The Power of Combining Big Data Analytics with Business Process Workflow

In this paper, we demonstrate how the combination of big data analytics, business process workflow and smart people with the right skills can deliver sustained and measurable success.

This can be achieved in many different industries. We take as our example the health care payments sector, where nearly $3 trillion dollars in payments flow from health insurers to providers every year in the U.S. alone. While most of these are straightforward reimbursements for agreed upon services, a proportion should not have been paid in the first place and need to be recovered.

Having a partner act as an intermediary to audit and recover these improper payments has many advantages. A crucial advantage is that the business process service (BPS) partner is also in an ideal position to harness big data analytics to produce improvements in recovery rates—for typical improper payments cases as well as instances of outright fraud—while minimizing the impact of audits on providers.

TABLE OF CONTENTS

Introduction

Big data is a hot topic across executive suites today. The amount of data that is available to businesses is increasing exponentially, with social media and machine-to-machine as just two of the leading sources. Storage is getting cheaper, driven by cloud computing, and processing power is getting faster.

How is all of this big data being structured and analyzed to deliver real business benefits? What are these tangible, measurable benefits?

The truth is that many organizations are already successfully harnessing the power of big data. It's best done where innovation is always done best—in the context of a real business problem to fix, with smart people who can structure the necessary data logic and interpret its results, and a business process workflow that turns these results into desired business outcomes.

In this white paper, we take an in-depth look at the impact of combining smart people, big data analytics and business process workflow in one particular market sector —the auditing and recovery of improper health care payments by insurers to providers. If you work in the health payment sector today, you will find insights that will directly impact your business. If you are in a different industry, we believe you will find the health payment example relevant for understanding how the combination of smart people, big data analytics and business process workflow could translate into cost-effective solutions for your own organization.

Audit services in health care

Health care is delivered in the U.S. and many other countries under variations of the same model, despite the complexity that can sometimes be apparent. An individual enrolls in a health care plan (which may be a government sponsored or private plan), visits physicians and other care providers, and perhaps has a hospital stay and takes medications.

The physicians, the hospital and the pharmacy are all providers of services and make claims on the health care plan for payment. This is happening across the world, millions of times a day. About $3 trillion in payments flows from health insurers to providers each year in the U.S. alone. By any measure, it's big business. While most of it is proper and accurate, the hard truth is that inaccurate, improper and sometimes fraudulent claims are made.

Let's assume, for example, we have a claim for a physician's appointment. The claim form says the visit lasted one hour. How can this be verified? What other data might be available to check that the claim is truthful?

Let's take another example: two claims, filed one after the other, for the same patient to have an appendix removed. It's easy to spot the impossibility here, making this a clear cut case of an improper claim.

As a final example, take the case of a medical group in a small town whose claims come mostly from patients with addresses far away from the town—much farther than is typical for similar providers. Does this group offer excellent services that explain the long trips to visit its doctors, or is something else going on?

In the U.S., many health care plans—both public and private—engage specialized auditors to review claims, find improper ones, and either prevent a payment before it takes place or recover a payment that has already been made.

The audit process

When a claim is made, there are typically two stages involved in processing the necessary data, especially in the case of a U.S. government sponsored plan. The physician will write the

medical records, and a coder will review those records and apply the appropriate code to the claim. The claim is then routed to the health care plan for payment.

A health care plan will typically audit a sample of claims to identify any that are improper, either on its own or through a third party specialist. CGI is one such third party auditor, working on behalf of a range of health care plans in the public and private sectors.

CGI uses its own software applications to audit these claims. Claims are run through the system and tested against business logic to identify counter-to-policy, unusual or logically impossible combinations of procedures and codes. Each item of business logic is called an "edit." The edits are frequently updated, taking into account recent trends in improper claims and the latest research.

At the same time, a broader predictive analysis is conducted to identify trends in the filing of claims by certain types of physicians or for certain types of procedures and to identify outliers that may suggest possible improper claims for which more investigation is required. There are many roles in the claims audit and analysis process; the two we will refer to most often in this paper are the auditor and data scientist roles:

- **Claims auditor:** Examines claims flagged as potentially improper, gathers additional information as necessary and recommends recovery action. This individual is skilled in the relevant process steps and policies and has specialized knowledge in medical coding and clinical practices.

