# Chapter 25
# Case Study: Ideal Insurance

**Deepak Agrawal and Soumithri Mamidipudi**

## 1 Introduction

Sebastian Silver, the Chief Finance Officer of *Ideal Insurance Inc.*, was concerned. The global insurance industry was slowing, and many firms like his were feeling the pressure of generating returns. With low interest rates and increase in financial volatility in world markets, Sebastian's ability to grow the bottom line was being put to test.

Sebastian started going through the past few quarters' financial reports. He was worried about the downward trend in numbers and was trying to identify the root causes of the shortfall in order to propose a strategy to the board members in the upcoming quarterly meeting. To support his reasoning, he started looking through industry reports to examine whether the trend was common across the industry or whether there were areas of improvement for his company. Looking at the reports, he observed that both profit from core operations, that is, profit from insurance service, and customer satisfaction rate were surprisingly lower than industry

**Electronic supplementary material** The online version of this chapter (https://doi.org/10.1007/978-3-319-68837-4_25) contains supplementary material, which is available to authorized users.

D. Agrawal (✉) · S. Mamidipudi
Indian School of Business, Hyderabad, Telangana, India
e-mail: a.deepak@outlook.com

© Springer Nature Switzerland AG 2019

B. Pochiraju, S. Seshadri (eds.), *Essentials of Business Analytics*, International Series in Operations Research & Management Science 264,
https://doi.org/10.1007/978-3-319-68837-4_25

standard. The data was at odds with the company's claim settlement ratio,[1] which Sebastian knew was higher than that of his rivals, and claim repudiation ratio,[2] which was lower than industry average. He also observed that claim settlement was taking longer than expected.

The head of the claims department at *Ideal Insurance*, Rachel Morgan, told Sebastian that there was a tremendous shortage of manpower in the settlement team and added that there was no focus on making innovation and improvements in the claim investigation process. She also reminded Sebastian that she had proposed an in-house department of analysts who could help improve the claim settlement process and support other improvement initiatives. Sebastian promised he would review the proposal, which had been submitted to Adam Berger's HR team in the beginning of the year, and set up a meeting with Rachel in the following week. He also asked her to contact an expert in claim settlement and investigation to provide a detailed review of performance and to suggest a road map for changes.

Following Sebastian's suggestion, Rachel reached out to an independent consultant to verify and analyze Ideal's healthcare policy claims. She knew that fraud prevention, one of the biggest reasons for profit leakage in the sector, would have to be a key priority area. However, she was also aware that improving the probability of detecting fraudulent claims could hurt genuine policyholders. The challenge facing Rachel was how to balance the need to deliver swift responses to customers with the knowledge that too many fraudulent claims would severely hurt the bottom line. It was to solve this challenge that the consultant advised Rachel to consider using advanced analytical techniques. These techniques, such as artificial intelligence and machine learning, could help to make claims processing more efficient by identifying fraud, optimizing resource utilization, and uncovering new patterns of fraud. The consultant added that such applications would improve customer perception of the company because genuine claims would be identified and processed more quickly.

## 2 The Insurance Industry

The global insurance industry annually writes trillions of dollars of policies. The nature of the insurance industry means that insurers are incredibly sensitive to fluctuations in local and global financial markets, because many policies involve coverage for decades. Premiums that are paid to insurance companies are therefore invested for the long term in various financial assets in order to generate returns and to use capital efficiently. This necessarily means that understanding the global

[1]Claims settlement ratio is calculated as the percentage of claims settled in a period out of total claims notified in the same period. The higher the settlement ratio, the higher the customer satisfaction.

[2]Claims repudiation ratio is calculated as the percentage of claims rejected (on account of missing or wrong information) in a period out of total claims notified in the same period.

insurance industry involves not just understanding the nature of the companies that operate in the sector but also its interconnections with financial markets.

Perhaps the most important distinction regarding insurance companies is the nature of the policies written. Large (measured by assets and geographical coverage) firms such as AXA and Prudential usually have an array of policies in every segment. Smaller companies, however, may restrict themselves to writing only in the life or non-life segments. Insurers may thus choose only to be involved in motor, disaster, or injury-linked claims. Companies may also reinsure the policies of other insurance companies—Swiss Re, Munich Re, and Berkshire Hathaway all engage in reinsuring the policies of other insurers. The operating results of any specific insurance company, therefore, will depend not just on the geography in which it operates but also on the type(s) of policy(ies) that it underwrites.

