Can Predictive Analytics Become Healthcare’s Crystal Ball?

Source: HIMSS Clinical & Business Intelligence Committee

April 2016

INTRODUCTION

As our healthcare system transitions from fee-for-service to fee-for-value and risk-based models, and providers go at-risk in payment models such as accountable care organizations (ACOs) in the Medicare Shared Savings Program (MSSP), commercial ACOs, bundled payments, Delivery System Reform Incentive Program (DSRIP) and others, it is essential that initiatives to improve quality and outcomes are aligned with efforts to optimize costs and utilization. This alignment is driven by rapidly advancing technology that enables healthcare stakeholders to leverage data across multiple, disparate settings into real-time, meaningful information to track and predict outcomes.

In the past, claims data aggregation and the analysis of data were in the domain of health plans. With greater adoption of electronic health records (EHRs) and other electronic reporting systems, and their capabilities to capture and analyze data, health systems, hospitals, and other provider organizations now have the ability to utilize clinical data to improve performance. Successfully leveraging these data along with pharmacy and laboratory data has enabled organizations to track and deliver a superior quality of care than was possible in the past. In conjunction with claims data, these organizations are also developing a greater understanding of utilization of resources and optimizing their costs.

Predictive analytics and the information they provide are central to this. Predictive analytics is increasingly being implemented by both providers and payers to power initiatives from improving resource utilization and cost optimization, to diagnosing diseases and coordinating care, to managing patient and staff workflow.

This compendium discusses the role of predictive analytics in improving outcomes and managing costs in three areas:

1. **Improving Care Delivery**: Improving clinical care by predicting potential failure before it happens. Examples are predicting sepsis, progression of renal failure, kidney injury, preventing suicide, predicting cancer outcomes, and diagnosing rare diseases in children;

2. **Cost Containment**: Cost containment by reducing fraud and abuse, optimizing patient and provider workflow, and resource utilization and predictive maintenance; and

3. **Reducing Preventable Readmission**: Utilizing predictive models to identify and manage high-risk patients and reduce preventable readmissions.
All of the three sections are extensively researched and provide links and citations to various articles for further learning. They provide an overview of how predictive analytics is being leveraged to improve outcomes for various conditions and optimize costs.

1. IMPROVING CARE DELIVERY
Up until several years ago, many of the clinical buzzwords about outcomes dealt with care paths, risk-management, evidence-based medicine, registries, and clinical trials. Expertise was concentrated primarily in the hands of insurance companies, finance, healthcare researchers, or the hospital quality departments. The vast amount of data was accessible only by a select group of people. With the advent of big data in concert with governmental regulations (ACOs, MSSP, Meaningful Use / EMR adoption, Hospital Readmissions Reduction Program (HRRP)) and advanced computing on ever-smaller devices, clinical information is now available to clinicians and patients in real time at their fingertips. Buzzwords have changed to population health, predictive medicine, neural networks, natural language processing, optimization, and predictive analytics. With this explosion of data, there has been movement on many fronts within the clinical arena. Let’s explore some avenues in the new frontier.

Predict Failure, Progression, Infection
The ability to predict risk of organ failure, disease progression, or infection is invaluable to a health care provider. Controlling or managing diseases before they become problems is instrumental in reducing costs and producing better outcomes in patients. Careful selection of clinical variables from the EHR is vital to accurately assess the patient’s risk for failure. Also, ensuring all applicable data is available in an easily understandable format. Poor model design or poor visualization may cause unexpected outcomes or wrong diagnosis. Therefore, it is important that the EHR infrastructure is able to support the use of predictive models in real time. The following provides examples in three disease areas.

- Renal Failure: Predicting Renal Failure Progression in Chronic Kidney Disease Using Integrated Intelligent Fuzzy Expert System
- Acute Kidney Injury: Utilizing electronic health records to predict acute kidney injury risk and outcomes: workgroup statements from the 15th ADQI Consensus Conference Predict AKI
- Parkinson’s Disease: Winning Data Challenge Entry Seeks to Use ‘Machine Learning Approach’ to Improve the Monitoring of Parkinson’s Disease
Clinical Algorithms Development
Over the past several years, the number of companies dealing with predictive analytics has exploded. To gain access to large amounts of data, companies have partnered with the Veterans Health Administration (VHA), professional organizations, health associations, or large hospital systems to develop predictive models. Collaboration among large hospital systems, specialized companies, and healthcare associations will help speed up the development of reliable clinical algorithms.

