splunk>

WHITE PAPER

USING HEALTHCARE MACHINE DATA FOR OPERATIONAL INTELLIGENCE

Introduction

The seemingly endless number of patient forms and documents filled out at hospitals, doctor's offices and clinics are a large part of the ongoing healthcare data generated today that becomes part of our permanent health record. But this is just the tip of the iceberg. Behind the scenes, each MRI procedure, pharmacy transaction, CAT scan, EKG record, insurance claim, billing record and blood analysis represents a huge amount of data that sits "below the waterline" and out of sight.

Healthcare data is generated by numerous systems and in a wide variety of formats—syslog, custom application logs, XML, HL7 and myriad other formats. Add to this business vertical an IT vendor technology landscape that is influenced by mergers, acquisitions and disparate and conflicting development processes. It's no surprise that most healthcare applications do not conform to a single data format. With so many unique formats to contend with, managing this data and deriving value from it represents an ongoing struggle for healthcare industry IT professionals.

The breadth of this data is one facet of the challenges facing healthcare organizations; the scale of this data is massive and preventing misuse is no small undertaking.

The U.S. healthcare system is under scrutiny like never before. The passage of the HITECH Act as part of the American Recovery and Reinvestment Act of 2009, the Affordable Healthcare bill in 2010 and President Obama's executive order signed in 2009 that places a bounty on healthcare fraud have begun to change this $3.0 trillion industry. These laws and the executive order challenge the healthcare sector on three separate fronts:

- Reducing fraud and billing errors
- Improving patient outcomes
- Supporting regulations to move from predominantly paper records to electronic heath records (EHR) and to ensure patient privacy

The 2009 economic stimulus set aside over $38 billion to be distributed as assistance grants to healthcare providers in support of organizations moving from paper records to EHR based on a "meaningful use" criteria set by the Department of Health and Human Services (HHS). These challenges have healthcare organizations reexamining business processes, questioning the volume of custom applications in use and looking for ways to deploy more effective data management strategies to meet new mandates.

Additionally, in 2013 the Department of HHS issued the Omnibus Final Rule, the final HIPAA amendment related to privacy and security protections for health information. The rule further increases obligations, enforcement and financial penalties for non-compliance.

Furthermore, Medicare Access and CHIP Reauthorization (MACRA) commenced on January 1, 2017. MACRA, established by the Centers for Medicare & Medicaid Services (CMS), ties medical reimbursements to improved care and better outcomes for patients, while lowering costs. Fees paid to physicians will be scored based on performance, quality metrics and HIPAA security risk—all in an effort to move to value-based care. Physicians who participate in the MACRA Merit-Based Inventive Payment System (MIPS) will be scored on their use of EHR systems and must prove that patient information is protected.

This paper provides a glimpse into some of the key IT and business challenges facing healthcare providers today and introduces a new approach for data management. While it is critical to achieve HITECH and HIPAA compliance, access to records cannot get in the way of providing healthcare in emergency situations or even in the normal course of business.

Reducing Fraud and Billing Errors

The cost of healthcare continues to rise. Healthcare costs in the U.S. have exceeded $2.9 trillion since 2014 and continue to spiral out of control. Of that, approximately $145 billion is estimated to be lost to fraud and another $609 billion due to billing errors. A portion of these errors is attributed to hospital bills that fail to be properly audited by insurance payers prior to payment.

Collaborated, Organized Fraud Schemes

Insurers process tens of millions of dollars of claims per quarter. On average they have seven days to either pay the claim or flag it as potentially fraudulent. Medicare alone pays $1.5 billion in claims per day to 1.5 million providers nationwide, based on 48.7 million beneficiaries. Given the volume of claims and the short window for flagging them, it is very difficult for insurers to consistently catch fraudulent behavior. It is estimated that only 1 percent of fraud is caught. This high success rate of Medicare/healthcare insurance fraud has spawned organized attempts to defraud the system. In many cases, by the time the insurance companies can follow up on potentially fraudulent claims, criminals have closed up shop—with millions of dollars—all paid for courtesy of insurance premium payers.

In recent testimony to congress, Lewis Morris, Chief Counsel to the Inspector General, U.S. Department of HHS shared an organized Medicaid fraud scheme. "This sophisticated group stole identities' of Medicare beneficiaries and doctors licensed to practice in multiple states. The group set up bogus clinics, opened bank accounts to receive funds and submitted fraudulent claims for services never provided. The funds received from Medicare were quickly withdrawn and laundered; sometimes sent overseas."

