Integration of Genetic and Familial Data into

Electronic Medical Records and Healthcare Processes[©]

By

Thomas Kmiecik and Dale Sanders

February 2, 2009

Introduction

Although our health is certainly impacted by lifestyle and environment, there are clearly genetic factors which can also have significant influence. Recent technology developments have made possible the quick analysis of genetic material and its relationship to an individual's health. Over 160 different healthcare-related conditions and diseases have now been clearly identified as having a significant genetic influence, from an individual's response to ACE inhibitors, to increased risk of venous thromboembolism.   While the pace of understanding genetic influence on health is increasing, the pace could move much faster if the underlying genetic data could be more effectively integrated with familial and clinical care data in electronic medical records (EMRs).

Healthcare information systems, particularly EMRs, need significant modifications to incorporate genetic information in a manner which allows clinicians to efficiently include the information into diagnostic and treatment decisions. This use of genetic information to specifically tailor treatment protocols to an individual is referred to as "personalized medicine." Among other major benefits, personalized medicine will empower physicians to select the best medication and dosage based upon specific genetic makeup.

In order to understand genetic data and personalized medicine, it is necessary to have a basic understanding of genetics. The cells in our bodies each contain 23 chromosomes,

which are made of strands of the chemical DNA. Those strands contain genes, which encode our inherited characteristics. The total store of genetic information that an individual carries is called the genome. The genes produce the individual proteins that function in the various biochemical pathways that are central to our health. Each gene is made up of individual subunits called nucleotides, and much of the genetic variation between humans results from the substitution of individual nucleotides. These variations are called Single Nucleotide Polymorphisms or SNPs. Once genetic variants are identified, it is necessary to establish which correlations exist between disease and known genetic variants by using databases that store information on all known genetic variants, along with information on any disease associations which might exists. For example, the Human Variome Project (http://www.humanvariomeproject.org/) is a global initiative to catalog all human genetic variation.

Familial data describes the relationships between family members. In a healthcare setting, these familial relationships are particularly important as a simple and easy step towards understanding the effects of heredity in diseases and other healthcare related issues, such as allergies and tolerances to medications. In the absence of specific genetic data, which is typically more difficult and costly to obtain, familial data can serve as a "poor man's" first step to genetic medicine. For example, ovarian cancer tends to show strong familial relationships among "first degree" relationships-- that is mother, sister, or daughter. A complicated genetic analysis is not necessary to observe the hereditary trend that exists between the female members of a family. Both tolerance and intolerance for anesthesia also show strong familial relationships. Armed with this information about familial trends in disease, physicians can adjust their treatment

protocols and preventive diagnostic screenings accordingly. Familial health data can act as an accelerant to understanding the more complicated factors influencing an individual's genetic influence on their healthcare. Unfortunately, most healthcare organizations and EMRs do not capture or store a computable version of familial data, even though the information is routinely captured in paper health history forms. Possessing a complete profile of a person's genetic, familial, and clinical outcomes data enables a much faster and more accurate understanding of the influence of genetics on a person's overall health profile. Incorporating all of this information and exposing it in a useful fashion at the point of care in an EMR allows physicians to make a fully informed and customized decision when delivering care to those patients.

Organizations such as SNOMED (Systematized Nomenclature of Medicine) are now beginning to make the necessary changes in terminologies and data standards that are needed to accommodate genetic information. A major effort is also underway at the National eHealth Collaborative (NeHC, formerly AHIC) (http://www.nationalehealth.org/). The NeHC is developing use cases for personalized healthcare and for family health history (a use case is a description of a system's behavior used as a starting point for software development), which the Health Information Technology Standards Panel (HITSP) then uses to produce a set of standards for implementing the functionality they describe. Those standards are typically recognized by HHS and used in the certification criteria that are produced by the Certification Commission for Healthcare Information Technology.

Current Information Systems

Information systems have been built by academic institutions and by commercial vendors and are in varying states of maturity. Following is a brief overview of some of those systems.

Partners HealthCare

The Partners HealthCare Center for Genetics and Genomics (www.hpcgg.org) has created an advanced computer system for integrating genetic and clinical data. On top of the underlying patient data warehouse is an enterprise Laboratory Information Management System called Gateway for Integrated Genomics-Proteomics Applications and Data (GIGPAD), which passes information to a genomic knowledge base (GeneInsight). There is a consent tracking system which serves as a gateway to the genomic data. From that point, different systems are used for research and clinical purposes. On the clinical side, there is a genomic variant interpretation engine which connects to the EMR and passes the information to a clinical decision support rules engine.

IBM

IBM has developed a Clinical Genomics solution to facilitate the integration of genomic and clinical data. The specific system design can be modified to fit existing architectures or strategies. The design is built upon the HL7 clinical genomics object model. The modular design of the IBM system will facilitate changes to the system that might be necessary. The IBM Clinical Genomics solution has been used by the Mayo Clinic to

perform research on a patient database with over 4.3 million records, establishing the scalability of the architecture.

Inforsense

The basic Inforsense software (www.inforsense.com) is a data mining package originally designed for business application. It has been extended to produce the Inforsense Translational Research system, which allows for integration of experimental and clinical data for translational research. The Inforsense system is built on a service-oriented architecture. The Inforsense system allows a complete system to be built up from separate modules, such as ClinicalSense, GenSense, and ChemSense. The Inforsense system can utilize data from EMRs such as Epic and Cerner, and can also incorporate data from Oracle, Microsoft, and IBM databases. As with other translational research systems, the Inforsense software exists as an independent application separate from any EMR which houses the original clinical data.