- IBM Watson is working with the VHA to pull data from post-traumatic stress disorder (PTSD) patients to help prevent suicide.
- Cerner’s predictive algorithm for sepsis improved the mortality rates at one hospital system, and they are now testing the model at 490 other facilities.
- American Heart Association (AHA) has selected IBM Watson and Welltok to help determine factors to optimize heart health by using science-based metrics and patient health assessments.
- American Society of Clinical Oncology (ASCO) and SAP are developing a system called CancerLinQ that will combine evidence based medicine, expert group knowledge, and patient experiences through a big data registry project to identify the best evidence-based course of cancer care, provide real-time benchmarking against national standards, and analyze unseen patterns in patient characteristics and outcomes.
  - Watch a video of ASCO President Dr. Peter Paul Yu discuss CancerLinQ. [PLAY]
- VHA\(^1\) developed point of care decision models that predict acute kidney injury (AKI)
  1. within the first 48 hours of admission
  2. following a coronary angioplasty
- Penn Medicine developed predictive models to improving Sepsis Mortality Index within their EHR

Diagnosing Diseases
A collaborative effort between IBM Watson and Boston Children’s Hospital is in the developmental stages of diagnosing rare diseases in children. In the world of pediatrics, diagnosing diseases can be difficult because research is scarce for factors that can influence the final clinical decision. Finding relevant articles and matching up specific information about a patient’s condition is extremely challenging. IBM Watson has the capability to quickly analyze and assess information from the vast amount of periodicals and journals to produce a relevant diagnosis in a much shorter period of time.

- Boston Children’s, IBM Watson take on rare diseases
- How Data Science is Fighting Disease

\(^1\) See also “VA Puget Sound Health Care System - Davies Enterprise/Organizational Award” and learn more about their Davies Award winning program.
**Risk Scores / Monitoring Current Condition / Alerts**

Risk scores primarily used in nursing assessments have been around for a long time: stroke, falls, coma, deep vein thrombosis (DVT). Big data analytics now allows for testing the accuracy of these predictive tools over a larger population and further refinement of these tools. Essentially, analytics can be used as a feedback loop that can help validate (or invalidate) specific variables used in the predictive risk tools.

- Risk Score DVT Study: [Risk Assessment of Recurrence in Patients With Unprovoked Deep Vein Thrombosis or Pulmonary Embolism](#)

**2. COST CONTAINMENT**

Although concern over the rising costs of healthcare has been an area of focus for years, it has only been recently – through efforts to transform the industry from the traditional fee-for-service structure to pay-for-performance – that cost containment has become a strategic imperative. Value-based business models require healthcare providers to focus on reducing total cost of care.

Cost in healthcare comes in various forms, including labor, supplies, facilities, services, etc. As healthcare is primarily a service industry, *people cost* is a major component and can account for as much as two thirds of the cost of a major medical center. With the complexity and variability inherent in healthcare, optimizing the various cost elements, especially people, is a big challenge. This is the kind of challenge that Big Data and predictive analytics have the potential to aid in the solution.

Additionally, cost is a function of perspective as to whether one is the provider, payer, employer, or consumer. Under a fee-for-service business model, the payers and employers have had a strong financial incentive to invest in efforts to uncover and detect fraud and inappropriate charges.

Predictive analytics has proven to be an effective tool to sift through the streams of claims data to highlight billing anomalies that warrant further attention, much in the same way that credit card companies can identify potential credit card fraud after a single questionable charge. With the shift to pay-for-performance models for providers and greater deductibles and out of pocket expenses for consumers, there is now a financial incentive for everyone to harness the power of data to better manage cost.

The following reflects areas of focus within healthcare where predictive analytics has been successfully used to manage and reduce costs, or, at a minimum, demonstrated to hold the potential for future cost containment strategies.
• **Fraud Detection** – *Monitoring claims data to detect and deter fraudulent charges.*
  Credit card companies have developed and refined complex algorithms that identify potentially fraudulent credit card transactions and automatically place holds on accounts until hard holders can confirm or deny the appropriateness of the questionable transactions. With national healthcare expenditures in the United States at $3.0 trillion, or $9,523 per person, as reported by Centers for Medicare and Medicaid Services (CMS) for FY2014, the potential for improper or fraudulent charges is significant. In a March 2015 Government Accountability Organization (GAO) study of government-wide improper payments, it was estimated that in FY2014 there were $45.8 billion in improper Medicare fee-for-service payments and $17.5 billion in improper Medicaid payments.