Organized criminal fraud isn't the only concern. Across the country, seminars on "How to make your practice more profitable" are advertised to small practices and medical groups. These seminars provide the latest information on how to "upcode" procedures to increase insurance reimbursements. Insurers "see" this happen when they notice claims from a cluster of zip codes suddenly start appending procedure code 59 to all their regular procedure codes.

Billing Errors

Not all overpayments are due to fraud. Pat Palmer, founder of Medical Billing Advocates of America, estimates that she finds "errors in eight out of every 10 hospital bills" she reviews. In many instances inadvertent human mistakes in coding or mismatches between systems can cause errors. The most common billing errors are:

- **Duplicate billing** – Patients are charged twice for the same service, supply or medication
- **Additional days in hospital care** – The days that patients spend in hospital care are recorded incorrectly (for example, the day a patient is discharged from the hospital is counted as a day in hospital care)
- **Incorrect room charge** – Patients are charged for a private room when they occupied a semi-private room
- **Additional operating room time** – Patients are charged for more operating room time than what actually occurred
- **Upcoding** – Patients are given lower-cost services or medications but are charged for a more expensive alternative
- **Canceled work** – Physicians order expensive tests and then cancel them, but patients are charged anyway
- **Services never rendered** – Services, treatments or medications that patients never receive
- **While not necessarily intentional, human errors in charting data and system misconfigurations** can send wrong information to billing systems. This has the potential to erode patient trust in the hospital and the healthcare system at large.

Insider Fraud

Approximately 15 percent of healthcare professionals go through some form of prescription drug dependency during their careers. With several million healthcare professionals in the U.S. alone, the potential exposure to fraud is in the hundreds of thousands. Effectively monitoring the dispensing of prescription drugs can reduce the potential for loss and contribute to improvements in patient care.

Improving Patient Outcomes

Improving patient outcomes is not really a new challenge but rather a healthcare industry charter. Both for-profit and non-profit hospitals and clinics struggle with balancing cost and service without compromising patient care.

Patient health data is "big data" and is getting bigger. Petabytes of intensive care unit (ICU) data and patient data are available in facility systems, but according to the Veterans Health Administration, this data is not mined for patterns. Pattern discovery can reveal trends to watch out for and suggest possible courses of action to take once a patient's condition begins to deteriorate.

Another opportunity area for healthcare is monitoring the delivery of patient services, which isn't currently measured. As a patient service is administered, time and date notations are made to the patient's electronic chart. Staffing issues and an overload of patients can cause delays that can affect patient health. If each of these events—doctor diagnoses, services prescribed on a schedule and services delivered—is monitored as a transaction, a correlation of this data can indicate the efficiency of the care measured as time-sequence event. These events can be trended over time to understand and compare efficiencies between different staffing shifts and different individuals. It is understood that in a dynamic hospital environment, "things happen," but over the longer term behavior patterns should begin to emerge—revealing patient care efficiencies and opportunities for improvement.

Ensuring Patient Privacy and Regulatory Compliance

The HITECH Act passed as part of the U.S. government stimulus in 2009 offers financial incentives for healthcare providers to move toward electronic health records (EHRs) and away from paper records. Healthcare providers must ensure and enforce the security and protection of patient records. As such, there are penalties and fines for inappropriate patient data exposure. The law allows these fines to be used to help finance additional HIPAA audits. In this way, agencies are incentivized to perform more audits, find violations, and expose data breaches when they occur. Being HIPAA compliant and keeping patient data private is a primary objective for healthcare providers and their business associates. There are several high-risk HIPAA scenarios that should be a primary concern for healthcare providers.

- Employee as patient (user and patient same name/same billing or home address) self-prescribing treatment
- Family member viewing patient records
- Neighbor viewing patient records
- Healthcare personnel snooping of VIP/executive patient records

Restricting access to patient records is a vexing problem. It goes beyond implementing a robust access management system. A healthcare facility is in the business of providing care without placing any restrictions on who can provide care to a patient in an emergency. While it is critical to achieve HITECH and HIPAA requirements compliance, record access cannot get in the way of providing healthcare in emergency situations or even in the normal course of business.

Operational Intelligence for Healthcare

When looked at from 10,000 feet, the three issues outlined in this paper are not all that different from those faced by many businesses. There's always a need to raise revenues and lower expenses, provide better service and protect the organization from issues such as data breaches or violations that might result in fines. What many organizations don't know is that their machine data is linked to these goals and provides a definitive record of all operational activities.

