C. Basole, Tennenbaum Institute, Georgia Tech) (http://jamia.oxfordjournals.org/content/early/2015/02/05/jamia.ocu016.abstract?ijkey=TIVAVstVaFOzjwz&keytype=ref)

find their way back into the EHR you use. This is also increasingly true of patients who have a growing potential to provide insights and nuances not traditionally found in provider-facing records. Thus, medical and personal electronic records, health information exchange and innovative analytics are all intertwined in ways we're only now beginning to broadly discover and exploit in clinical practice. A primary purpose for writing this book was to help physicians and other care providers, as well as their patients, gain an early appreciation of this future landscape.

As a result, I hope the recognition of these connections provides encouragement to those readers engaged in direct patient care or who are the current or potential future recipients of it (e.g. all of us) to pay added attention to the increasing importance of the scope and accuracy of the digital clinical data we all create. The extra time and effort, when combined with similar efforts by thousands of other dedicated providers and millions of other patients and consumers, increasingly has the potential to create the multifaceted, multidimensional view of patient care we have learned in this final section has the greatest potential to help improve the quality and efficiency of healthcare delivery. As a result, over the long term, we all stand to benefit from the insights and knowledge that are increasingly being gleaned from our collective effort.

Indeed, together, as we intimated at the start of this section, we may well be standing on the threshold of actually creating the IOM's long sought goal of a "Learning Health Care System in America." I think it is useful to repeat that vision,

which we first encountered way back in the discussion of the U.S. healthcare system. I hope you can now visualize more clearly than before how health informatics, if designed, implemented and used well, can contribute to: "a learning health system designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety and value in health care."[26]

[26]http://www.nap.edu/catalog/11903/the-learning-healthcare-system-workshop-summary-iom-roundtable-on-evidence.

The Road Ahead

So, we've now reached the end of the story … at least the part that has played out so far.

We began with the nature and problems of our healthcare system and the key role that chronic disease plays in it. We made the case for digital records and data sharing and we learned what the federal government is doing to encourage their adoption. We examined the key underlying technologies of health information exchange; privacy, security and trust; and data and interoperability standards. We then looked at how these technologies are being used in actual systems and tools for providers, patients, public health and clinical research.

With this background we focused in particular, on the challenges faced in designing and using electronic health record systems and some promising approaches to overcoming those challenges. We also looked at the dynamic field of technology for engaging and supporting patients outside what we normally think of as the healthcare delivery system. We then explored how data can be aggregated for analysis and concluded with a number of exciting examples of the results that can be achieved. These early examples suggest tools that every practicing healthcare provider could find useful in order to diagnose and treat patients earlier and more accurately and effectively. They have the potential, as envisioned by the IOM, to help providers avoid mistakes and provide each of their patients with the best available care, given those patients' unique clinical, genomic and personal factors.

At the same time, not only must providers adapt to using informatics systems, but those systems must adapt to the real-world demands of practice. They cannot require significantly more time to use than the paper tools they seek to replace. We've seen some examples of how more intelligent and adaptive electronic records may be able to do that. Providers can't be asked to flip from system to system and among many different user interfaces in order to get the full benefits of informatics tools. We've seen how the JASON proposals might be used to seamlessly integrate tools, no matter who developed them, into a single and efficient user experience.

Of course, we aren't there today. Providers are right to be frustrated. They also need to recognize that no tool, no matter what its potential, will produce a

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8_14

satisfactory result if it isn't used properly. Far too often, informatics systems are viewed as something magical. Sometimes we act as if merely implementing them will solve problems that are largely the result of poor workflow and process. This is never the case. We've also seen some very early attempts to help understand clinical processes far more efficiently than the traditional "walking around with a clipboard" approach.

Finally, it is important to remember that, as recently as 2008–09, a very small percentage of hospitals and physician's offices had a clinically meaningful EHR. Today those percentages are above 60 percent for providers and 90 percent for hospitals. This does not mean that these systems are being used optimally, nor does it mean that their design is the best we could hope for. Nevertheless, we've come a long way. We have an even longer journey ahead of us. I have a real sense, as I write this, that the frustration generated by our current situation is leading to a serious national dialogue about solutions and that they may not be all that long in coming.

It is my hope that this book will encourage you to persevere and even come along and be part of the generation that actually achieves the long-held dream of transforming healthcare, in large part through fulfilling the promise of health informatics.

Index of Innovative Analytics in Healthcare

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8

Glossary of Health and Information Technology Terms and Acronyms

Accountable Care Organization (ACO): Medicare's outcomes-based contracting approach.

American Recovery and Reconstruction Act (ARRA) the Obama administration's 2009 economic stimulus bill.