- **Data scientist:** Collects and analyzes data from large numbers of claims, audits, and recovery results. This effort generates useful insight into trends and patterns that can be put to work in making improvement to edits at the most basic level and also to create capabilities for further advances through pattern matching, predictive models, geospatial mapping and social network analysis.

Claims that hit one or more of the edits are flagged in the system, indicating a heightened probability of error or impropriety. If the nature of the data suggests that the probability of an improper claim is very high—for example, in the case of claims for two appendices being removed from the same patient—then CGI's software routes the claim through an automated workflow process and a notice is automatically sent to the provider. This business process workflow is key to ensuring that necessary recovery actions are taken. Of course, the automation ensures that this is done cost-effectively, without diverting the focus of the audit team from more complex cases.

If the probability is less certain, then the system will generate an automatic workflow request for the medical records to be sent to CGI, and the audit team will be automatically alerted once these arrive. Experienced medical specialists and coders then review the data and match it against other information—for example, statistics on other claims for similar procedures or the patient's medical records, which average 300 pages.

Taken together, the information sources used in auditing and researching claims amount to big data indeed. There's a high *volume* of data, often in the range of hundreds of terabytes. New claims and medical records for analysis and auditing arrive constantly to, and claimants change their behavior quickly, so the *velocity* of information is high as well. In addition, the claims and medical data contain both structured and unstructured elements, so there is a great deal of *variety* in the data.

Analyzing big data is more challenging than analyzing simple relational (structured) records, but it is worth doing because the additional information locked up in large, unstructured data sources has a great deal of value in determining the validity of claims.

After reviewing the data related to a claim, the auditors will then enter their findings back into the workflow and analysis system, which will initiate the necessary processes—from issuing provider notices to the recovery of funds. The system will take these findings, analyze trends,

4

and use the results to influence the design of future edits. In this way, the system captures what has been learned and enables the team to maximize future effectiveness.

Because of stringent regulatory rules in the U.S. for the maximum time to pay a claim, the majority of repayment claims today are done "post-pay," whereby the provider is typically given the option to offset the money due against future claims, make a repayment, or follow an appeal process. Increasingly, auditors such as CGI are speeding up their processes to catch improper claims at "prepay" time, i.e., before the funds reach the claimant.

This represents a key opportunity to increase effectiveness, as we will discuss later in this document. While the auditing team learns through experience and adjusts the edits and predictive models accordingly, the providers (claimants) also learn. Auditors will offer education services to providers, ideally to reduce the number of improper claims in the future. However, in a scenario where a claimant is actively looking for opportunities to increase the value of claims, the claimant may try new over-claim strategies, so the auditor needs to keep ahead of the game and anticipate these in the edits.

There is a continuum of fraud, abuse and waste that underlie improper claims. An improper claim could be a simple mistake by a coder. To demonstrate real fraud, there has to be intent, and big data itself can rarely demonstrate that intent—this has to be done by smart people parsing and interpreting the data. While true fraud needs to be stopped, of course, usually through legal action, the commercial need is to recover the money in any event, whatever the nature of the improper claim.

The power of analytics

The economics of the situation are clear; there are billions in improper payment to be recovered, and there is a cost to recover them. The most straightforward cost is in the effort it takes to perform the audits and the follow-up actions necessary to recover payments. On top of that, audits provoke reactions in providers, including irritation in complying with auditor requests, known as "provider abrasion," and more subtly the shift in behavior to try to reduce the likelihood of an audit. There are regulatory requirements to minimize provider abrasion as well.

Against this backdrop, more and more advanced analytics have come into play with the goal of maximizing the effectiveness of the claims audit function. Because it's not known whether claims are proper, improper or even fraudulent before they are evaluated, data scientists treat each one as having a probability of recovery and an expected recovery amount. The challenge then is to create audit trigger rules—the edits—which identify claims with a high probability of having a recoverable amount, a high dollar value of potential recovery, or both.

In the simplest sense, any set of edits will produce a number of audits when applied to a group of claims, and the resulting recoveries can be measured. If a change is made to the set of edits, for example by a data scientist proposing a new rule discovered in his or her analysis, that will change the way the system performs and its results.