Operational Intelligence is a real-time view into the operations of an organization—understanding the "who, what, when, and why" of activities within a context unique to the industry. The record of human-to-machine and machine-to-machine interactions plays a critical role in understanding what's really

going on at any given time. Time-based correlations between sources of machine data can be applied and metrics can be gathered. Further enhancement of the information can be achieved by adding other sources of authoritative information within the business, such as HR employee data, IDC-10 codes and asset database information.

In her special report called "The New Realities of IT," Gartner VP Distinguished Analyst Kathy Harris explained how the role of IT is expanding and becoming a strategic partner in helping the business to achieve its goals: "Triggered by significant changes in the business climate and environment, the new realities of IT are rapidly erasing the line between IT and the rest of the business. In this new normal, IT is pressured to do it all—innovate, drive growth, optimize costs, understand and manage risk—and to refocus governance toward business change instead of IT supply and demand."

The business of providing healthcare exemplifies these new and near-term challenges for CIOs, CTOs and the challenges they face. Only Splunk software and its unique analytics engine for machine data delivers the Operational Intelligence needed to get there.

Real-Time Fraud Detection

In testimony to congress, Louis Saccoccio, Executive Director, National Health Care Anti-Fraud Association said, "Clearly, the only way to detect emerging fraud patterns and schemes in a timely manner is to aggregate claims data as much as practicable and then to apply cutting-edge technology to the data to detect risks and emerging fraud trends. The "pay-and-chase" model of combating healthcare fraud, while necessary in certain cases, is no longer tenable as the primary method of fighting this crime."

One way to begin to tackle fraud perpetrated from external sources is to start analysis and validation of healthcare claims for patterns of incongruence in the claim itself. In many instances, fraud perpetrators do not have the background or the time to carefully fill out the claims they issue to insurance companies. These perpetrators know that attempts will be made in the old "pay-and-chase" model to eventually track them down, so speed is more important than accuracy.

Claims metrics can be used to watch for abnormal numbers of specific types of claims over specific periods of time. Jumps in the numbers of prescriptions, specific treatments and/or the number of doctors prescribing for a single treatment can also mean potential fraud.

Claims processed for approval follow a process of acceptance or receipt, adjudication and then payment. There are several solutions that will monitor claims patterns post-payment; there are just a few that will monitor post-receipt and pre-adjudication. There are none that can monitor for fraud patterns in real time in electronic reimbursement forms while looking up data from other systems that have additional caregiver or authorized lab information.

Enter Splunk

Splunk software is a massively scalable real-time engine for machine data. The Splunk platform provides the ability to collect virtually any type of ASCII machine data, watch for data patterns, perform real-time analytics and correlation and has the ability to look up data from other third-party sources and systems. It can also enable the rapid analysis of claim forms, including:

- Monitoring and correlating the doctor's specialty with the treatment code, watching for logical mismatches
- Watching for specific treatment codes that would mean the patient was elderly and correlate against the age of the patient for outliers
- Watching for spikes in the amount of specific kinds of treatments against a rolling 30-day average
- Watching for multiple payment requests with the same patient address
- Watching for multiple payment requests for the same treatment for the same patient from different doctors

- Qualifying healthcare providers who are "new" to the system
- Watching for claims that may be from the same doctor from several service addresses over a given period of time
- Monitoring the amount requested to be paid vs. an average cost of the typical service
- Watching for the same physician requesting payments to many different banks
- Monitoring for the same physician using multiple places of service
- Monitoring the place of service to see if it has been used for previously detected fraud
- Examining the physician or clinic name to determine if either might be on a government services exclusion list
- Qualifying the patient name to determine that it is or is not the same as on a prior investigation or on a Medicaid payment suspension list

Many of these examples support what is called by Omar Perez, Assistant Special Agent in Charge, Office of Inspector General, HHS, an "investigative snapshot." Claims data can help identify important information in assessing whether a fraudulent scheme is underway, including:

- Total amount paid
- Dates of service
- Referring/ordering physician
- Beneficiaries
- Claim dates
- Types of procedures billed
- Place of service
- Provider banking information
- Ownership status

Monitoring and correlating claims data is also in keeping with the first four of the Department of HHS Center for Program Integrity's vision for combating fraud.