Insurance to a layperson is nothing but the business of sharing risk. The insurance provider agrees to share the risk with the policyholder in return for the payment of an insurance premium. Typically, an underwriter assesses the risk based on age, existing health conditions, lifestyle, occupation, family history, residential location, etc. and recommends the premium to be collected from the potential customer. The policyholder secures his unforeseen risk by paying the premium and expects financial support in case the risk event takes place.

The insurance business depends for survival and profitability on spreading the risk over proper mix and volume and careful planning over a long horizon. The collective premium is either used to honor the claims raised by policyholders or invested in long-term assets in expectation of significant profit. Thus, the insurance provider has two major avenues to earn profit, namely, profit from the core operations of risk sharing and profit from investments. It has been observed that profit from core operations is generally very low or sometimes even negative; however, overall profits are high due to investment strategies followed by the firm—such as value investing, wealth management, and global diversification. Most insurance businesses have an asset management entity that manages the investment of such collected premium. The insurance providers also protect themselves from huge losses from the core business by working with reinsurers such as Swiss Re and Berkshire Hathaway, who share the risks among the insurance providers.

## 3 The Healthcare Insurance Business

The health and medical insurance industry is a fast-growing segment of the non-life insurance industry. The compound annual growth rate of the global health insurance market is expected to be around 11% during 2016–2020.[3] The revenue from the global private health insurance industry, which was around US$1.45 trillion in 2016,

[3] https://www.technavio.com/report/global-miscellaneous-global-health-insurance-market-2016-2020 (accessed on Aug 17, 2018).

is likely to double by 2025 (Singhal et al. 2016). While the USA occupies the top rank in gross written premium revenues, with more than 40% of the market share, an aging population and growing income are expected to lead to major increases in the demand for healthcare services and medical insurance in Latin America and Asia-Pacific regions in the coming years. The major driving forces and disruptive changes in health and medical insurance markets are due to the increase in health risk from the rise in noncommunicable diseases and chronic medical conditions, advances in digital technology and emerging medical sciences, improved underwritings and changes in regulatory environments, and increased consumerism and rise in aging population in developing economies.

Health insurance accounts for billions of dollars of spending by insurance companies. The *Centers for Disease Control and Prevention* estimates that about 15% of spending by insurers is on healthcare plans.[4] Advances in healthcare technology are likely to be balanced by increased demand. As populations continue to age in the USA, Europe, and Japan, spending on this sector will remain a cornerstone of the insurance business for decades to come.

Most healthcare insurance divides the cost of care between the policyholder and the insurer. Depending on the type of policy and the nature of treatment, this division can take place in a number of ways. Some policies pay out once spending crosses a certain threshold—called a "deductible." Others split the cost with the policyholder in a defined ratio—called a "copayment."

## 4 Claims Processing

Speed is at the heart of designing the processes in a health insurance firm. By its definition, insurance is required in situations that are unforeseeable. In the case of health insurance, receiving the money that the policy guarantees as soon as possible is vital to the policyholder. Yet this constraint means that the timeline of claims processing must be necessarily as short as possible, hurting the ability of insurance companies to verify claims.

The process for claiming health insurance normally proceeds in several stages. The initial contact between the firm and the policyholder is made via the call center or local office/agent. Often, this step is undertaken by a person close to the holder or by the healthcare provider, because the holder may be incapacitated.

The firm's call center obtains and records basic information from the client, including details regarding the type of policy that the holder owns, the hospital to which the holder has been taken, and the injury/ailment with which the holder has

[4]Eynde, Mike Van Den, "Health Insurance Market Overview," State Public Health Leadership Webinar, Deloitte Consulting LLP, August 15, 2013, URL: https://www.cdc.gov/stltpublichealth/program/transformation/docs/health-insurance-overview.pdf (accessed on May 25, 2017).

been afflicted. Armed with this information, the call-center employee forwards the necessary details to the claims processing team.