We can illustrate this with a simple table showing the value of more effective selection of claims for audit via the "swap set." The idea of a swap set is that a change to the set of edits changes the set of audit decisions made, so some claims "swap" from one decision to a different one. Unchanged decisions make no difference. Comparing the results in the swap set will reveal the effectiveness of the changed edits.

Decisions Using Current Edits	Decisions Using Improved Edits		Totals
	Audit	Don't Audit	Totals
Audit	Total Audits: 80,000 Success Rate: 50% Total Recoveries: $40,000,000 Recoveries per Audit: $500	Total Audited: 20,000 Success Rate: 10% Total Recoveries: **$2,000,000** Recoveries per Audit: $100	Total Audits: 100,000 Success Rate: 42% Total Recoveries: $42,000,000 Recoveries per Audit: $420
Don't Audit	Total Audited:20,000 Success Rate:40% Total Recoveries **$6,400,000** Recoveries per Audit: $320	Total Audited: -0- Success Rate: NA Total Recoveries: $0 Recoveries per Audit: NA	
Totals	Total Audits: 100,000 Success Rate: 48% Total Recoveries: $46,400,000 Recoveries per Audit: $464		

Figure 1: Implementing more effective decision criteria (edits) produces a swap set with measureable improvements in results. Drawn from an actual case; figures are examples for illustration only.

In this example, there are 100,000 audits being performed per month using the current set of edits, with recoveries totaling $42 million. By analyzing the patterns of claim characteristics that indicate higher probabilities of improper payments, the audit partner's data scientists are able to sharpen the set of edits. As a result, some claims that would have been previously flagged now pass through without an audit (the red box group in the figure), and others are flagged for audit that would have been passed through previously (the green box group).

The success rate and total recovery amounts are much higher for the green set, so total recoveries go up by $4.4 million to $46.4 million. Because the total number of audits is the same, there is no increase in the operational cost or provider abrasion as a result. That's more than $4 million a year—a more than 10 percent improvement—from one edit change.

Because both the health care plan and the BPS partner have incentives to make sure the process is done as effectively as possible, both cooperate in the search for improvements. An improved edits-and-discovery process will reduce false positives, which will in turn increase auditor productivity, reduce provider abrasion and decrease costs. Whenever ways are found to recover more improper claims, or to do so with fewer or less intrusive audits, both partners benefit as a result.

SETTING THE STAGE FOR EFFECTIVE EDITS AND ANALYSIS

In the early stages of a BPS relationship, we agree on the scope and range of the work to be done by the CGI team. Included in this scope is the definition of the data to which we will have access, the privacy and confidentiality controls, and the responsibilities for reporting and analysis each partner will assume.

Another important step at this stage is identifying the initial set of edits. There are two sources of input for the starting edits: the client organization's policies and CGI's knowledge of edits that have worked in the past for similar organizations.

Once the process is underway, data begins to accumulate showing actual results of edits and various levels of audit and recovery actions applied to the wide variety of claims. At this point, the data scientists can go to work, using their knowledge of the data, the clinical domain, the

An improved edits-and-discovery process will reduce false positives, which will in turn increase auditor productivity, reduce provider abrasion and decrease costs.

existing edit rules, and their ability to creatively construct new hypotheses to test. Their analysis includes the following:

- Details of claims and recoveries
- Trends in the effectiveness of existing edits, especially in terms of the tendency for them to "wear out" as providers change their behavior
- Medical records (scanned paper files and newer electronic medical records, which are used in the audits themselves and in analytics)
- News stories about health care fraud, which give ideas for new hypotheses and may also inspire copycats
- Anonymous tips (these are routed to the data scientists, so they can aid in the follow-up investigations)

Improper medical claims fall on a spectrum from simple errors, to mild types of claim inflation, all the way to large scale and sophisticated fraud operations. These different kinds of claim situations call for different analytical approaches.

Erroneous claims: These are often identified by logical consistency and policy rules. For example, root canals are not performed by ophthalmologists. Because health diagnoses and medical practice are complex by nature, analysis is often needed to discover these relationships.