Through funding by Congress within the framework of the 2010 Small Business Jobs Act, CMS was able to explore the adoption of predictive analytics tools to uncover and stop fraudulent Medicare and Medicaid payments. Several years into the program, the predictive analytics tools focus on four models to detect fraud:

1. rules-based
2. anomaly
3. predictive
4. social networking

During the second year after implementation, the program found 469 new investigations and 348 assisted ones for a total of $211 million in savings. Because of the complexity of healthcare, the system does not provide the automatic, real-time protection that the credit card industry has realized, but it is a step in the right direction toward addressing this significant cost problem within the healthcare industry.

Additional research and development will be required to improve the sensitivity and specificity of this analytic tool in order to ensure that fraudulent charges are caught before they are paid without adversely impacting legitimate charges that might trigger a false “positive” by the analytics.

• **Health Care Resource Utilization** – *Predicting patient health care utilization needs for optimal care planning.*
  As hospitals take on more risk through bundled payments, ACOs, DSRIP, etc., they will need to better understand their costs and cost drivers. This is especially true for health care resource utilization. Provider report cards and benchmarking are playing a significant role in managing resource utilization and costs especially where payments are fixed (DRG-based payments or bundled payments). Even retrospective analysis such
as risk adjusted length of stay analytics and readmissions rates by provider and floor can play a significant role in resource utilization and optimization.

For example, University of Pittsburgh Medical Center (UPMC) uses a variety of modeling tools to identify patients who are high utilizers of its services and—significantly—are likely to continue to be high utilizers in the next year. They look to find the 20% of patients who won't get better on their own but who could respond to intensive, coordinated care.

The VHA analyzes 120 variables to predict six separate probabilities for each veteran, calculating 90-day and one-year “all risk” scores for hospitalization, death, and hospitalization and/or death. The risk scores are used to guide Patient-Aligned Care Teams as they apply the principles of patient-centered medical homes throughout their network of 152 hospitals and 990 outpatient clinics.

The Maine HIE, HealthInfoNet, performed a retrospective study of 1,273,114 patients over a 12-month period of EMR records to develop a predictive model for forecasting each patient’s next 6-month risk of resource utilization. A prospective validation was conducted of 1,358,153 patients in the following period. The study demonstrated a strong correlation between care resource utilization and predictive algorithm-based risk scores. The desire is to be able to apply these predictive algorithms to enable more effective care management strategies that drive improved patient outcomes while reducing overall cost of care.

- **Patient and Staff Workflow Optimization** — Optimizing patient flows and staff processes through the power of predictive data.

Significant gains in streamlining processes and eliminating waste from patient, staff, and material flows have been achieved by ongoing operational excellence and performance improvement strategies harnessing Lean and other proven quality systems throughout hospital and health system organizations. Much of this work is team based and results from a blend of qualitative and quantitative understanding and analysis. Significant gains in patient, staff, and material workflows can be attained by the use of Lean and other organization-wide continuous improvement initiatives. As Healthcare processes tend to be complex systems, striving for continual improvement will require greater amounts of quantitative analysis to fully understand underlying trends and root causes.

With the focus on interoperability of electronic health record systems, smart medical equipment, and other health IT systems, there is a growing source of data involving the location, condition, and association of patients, staff, medical equipment, and critical
supplies. This data results from a tapestry of bar code scans, electronic record entries, proximity reads of ID badges, and location data from real time location systems (RTLS).

Dr. Aviv Gladman, Chief Medical Information Officer, at Mackenzie Health in Ontario, Canada, shared an insightful presentation entitled “Predictive Analytics and Workflow Optimization: Big Metadata and the Internet of Healthcare Things” at HIMSS Annual Conference in Las Vegas on February 29, 2016. Mackenzie Health’s Innovation Unit has been exploring the use of metadata created by a breadth of systems, including: RTLS, Nurse Call, Smart Beds, Mobile Devices, and Hand Hygiene, combined with industrial engineering process simulation methodologies to understand trends and potentially develop predictive analytics to optimize care delivery processes.

Examples of harnessing data to predict future events as an early warning to change current operational practices is not limited to futuristic “Internet of Things” type efforts. The VHA supported a research project to explore the ability to predict and aggregate in real time the probabilities that emergency department patients will need to be admitted to a hospital inpatient unit. The concept was to utilize patient data collected at the time that they first arrived at the emergency department in order to predict hospital admission needs later in time. By establishing an “early warning” of future in bed and other patient admission needs, hospital staff could proactively prioritize activities to minimize delays and reduce emergency department crowding.