Arden Syntax: an approach to specifying medical knowledge and clinical decision support rules in a form that is independent of any electronic health record (EHR) and thus sharable across hospitals.

Area under the Curve (AUC): a measure of how well a model will rank a randomly chosen positive instance higher than a randomly chosen negative example.

Blue Button: an ASCII text-based standard for heath information sharing first introduced by the Veteran's Administration to facilitate access to records stored in VistA by their patients. The newer Blue Button+ format provides both human and machine readable formats.

Centers for Disease Control and Prevention (CDC): the federal agency focused on disease in the community.

Centers for Medicare and Medicaid Services (CMS): the component of the Department of Health and Human Services that administers the Medicare and Medicaid programs.

Certificate Authority (CA): an entity that digitally signs certificate requests and issues X.509 digital certificates that link a public key to attributes of its owner.

Clinical Context Object Workshop (CCOW): an HL7 standard for synchronizing and coordinating applications to automatically follow the patient, user (and other) contexts allow the clinical user's experience to resemble interacting with a single system when the user is using multiple, independent applications from many different systems.

Clinical Document Architecture (CDA): an XML-based markup standard intended to specify the encoding, structure and semantics of clinical documents.

M.L. Braunstein, *Practitioner's Guide to Health Informatics*,
DOI 10.1007/978-3-319-17662-8

Clinical Information Modeling Initiative (CIMI): an independent collaboration of major health providers to improve the interoperability of healthcare information systems through shared and implementable clinical information models.

CommonWell Alliance: a group of major HIT companies that is working to achieve interoperability among their respective software products and services.

Complete EHR: an EHR software product that, by itself, is capable of meeting the requirements of certification and Meaningful Use.

CONNECT: ONC supported open source software for managing the centralized model of HIE.

Consolidated Clinical Document Architecture (CCDA): the second revision of HL7's CDA architecture that attempts to introduce more standard templates to facilitate information sharing (a mandate of Meaningful Use Stage 2).

Continua Health Alliance: a group of over 200 companies working on interoperability for health and fitness devices.

Continuity of Care Document (CCD): an XML-based patient summary based on the CDA architecture.

Continuity of Care Record (CCR): an XML-based patient summary format that preceded CDA.

Cross-Enterprise Document Sharing (XDS): the use of federated document repositories and a document registry to create a longitudinal record of information about a patient.

Current Procedural Terminology (CPT): the American Medical Association's standard for coding medical procedures.

De-identified Patient Health Information: PHI from which all data elements that could allow the data to be traced back to the patient have been removed.

Digital Imaging and Communications in Medicine (DICOM): a widely used standard for creation and exchange of medical images.

Direct: a set of ONC-supported standards for secure exchange of health information using e-mail.

Domain Name System (DNS): the naming system for computers, services or any resource connected to the Internet (or a private network). Among other things, it translates domain names (for example, eBay.com) to the numerical IP addresses needed to locate Internet connected resources.

EDI/X12: a format for electronic messaging that utilizes cryptic but compact notation primarily to support computer-to-computer commercial information exchange.

eHealth Exchange: a set of standards, services and policies that enable secure nationwide, Internet-based HIE using CONNECT or one of the commercial HIE products that support eHealth Exchange.

Electronic Health Record (EHR): a stakeholder-wide electronic record of a patient's complete health situation.

Electronic Health Record Certification: a set of technical requirements developed by ONC that, if met, qualify an EHR to be used by an eligible professional to achieve Meaningful Use.

Electronic Medical Record (EMR): an electronic record used by a licensed professional care provider.

Eligible Professionals (Medicaid): health providers who are eligible for Medicaid Meaningful Use payments: doctors of medicine, osteopathy, dental surgery, dental medicine, nurse practitioners, nursecertified, nurse-midwives and physician assistants who work in a federally qualified health center or rural health clinic that is led by a physician assistant.

Eligible Professionals (Medicare): health providers who are eligible for Medicare Meaningful Use payments: doctors of medicine, osteopathy, dental surgery, dental medicine, podiatry, optometry and chiropractic.

EMPI: an enterprise master patient index.

Electronic Healthcare Network Accreditation Commission (EHNAC): an independent, federally recognized, standards development organization focused on improving the quality of healthcare transactions, operational efficiency and data security.

Extensible Markup Language (XML): a widely used standard for machine- and human-readable electronic documents and the language used to define CDA templates.

Fast Health Interoperable Resources (FHIR®): an HL7 initiative that seeks to use modern web standards and technologies to simplify and expedite real-world interoperability solutions.

Health System: a network of providers that are affiliated for the more integrated delivery of care.