Ordinary claim inflation: This falls in a grey area between error and intentional misrepresentation. Inflated claims usually make logical sense on paper, but are improper because the claim does not match the reality of the care situation. An example is a set of codes indicating a physician office visit. Sometimes a longer or more intensive visit is claimed, where a short and uncomplicated visit was actually delivered. This is an area where predictive models can be employed to great effect to find these improper payments, where claim characteristics are correlated to a high probability of a recoverable payment.
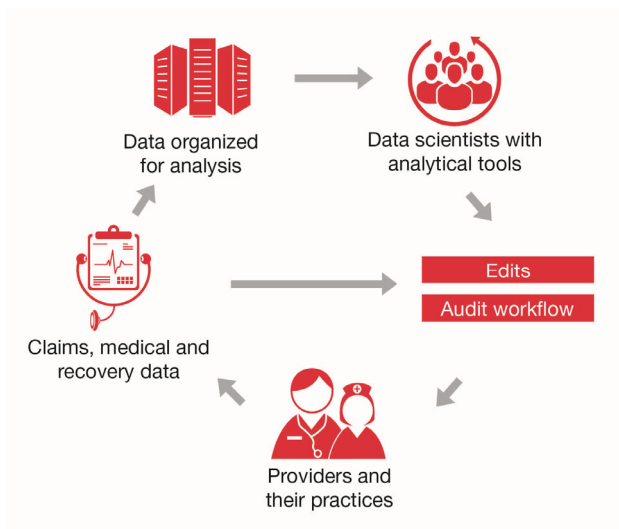
Fraud: Serious cases of intention misrepresentation are difficult to detect, and fall outside the normal activity of claims audit. When situations arise that present reasonable suspicion of criminal activity, those cases are turned over to a special unit for further investigation and possible legal action. The analytical techniques needed to uncover potential fraud are equally sophisticated, including pattern recognition and social network analysis. Even then, data and models can't prove intent, only uncover the evidence.

Just as the types of analysis vary for the different types of improper claims on this spectrum, the possible recovery actions also vary. As we have seen, for simple errors caught by a policy rule, recovery can be automated by generating a letter. Grey-area, inflated claims may require expert review and more information from the provider to determine the right action and amount to recover. When a pattern of similar errors is found at a particular provider, the large number of individual claims may be bundled together into a bulk recovery package, which has the effect of recovering the money that is due without incurring the expense of processing multiple recovery actions.

THE IMPORTANCE OF HAVING A JOINED-UP PROCESS

The claims audit workflow supported by analytics is a closed-loop process:

Improper medical claims fall on a spectrum from simple errors, to mild types of claim inflation, all the way to large scale and sophisticated fraud operations.

Figure 2: Schematic diagram of the claims audit and data analysis functions integrated to form a closed-loop process

It should be clear from these examples that it's the combination of a rules-driven business process, analytics and people with the right skills that drive great results and continuous improvement in recovery auditing. A particular set of edits, predictive models and pattern-matching algorithms may work well for a time, but the world is dynamic and will change in ways that require smart people constantly striving to keep ahead of it.

A BPS partner, such as CGI, is in a great position to help health insurance providers to stay ahead of the game by investing in software, processes and people, and by spreading these costs across multiple clients.

In working with a BPS partner, CGI's clients have found that the nature of their business agreement has a big influence on the success of the business process / analytics collaboration as well. A contract where both parties benefit from increased recoveries as well as from reduction in unnecessary audits, provider abrasion, and other costs provides the harmony of interests that produces the best results.

In summary, a fully effective recovery audit process includes designing an effective set of audit-selection business rules (edits), continually improving them over time by analyzing the audit and recovery results, and then ensuring, through efficient workflows, that the improper claims, once identified, are followed all the way to payment recovery. It's the combination of three essential ingredients—business process workflow, data analytics, and smart people with the right skills.

The future

The worlds of health insurance, business process services, data and analytics are all evolving at a rapid pace. Some aspects of recovery audit operations will remain relatively unchanged over the next few years, including the need for such operations in the first place. As long as there are insurance reimbursements for claims, there will be providers who submit erroneous, inflated or fraudulent claims.