- **Predictive Maintenance** – *Condition monitoring to predict optimal scheduling of service and preventative maintenance to eliminate unscheduled or excessive downtime.*

The commercial aircraft industry has developed sophisticated predictive maintenance programs that deliver value each and every day. The Airbus system called Aircraft Maintenance Analysis (Airman) is used by 106 customers and constantly monitors the health of the aircraft and transmits faults or warning messages to ground control in order to provide timely access to maintenance documents and troubleshooting steps which are prioritized by likelihood of success. Boeing has a similar system entitle Airplane Health Management (AHM) which is used on 2000 aircraft and 53 customers. These systems are estimated to increase aircraft availability by up to 35%, enabling capacity increases without the cost of purchasing extra aircraft while minimizing the impact and cost of flight delays due to unscheduled aircraft maintenance issues.
Additional insights:
- How artificial intelligence could lead to self-healing airplanes

At British Petroleum (BP), predictive models can calculate the probability of anything from machinery breaking down to mistakes being made. As long as critical refinery or drilling units are all sensored and monitored in real-time, the resulting data can be analyzed on the fly. Patterns in vibrations or temperature can be fed in real time into algorithms that compare data from previous episodes or to manufacturers’ guidelines in order to predict future maintenance needs prior to a failure causing unscheduled downtime.

Additional insights:
- Predictive Analytics for Energy: Refine your Processes
- The BP Group Launches Smart Integrated Building Technologies Division
- Monsanto bets nearly $1 billion on big data analytics

The Director of the Department of Engineering at The Joint Commission spoke at AAMI 2015, in which he recommended the use of predictive medical equipment maintenance to help busy healthcare technology management professionals. He referenced ANSI/AAMI EQ56:2013, Recommended Practice for a Medical Equipment Management Program as the guiding standard for any healthcare organization developing a predictive maintenance program.

Additional insights:
- New Standard Provides Guidance on Medical Equipment Maintenance Strategies

3. REDUCING PREVENTABLE READMISSION
The interest in the ability to predict readmissions has increased since CMS established the Hospital Readmission Reduction Program (HRRP) as part of the Affordable Care Act (ACA) in 2009 and beginning with discharges as of October 1, 2012. Hospitals are starting to see their revenue erode as the ever-increasing penalties have been assessed over the past several years: 1% of payments in FY2013, 2% in FY2014 and 3% in FY2015.

However, the internet is also rife with many examples of hospitals successfully implementing models that help identify at-risk patients and thereby effectively decrease readmission rates. Each goal of the models is the same, but the methodology varies. Here is a list of several factors to consider when developing a readmission risk predictive model (RRPM):
• **Complexity** – *Simple (previous utilization) to intermediate (LACE) to difficult (deep learning models).*

Penn Medicine was able to drop readmission rates 2-3% by using one simple variable: previous utilization. Expanding to four variables, LACE has become a widely used model and also the comparison standard to determine efficacy of newly-developed models in studies such as the KPSC model or institution-specific risk models. Last year, IBM incorporated deep learning in their commercial version of Watson, SAS integrates deep learning process in their predictive model software. Costs to implement and maintain these models increase with the degree of complexity.

- Improving the Efficacy of Predictive Models
- Case Study: A Readmissions Prediction Model Using LACE

• **Measure Sets** – *Condition-Specific, Procedure-Specific, Hospital-Wide All-Cause Unplanned Readmission (HWR).*

Many predictive models are based on the CMS readmission measures which are comprised of five condition-specific codes...

1. Acute myocardial infarction
2. Heart failure
3. Pneumonia
4. Chronic obstructive pulmonary disease
5. Stroke

...and two procedure-codes:

1. Hip replacement
2. Knee Replacement.

A predictive model for coronary artery bypass grafting (CABG) will be introduced in 2017.

For HWR, studies have shown that comprehensive institution-specific risk prediction models tend be better predictors for potential patient readmissions rather than using a one-size fits all model.

• **Data Source** – *Administrative Claims, Clinical EHR, Disease Specific System, Pharmacy Meds, Health Information Exchange.*

The CMS condition/procedure-specific readmission rates are claim-based via administrative diagnostic codes which can be downloaded from CMS’s Readmission Reduction Program webpage. Claims data used to calculate expected vs actual have been used to create regional prediction models. Using generalized risk factors and local
data derived from a clinical EHR are useful in developing a real-time predictive model in a **community setting**. UPMC Shadyside pharmacists use predictive models from various data sources to identify high risk patients and provide enhanced interventions.

> “The tool continually calculates a patient’s readmission risk based on the following factors: hemoglobin and sodium levels, whether the admission was elective or the patient underwent a procedure, the number of prior admissions this past year, the length of stay greater than 4 days, admitting service, and whether the patient is going to be discharged from an oncology service,” said **Amber O’Malley, Pharm D.**

For a transient population in large metropolitan areas, access to data across multiple facilities is key in pulling all-inclusive information. One recent study examined whether Health Information Exchange (HIO) data can be used to determine readmission risk. Results found that it is feasible if hospitals are actively engaging in the health data exchange.