Health Information Exchange (HIE): the sharing of digital health information by the various stakeholders involved, including the patient.

Health Information Service Provider (HISP): a component of Direct that provides a provider directory, secure e-mail addresses and public-key infrastructure (PKI).

Health Information Technology (HIT): the set of tools needed to facilitate electronic documentation and management of healthcare delivery.

Health Insurance Portability and Accountability Act of 1996 (HIPAA): legislation intended to secure health insurance for employees changing jobs and simplify administration with electronic transactions. It also defines the rules concerning patient privacy and security for PHI.

Health Level 7 (HL7): a not-for-profit global organization to establish standards for interoperability.

Health Maintenance Organization (HMO): an organization that provides managed healthcare on a prepaid basis. Employers with 25 or more employees must offer federally certified HMO options if they offer traditional healthcare options.

Healthcare Information Technology Standards Panel (HITSP): a public and private partnership to promote interoperability through standards.

Healtheway: an ONC-supported public-private partnership to promote nationwide HIE via the eHealth Exchange.

HIMSS: describes itself as "a global, cause-based, not-for-profit organization focused on better health through information technology (IT)."

HL7 Development Framework (HDF): the framework used by HL7 to produce specifications for data, messaging process and other standards.

hQuery: an ONC-funded, open source effort to develop a generalized set of distributed queries across diverse EHRs for such purposes as clinical research.

Hypertext Transfer Protocol (HTTP): a query-response protocol used to transfer information between web browsers and connected servers. HTTPS is the secure version.

i2b2 (Informatics for Integrating Biology and the Bedside): a scalable query framework for exploration of clinical and genomic data for research to design targeted therapies for individual patients with diseases having genetic origins.

IHE Cross-Enterprise Document Media Interchange (XDM): a standard mechanism for including both documents and metadata in zip format using agreed upon conventions for directory structure and location of files.

IHE Cross-Enterprise Document Reliable Interchange (XDR): a standard mechanism for exchanging both documents and metadata using SOAP web services as the transport mechanism.

International Classification of Diseases (ICD): the World Health Organization's almost universally used standard codes for diagnoses. The current version is ICD-10, but ICD-9 is used in most U.S. institutions. The conversion target, set by CMS, is currently October 1, 2015.

International Health Terminology Standard Development Organisation (IHTSDO): the multinational organization that maintains SNOMED.

Internet of Things (IoT): in healthcare this describes a profusion of Internet-connected devices, sensors and other equipment that has the potential to transform care delivery.

Interoperability: the ability of diverse information systems to seamlessly share data and coordinate on tasks involving multiple systems.

IP Address: a 32-bit (the standard is changing to 128-bit to accommodate Internet growth) number assigned to each device in an Internet Protocol network that indicates where it is in that network.

JASON: an independent group of some 30–60 scientists that advises the United States government on matters of science and technology.

JavaScript Object Notation (JSON): is a relatively simple, human readable data-interchange format for packaging a group of data items that is also easy for computers to parse and generate. It is based on a subset of the JavaScript programming language that is widely used on the web.

Lightweight Directory Access Protocol (LDAP): a protocol for accessing (including searching) and maintaining distributed directory information services (such as an e-mail directory) over an IP network.

Logical Observation Identifiers Names and Codes (LOINC): the Regenstrief Institute's standard for laboratory and clinical observations.

Massachusetts General Utility Multi-Programming System (MUMPS): an integrated programming language and file management system designed in the late 1960s for medical data processing that is the basis for some of the most widely installed enterprise health information systems.

Master Patient Index (MPI): software to provide correct matching of patients across multiple software systems, typically within a health enterprise.

Meaningful Use: a set of usage requirements defined in three stages by ONC under which eligible professionals are paid for adopting a certified EHR.

MEDCIN: a proprietary vocabulary of point-of-care terminology, intended for use in electronic health record systems (as a potential alternative to SNOMED-CT) maintained by Medicomp Systems.

Medicaid: the joint federal and state program to provide healthcare services to poor and some disabled U.S. citizens.

Medical Dictionary for Regulatory Activities (MedDRA): the International Conference on Harmonisation's classification of adverse event information associated with the use of biopharmaceuticals and other medical products.

Medical Logic Module (MLM): the basic unit in the Arden Syntax that contains sufficient medical knowledge and rules to make one clinical decision.

Medicare: the federally operated program to provide healthcare services to U.S. citizens over the age of 65.

Multipurpose Internet Mail Extensions (MIME): the Internet standard for the format of e-mail attachments used in Direct. S/MIME is the secure version.

Modular EHR: a software component that delivers at least one of the key services required of a Certified EHR..