There are many aspects of these operations that are likely to experience the forces of change in the near future, however, including the following:

- **New data:** New data sources will enrich the claims data and scanned medical records used today. Electronic medical records have great promise because they provide an independent and, in principle, an accurate verification of the diagnosis and procedure that can be matched against a claim automatically. Sources like death records, residency records and driver licenses provide a way to verify that patients are who they say they are through list matching and link analysis. Also, in the U.S., the National Claims Benchmarking Data facilitates trending and benchmarking across patients and providers.

- **Text analytics:** Where useful information is to be found in text data sources such as scanned paper records, text analytics methods like keyword extraction and document classification are already starting to be used to improve the efficiency of audits. For example, document classification can automatically take an auditor to the right page to read in the record, where the average medical record is 300 pages long.

- **Predictive modeling:** Predictive modeling is still in its early stages of use for improper payment recovery and fraud detection. We expect this technique to grow rapidly over the next few years.

- **Market changes:** Payment models will continue to evolve into, for example, outcome based payments, which will create new opportunities for over-claiming. Auditors will need to stay ahead of this new game.

- **Regulatory changes:** New coding schemes (for example, the introduction of ICD-10 in the U.S. in October 2014) will increase the amount of data and require the development of new, more complex edits.

- **Increased cyber security threats:** Health care providers, insurers and their partners have a fiduciary responsibility to protect data. As cyber security threats increase, so will the range of measures available to protect against them.

- **Mobility:** With the increase in home care treatment, there comes the challenge of proving that home care workers have fulfilled an appointment as claimed. We expect to see increasing use of mobile technology in this area.

- **Prepay:** The processes and rules described in this white paper relate to the most common form of recovery in the market today—"post-pay" audits and recoveries, or what is sometimes called "pay and chase." Over time, there will be an increasing focus on prepay audits, which must be done within short time limits to comply with regulations, but will prevent many improper payments from going out the door in the first place.

Each of these future developments highlights another reason that smart people who understand the claims audit domain are part of the winning combination. Process discipline and analytics are great for increasing the effectiveness of an existing function. But it takes people who have the ability to envision a completely different way of doing things to change the game rather than to just play the game more skillfully.

At CGI, we firmly believe that there is an opportunity for organizations in many industries to get more business benefits from big data by combining it with business process workflow and people with the right skills.

Conclusion

In this white paper we have described how the combination of business process workflow, data analytics and smart people with the right skills can produce sustained and measurable business benefits. We have used as our example the high volume, fast moving world of health care payments and how CGI is supporting health insurers to audit and recover payments made for improper claims. We have shown how we build continuous improvement into the service and illustrated some likely future developments in the market.

At CGI, we firmly believe that there is an opportunity for organizations in many industries to get more business benefits from big data by combining it with business process workflow and people with the right skills. We trust that the health care example we have given will motivate you to further explore the opportunities. As big data and the techniques to harness it continue to expand, the ability to drive business improvements from analytical insights has become an increasingly vital driver for success.

Whether you are In the health sector or another industry, we welcome the opportunity to discuss how we can help you bring together the power of big data analytics and business process workflow management to address the challenges you are facing. Please get in touch with us at info@cgi.com.

ABOUT THE AUTHORS

Robert Rolf: With more than 17 years of recovery audit experience, including payment consulting, improper payment auditing and related technology solutions, Rob serves as vice president of CGI's Healthcare Compliance practice. He oversees recovery audit and related projects for the Centers for Medicare & Medicaid and several state Medicaid plans. Based on his expertise, he has been called upon to testify before Senate committees and sub-committees.

Christopher Sutton: Chris is vice president of CGI's global BPS practice. Based in London, England, Chris has 15 years of BPS experience gained across North America, Europe and the Asia Pacific. He holds a Master of Arts in Literae Humaniores from Oxford University, is a Fellow of the Institute of Chartered Accountants, and is a Six Sigma black belt.

Matthew Kuckuk: With 25 years of experience in data analytics and business intelligence, Matt leads CGI's global business intelligence and big data practice from his base in San Francisco, California. Matt has Bachelor of Science and Master of Science degrees in Operations Research from the University of Wisconsin and the University of California (UC) and serves on the Industry Advisory Board for the Department of Industrial Engineering and Operations Research at UC Berkeley.

cgi.com