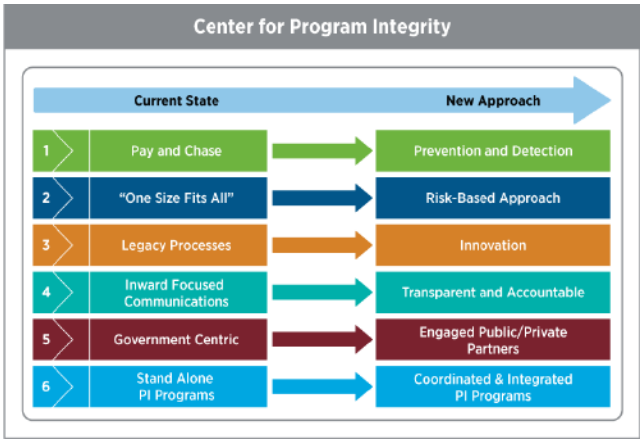

Figure 1. Exhibit presented in testimony to congress March 2, 2011, by John Spiegel, Director, Medicare Program Integrity Group, indicating the new approach in Medicare fraud.

Reducing Billing Errors

A hospital or other healthcare facility represents a complex system of integrated devices and applications that all need to communicate what services have been rendered to the patient. All of these systems communicate to the billing system using a common language called Health Level 7 (HL7). HL7 is an open and extensible standard used by 90% of the information systems vendors serving healthcare. HL7 is the language of billable healthcare events.

In nearly all medical-care facilities worldwide, any billable event in a healthcare IT application generates an HL7 event and that event is transferred to a billing system. It's important to note that an HL7 event can also be the refill of an automated pharmacy, the indication that the patient has health insurance, the addition of patient notes, or other hospital inventory event.

With all the IT systems in a hospital, a single error performed in data entry or system misconfiguration can result in billing errors, resulting in a patient being over or undercharged for a treatment or an entire hospital stay.

Enter Splunk

Splunk software can act as a checks and balance system to compare what was charted to what was billed. Splunk can also be configured to do statistical analysis over time for a specific procedure offered to many different patients looking for billing outliers beyond specific tolerances. In addition, Splunk software can monitor all the systems sending HL7 data monitoring transactions between systems to reduce the most common billing errors and address the following issues:

- **Consistent nomenclature** – Ensure that treatment codes match from one system to another
- **Receipt of data** – Systems acknowledge receipt of data and alert when this doesn't occur
- **Bottlenecks** – Reduce potential bottlenecks between systems that might affect system resiliency
- **Duplicate billing** – HL7 transactions, recorded as log data, are sent to the billing system and are monitored for duplicate billing events for the same patient. This can result from having the patient represented by their full name in one system, first initial and last name in another system and last name, first name in another
- **Number of days in care** – Charting data can act as a check to know when a patient is to be released and can be compared to patient billing data
- **Incorrect room charge** – Patient location information in charting data can be compared to billing information as a double check
- **Operating room time** – Compare billing information to the anesthesiologist's records can reveal differences in operating room
- **Upcoding** – Comparing data from the automated pharmaceutical dispensary (or actual pharmacy) to the charting data can indicate whether a less expensive medication was administered in lieu of a pricier brand name
- **Canceled work** – Physician charting data should be compared to other data where the particular service was to be rendered. For example, an MRI ordered for a patient should mean a record at the MRI machine
- **Consistency in treatment** – The same procedure for two different patients with the same condition with the same demographics information may produce widely varying billing amounts
- **Billing outliers** – Splunk can be configured to do statistical analysis over time looking for billing outliers beyond specific amount tolerances and flag those for follow-up

Monitoring Drug Dispensing

It is critical for healthcare facilities to know how, when, how much and to whom medications are being given. Drugs to be given to the patient are charted by the doctor to be given in certain quantities at specific times. Correlations of charting data with data from automated pharmacies can examine the quantities retrieved and match the amount and type prescribed by the doctor. Splunk can provide real-time alerts to monitor for drug dispensing mismatches and discrepancies.

There is also a need for monitoring the electronic records that are created when medications are received at a loading dock or removed from medical storage that requires badge access. These medications are distributed throughout the facility and checked at specific locations. The purchase records, amounts received and the distribution records need to be compared to protect against loss or theft.

Enter Splunk

Splunk software can be used to correlate the electronic pharmaceutical records of what the healthcare facility received, what is in current inventory, what drugs hospital personnel have collected from automated pharmacies and what has been dispensed to patients based on electronic charting data. Patterns of abuse by certain staff

members can be followed for possible human resource actions. Also, by comparing chart data, when medication was prescribed versus when the medication was administered, we can watch for patterns that may reveal service anomalies based on shift. This information may lead to the review of staffing levels or management issues.