National Drug Codes (NDC): the Food and Drug Administration's numbering system for all medications commercially available in the U.S.

Office of the National Coordinator for Health Information Technology (ONC): the agency created in 2004 within the Department of Health and Human Services to promote the deployment of HIT in the U. S.

Open mHealth: a non-profit collaboration seeking to address interoperability in the mobile health app and device space.

OpenNotes: is a national initiative (not a technology) working to give patients access to the visit notes written by their healthcare providers.

Outcomes-Based Contract: an approach to pay for healthcare that rewards physician performance against certain defined quality metrics when combined with a lower-than-predicted cost of care.

Patient-Centered Medical Home (PCMH): a team-based healthcare delivery model often particularly focused on the management of chronic disease.

Pay-for-Performance (P4P): an approach to pay for healthcare that rewards physician performance against certain defined quality metrics.

Personal Health Record (PHR): typically a web page where health data and information related to their care is maintained by the patient.

Physician Group Practice (PGP) Demonstration: the first pay-for-performance initiative for physicians under the Medicare program.

Preferred Provider Organization (PPO): a network of providers who have contracted to provide care to patients (usually at a discounted price) under an insurance plan.

Primary Care Physician (PCP): the generalist in a patient's care team who assumes overall responsibility for all their health issues and often the gatekeeper who must generate referrals to specialists.

Private Key: the protected (known only to its owner) part of the special pair of numbers used to encrypt documents using PKI.

Protected Health Information (PHI): health or health-related information that can be linked to or used to identify a specific patient. PHI is subject to strict HIPAA regulations.

Provider: health professionals, including physicians, nurse practitioners, physicians' assistants, that are engaged in direct patient care.

Public Key: the public part of the special pair of numbers used to encrypt documents using PKI.

Public Key Infrastructure (PKI): a widely used system for protection of documents, messages and other data that rests on a pair of public and private keys to allow for a variety of use cases.

Read Codes: a hierarchical clinical terminology system used in general practice in the United Kingdom.

Reference Information Model (RIM): a pictorial representation of the HL7 clinical data (domains) that illustrates the life cycle of an HL7 message or groups of related messages.

Registration Authority (RA): an entity that collects information for the purpose of verifying the identity of an individual or organization and produces a certificate request.

Resource Description Framework (RDF): a method for describing or modeling information on the web using subject-predicate-object expressions (triples) in the form of subject-predicate-object expressions that could be used to represent health ontologies (SNOMED, ICD-10).

Representational State Transfer (REST): Web interoperability principles proposed by Roy Fielding as a simple, consistent implementation of HTTPS basic commands (GET, PUT, POST or DELETE) for transfer of media (which can be data, images or other forms of digital information) between a server and a client. The ease and speed of REST development and led to its growing use for web interoperability. REST is FHIR's preferred transport protocol implementation for exchanging FHIR Resources.

Semantic Web: the proposed next generation of web in which technologies like RDF would create a "web of data" in which browsers (and other tools) could "understand" the content of webpages.

Simplified Mail Transport Protocol (SMTP): the Internet standard for e-mail used by Direct. The secure version is S/SMTP.

Simple Object Access Protocol (SOAP): a simple protocol for exchanging XML formatted information between applications using the Internet.

Systemized Nomenclature of Medicine (SNOMED): a comprehensive, hierarchical healthcare terminology system.

Systemized Nomenclature of Medicine Clinical Terms (SNOMED CT): SNOMED subset for the EHR.

Synthetic Health Data: facsimile clinical data created by a software system to realistically resemble actual patient data.

Templates: the reusable basic XML-based building blocks of a CDA document that can represent the entire document, its sections or the data entries within a section.

Transition of Care Initiative (ToC): the effort to develop a standard electronic clinical summary for transitions of care from one venue to another.

Treatment, Payment or Operations (TPO): HIPAA exception for providers, insurance companies and other healthcare entities to exchange information necessary for treatment, payment or operations of healthcare businesses

Unified Medical Language System (UMLS): a service of the National Library of Medicine, it links many health and biomedical vocabularies and standards to facilitate interoperability.

Veterans Health Information Systems and Technology Architecture (VistA): the VA's system-wide, MUMPS-based health information infrastructure.

View, Download, Transmit (VDT): a requirement of Meaningful Use Stage 2 that patients view, download or transmit their health information.

Web services: a method of communicating between two devices or software applications over the Internet.

X.509 Digital Certificate: the technical name for an electronic document issued by a CA that uses a digital signature to bind a public key with an identity based on information from an RA.

XMPI: a cross-organizational master patient index capable of dealing with many unaffiliated hospitals and health systems.