Xenobase

The Xenobase system (www.xbtransmed.com) is from the Van Andel Research Institute. The Xenobase system can integrate data from EMR providers like Cerner and Epic and provide an external translational research computing environment. The strength of the Xenobase system is in the level of integration, and the ability it provides an organization to quickly establish the IT structure necessary to allow research on genetic/clinical questions.

Cerner

The EMR system produced by Cerner (www.cerner.com) includes the Millennium Helix for Laboratory software for clinical genomics. The software enables clinical labs to integrate molecular diagnostic data with electronic medical records. The software uses the Cerner Clinical Bioinformatics Ontology to provide a standardized vocabulary needed for molecular diagnostics and cytogenetics. The Cerner software allows the immediate application of genetic data to diagnostic and therapeutic decisions, but is specific to the Cerner EMR software.

Indivo

The Indivo system has recently been adopted as the Payroll and Human Resources (PHR) system for a group of employers such as Wal-Mart, AT&T, Pitney-Bowes, BP, and Cardinal Health known as the Dossia Consortium (http://www.dossia.org).

The Indivo system will import data from EMRs such as Cerner and Allscripts. The Indivo system can use a subscription agent, which will periodically identify new or changed data in an EMR system and transfer it to the Indivo database. However, patient-defined access policies limit the actions that may be performed, and can override any privileges that may be assigned to the role.

The Indivo system has been extended to handle genomic data. The design of Indivo is such that it is not possible to aggregate the data in the system to perform queries that retrieve all the patients with pancreatic cancer, etc. Only accessing an individual patient

is efficient. The system therefore does not directly provide any support for research or quality control queries.

Familial health data can be used as a relatively simple, low cost proxy for genetic data, and healthcare IT systems and processes could be modified relatively quickly to accept such data. The Surgeon General's Family History Initiative has developed a web accessible software tool (https://familyhistory.hhs.gov) that allows individuals to enter in their family history. The freely downloadable RiskApps application, built by Dr. Kevin Hughes at Massachusetts General Hospital to identify and manage women that are at high risk for hereditary breast and ovarian cancer (http://hughesriskapps.com), will import family history data from the Surgeon General software application.

Intermountain Healthcare is building a family history software tool which uses Microsoft's HealthVault as a framework and will store the data in a standardized manner, therefore allowing it to be read by different electronic record systems.

The National eHealth Collaborative (NeHC) Personalized Health Care Workgroup has initiated a Family Health History Project, which intends to produce a web-based tool that would collect data. The workgroup then plans to partner with organizations that would build capabilities into their EMR or PHR systems to accept the file output from the NeHC tool.

HITSP has developed a Clinical Decision Support Component that is intended for use in communicating genetic and family history data from a healthcare IT application to a

clinical decision support system that would then provide a risk assessment to the patient.

Partners Healthcare is currently working on producing a 'Genetics Enabler Kit' (GEK), which would combine a genetic marker repository, test definition catalog, genome explorer interface, and a genetic variant interpretation engine. Their intent is to use the GEK to enable EHRs, PHRs, or Pharmacy systems to utilize genetic data.

Summary and Recommendations

In the future, a patient's longitudinal EMR health record will include a complete sequence of their genome. That genome will be analyzed in the background for general, population-based risks to their health, and the results of this analysis will be displayed to the physician and the patient, e.g., "This patient shows a genetic predisposition to warfarin metabolism anomalies which could result in abnormal bleeding. Lower dosing and attentive monitoring are advised." In the meantime, while the healthcare industry struggles through the ethical and technical hurdles to achieve this vision, familial data offers a simple, low cost alternative to more personalized medicine that recognizes the influence of heredity, e.g., "This patient shows two first-degree familial relationships with ovarian cancer, increasing her lifetime risk of ovarian cancer by 27%. Attentive monitoring and lifestyle awareness are advised."

To be prepare for the future of personalized medicine, healthcare CIOs should start planning now. All organizations would be best served by first taking advantage of family history data, monitoring progress on government-funded family history tools, and then

begin using those tools in their practices when they become available. Organizations should begin adapting their laboratory information systems to handle genetic data, should begin considering their approach to the data storage requirements that personalized medicine will introduce, and should establish the governance policies that will be necessary to properly handle the data. CIOs should also start familiarizing themselves with software tools produced for open source collaboration, such as those at Partners Healthcare. Academic medical centers should investigate collaborations with commercial drug or diagnostic partners to take advantage of synergistic combinations of organizational abilities and resources. Organizations should follow HITSP guidelines for family history data, genetic data, and clinical decision support. Any organization with a custom built clinical information system should consider the Partners Healthcare 'Genetics Enabler Kit' when it becomes available.

Authors:

Thomas Kmiecik is a Project Manager in the Department of Surgery at Northwestern University's Feinberg School of Medicine. He is also a graduate of NU's Masters in Medical Informatics program.

Dale Sanders is Chief Information Officer and Vice President of Information Services at Northwestern University's Medical Faculty Foundation. He is also a professor-lecturer, thesis advisor, and curriculum advisor for NU's Masters in Medical Informatics program.