Using System Data for HIPAA Compliance

The HITECH Act put real teeth into HIPAA. Data breaches now cost healthcare organizations real money. Agencies tasked with enforcement get to keep the fines they collect to help fund audits. In essence, agencies are incentivized to go out and find issues—especially when there can be large sums of money involved. The table below represents the change in the fine amounts from pre-HITECH to post-HITECH.

	For Violations Occurring Prior to 2/18/2009	For Violations Occurring on or after 2/18/2009
Penalty Amount	Up to $100 per violation	$100 to $500 or more per violation
Calendar Year Cap	$25,000	$1,500,000

The unique charter of the healthcare vertical means that very few (if any) controls exist on who can view patient information. While most off-the-shelf solutions offer a "one-or-zero" approach to data access (either you have access or you don't), this really isn't a workable model. For the healthcare vertical, access needs to be broken down into "qualities of access." These can be broken down into four areas:

- **Temporal** – Is the person viewing the patient records actually on-duty or on-shift?
- **Physical** – Is the person viewing the records in the facility?
- **Situational** – Is there a positive assignment of the caregiver to the patient?
- **Appropriate** – Is there a relationship between the caregiver and the patient that should preclude the care-giver from seeing the patient's data?

Enter Splunk

Splunk software can preserve data integrity while normalizing specific actions across data types. For example, record access can be reported in many ways by the myriad of custom and off-the-shelf applications in a healthcare facility. The Splunk platform can report all of these record accesses across all system data using the single term "patient data access," without regard to whether the data was from a charting application or a CAT scan system. This makes presenting the patient with an audit of all accesses of their health data a single report from a single system.

Splunk's search language facilitates data correlation in five different ways:

- **Time based** – Identify relationships based on time proximity or distance
- **Transaction based** – Track a series of related events together and display a single event and produce a "duration," "event count" or both
- **Sub-searches** – Taking the results of one search and using them in another
- **Lookups** – Correlations of data to external sources
- **Joins** – Support for SQL-like inner and outer joins

Using data correlations, Splunk software can continuously monitor access to private data and the qualities of access alerting on potential misuse and trending over time to see if inadvertent access leads to a pattern of abuse. Solving this issue means correlating the right data sets.

Examining the Qualities of Access

The **Temporal** quality means bringing together the EHR access information with time management system information that would indicate the shift times of the individual. Splunk's ability to perform time-based correlation of record access while reviewing time management data can isolate off-hours views of patient data.

The **Physical** quality means correlating badge access data with medical record access data to know whether the person is in the facility at the time they viewed the patient data. A record access without a corresponding physical access log can mean credential sharing among the staff, or worse, identity theft.

The **Situational** quality refers to whether a positive assignment of the healthcare professional to the patient has been established. The correlation of medical records access and charting data would be one way to confirm the situational quality and a way to find casual record viewing by staff members.

The **Appropriate** quality of access is the appropriateness of the record viewing based on a relationship between the patient and the caregiver. This means looking at the access data and comparing a combination of patient information (name, address, city, state, zip code, etc.) to that of the patient to determine family or neighbor relationships that can mean a HIPAA violation. The patient data can be compared to data available about the caregiver contained in HR databases to determine inappropriate patient data views.

Understanding and monitoring access for these qualities can mean the difference between convincing an auditor that the right person saw the right data at the right time to provide the right service versus, a finding that can affect the organization's reputation. Monitoring and correlating data watching for these qualities of access is the only way to reveal whether or not patient record access is appropriate while not interfering with the "do no harm" healthcare charter.

Conclusion

The healthcare sector is facing many challenges. Costs continue to rise, there's a constant demand for efficiency and service improvement and new regulations protecting patient data are being heavily enforced. Ransomware is also top-of-mind problem for the healthcare C-suite, as several hospitals fell victim to such attacks in 2016. Yet machine data, the transactions between humans and machines and machine-to-machine interactions, contain most of what we need to take a few steps in the right direction to ensure patient information is protected.

Bringing together the terabytes of structured and unstructured data already available from the applications in a healthcare facility for analysis allows Splunk to provide insights into some of the most pressing challenges facing the healthcare sector—lowering costs, providing better patient outcomes, protecting privacy of patient information, catching fraud and achieving HIPAA compliance.

The chief privacy officer, the CIO, the CSO, and ultimately the CEO and board of directors all benefit by unlocking the value of the machine data they already have and correlating it as needed to better understand privacy issues, find efficiency, recover revenue, and combat fraud and waste. Better utilization of system data in the healthcare vertical provides Operational Intelligence needed to face the sector's current challenges and facilitate better business decisions.