- **Number of Models / Algorithms and Sample Size.**
  While multiple models do better than single models, a balancing act is needed between number of models and sample size when developing models internally. Breaking down the total inpatient population by specific diagnoses increases the number of models but will result in smaller sample sizes that may compromise the effectiveness of the model. Increasing the time frame may help strengthen the sample size but can also introduce other variables that need to be incorporated or considered before developing the model. Other variables can include improved technology, techniques, medications, operations and systems.

- **Discharge Facility or Destination – Acute Care Facility.**
  The quality of care provided by discharge facilities varies by region/institution, which is another reason why institution-specific models are more reliable predictor for readmissions when taken into consideration. For instance, several studies have shown that surgical patients may benefit from going to an alternative facility before going home. In one study, patients with total hip replacements discharged to an inpatient rehab had the lowest risk for readmission within 180 days. Quick improvements can be made by using the study’s criterion to identify which patients would have benefited by going to an inpatient rehab instead of home with home health or SNF. Refinement of the model can be made through adjustments to variables that affect readmissions in your specific institution. In another study of general medicine, patients concluded that even though patients understood their discharge plans, they had difficulty in adhering to or carrying out the plan. Only a small percentage of patients ascribed their readmission...
to faults in the discharge facility. In this case, determining how much instruction and follow-up is needed prior to discharge will help decrease readmissions.

“Potentially avoidable hospitalizations among nursing facility residents stem from multiple system failures, including inadequate primary care in nursing facilities and lack of skilled staff in nursing facilities,” said Emma Sandoe, a policy analyst.

- **Other Factors / Variables To Consider in Model Development**
  - Patient Characteristics: Socioeconomic status, community-level resources, patient compliance, patient family, patient frailty
  - Quality of Follow-up Provider
  - Comorbidities
  - Quality of Discharge Facility
  - Medications

- **Readmission Initiatives / Regulatory Models**
  - **RARE**: Reducing Avoidable Readmissions Effectively—Minnesota hospitals and care providers identified 5 key areas that prevent readmissions:
    1. Comprehensive discharge planning
    2. medication management
    3. patient & family engagement
    4. transition care support
    5. transition communications
  - **STAAR**: State Action on Avoidable Rehospitalizations—Massachusetts, Michigan, and Washington four year initiative was to “reduce rehospitalizations by working across organizational boundaries and by engaging payers, stakeholders at the state, regional and national level, patients and families, and caregivers at multiple care sites and clinical interfaces.”

Findings:

1. Improve transitions of care by cultivating a cross-continuum learning collaborative.
2. Engage state-level leadership to understand and mitigate systemic barriers to change.

- **HSR&D Collaborative Research to Enhance and Advance Transformation and Excellence (CREATE) Initiative**:
  - From the website: “CREATE is defined by a group of coordinated research projects conducted in a focused research area by independent, collaborating investigators and VA partner(s) who may be local, regional, or national clinical, operations, or other healthcare system stakeholders. CREATEs that target critical areas of interest for Veterans..."
healthcare, including post-traumatic stress disorder, long-term care, women's health, pain management, patient-aligned care teams, and substance use disorders. See below for descriptions of each funded CREATE, as well as individual CREATE projects.”

Contributors
This paper is a product of the HIMSS C&BI Committee and its Task Forces.

Committee Chair
John Middleton, MD, MSc
VP/CMIO
SCL Health

Committee Vice-Chair
David Butler, BSME, MBA, FHIMSS
President
Heartland Innovations, LLC

Its contributors include:

Workgroup Lead
Louise Sulecki, MBA
Systems Analyst III
Cleveland Clinic

Authors
David Butler, BSME, MBA, FHIMSS
President
Heartland Innovations, LLC

Deborah Rupe, RN, MS, FHIMSS
Clinical Analyst
Shriners Hospital for Children

Raj Lakanpal, MD, FRCS, FACEP
CEO
SpectraMedix

Louise Sulecki, MBA
Systems Analyst III
Cleveland Clinic

HIMSS Staff
Shelley Price, MS, FHIMSS
Director, Payer & Life Sciences
HIMSS

For more information about this or other HIMSS Clinical & Business Intelligence resources or activities, please contact Shelley Price, Director, Payer and Life Sciences, HIMSS at sprice@himss.org, @SPriceHIMSS