The claims processing team is responsible for the first line of inquiry into the claim. They check the policyholder's coverage and expenses, verify the hospital and network at which the client has been admitted, and ask for proof such as bills and prescriptions. Upon doing so, the team either classifies the claim as Genuine, Discuss, or Investigate. Genuine cases are processed without any further clarification. Discuss cases are forwarded to a data collection team in order to collect more information and verify details. Investigate cases are forwarded to the claims investigation team and can take a long time to be processed (settled or rejected) further.

At present, the claim processing team at Ideal Insurance examines the following points to create a fraud score:

1. Is the policyholder raising reimbursement claims from a non-network hospital?
2. Are multiple claims raised from a single policy (except group policies) or policyholder?
3. Are there multiple group claims from the same hospital?
4. Is the claim raised close to the policy expiry date?
5. Is the claim raised close to the policy inception date?
6. Is there no pre- or post-claim to the main claim?
7. Is there misrepresentation of material information identified in the report?
8. Was the claim submitted on a weekend?
9. Are there "costlier" investigations?
10. Are there high doctor fees?
11. Was the claim reported one day before discharge?
12. Did the claim intimation happen after 48 h of admission?

Each indicator carries a weight assigned based on prior research and experience of the investigation team. The maximum weighted score is 40. If the weighted score (or fraud score) of a claim is more than 20, then the claim processing team forwards the claims to the investigation team to investigate potential fraud. If the fraud score is between 16 and 20, then the claim processing team seeks additional data from the information collection team and healthcare service provider. The claims with fraud scores of less than or equal to 15 are considered genuine and forwarded to the settlement team for payment to the policyholder or service provider.

Investigating claims requires firms to verify a host of corroborating details. In order to satisfy themselves of the genuine nature of the claim, investigators check the types of medications prescribed, years of patient history, and the specific nature of the ailment. Depending on whether the type of fraud suspected is hard or soft, investigators could choose to examine different levels of data. While soft fraud might be identified by conclusively proving that certain kinds of treatment were not appropriate for the disease diagnosed, hard fraud would need larger and more complex patterns to be uncovered.

Longer, more complicated, claims processes are a double-edged sword for insurance firms. As the number of claims that are investigated in-depth rises, the

chance of both inclusion and exclusion errors falls. Yet investigating a large number of claims takes up time, resources, and risks, causing delays to genuine customers.

A claim gets further complicated if the policyholder decides to file a litigation suit due to delay or rejection. Though it may create pressure for quick settlement, providing strong argument for minimizing delay/rejection, a suit by itself does not necessarily mean that the claim is genuine. Litigation is a crucial and important tool for insurance firms. As health insurance is vital to most clients, the decision to classify a claim as fraud can potentially open the door to a host of lawsuits. These lawsuits can be on the behalf of either individual customers or a host of clients. It is usually the company's responsibility to justify its opinion that a claim is fraudulent. Because legal standards for fraud may be different from the company's internal standards, ensuring that the company can win such cases can become complicated. In addition, court costs in themselves can be prohibitive. The company may have to follow different rules as prescribed by the law of the respective land. Large number of cases or long pending cases can also potentially damage the firm's reputation. Avoiding such challenges is the best bet. It is also important not to back off from litigation when the occasion demands to prevent potential fraudsters from taking advantage of the firm.

## 5 Stakeholders in the Health Insurance Industry

The transformation in the health insurance industry involves and requires influencing numerous stakeholders and market participants. These include[5] the following:

1. Consumers, patients, caregivers, and patient advocacy organizations: These are people experiencing the health problems and who would be beneficiary of various health services and treatments.
2. Clinicians and their professional associations: These are major medical decision-makers; and their skills, experience, and expertise matter the most.
3. Healthcare institutions, such as hospital systems and medical clinics, and their associations: Major healthcare decisions are structured and determined by choices of institutional healthcare providers as they often have a broad view on what is causing a particular health problem.
4. Purchasers and payers, such as employers and public and private insurers: Coverage by insurer and purchaser of healthcare plays an important role in diagnostic and treatment decisions and choices of insured.
5. Healthcare industry, pharmaceutical companies and drug manufacturers, and industry associations: Manufacturers of drugs and treatment devices and their suppliers and distributors influence the quality of healthcare services available in a region.

[5]Agency for Healthcare Research and Quality (AHRQ). 2014. Stakeholder Guide. https://www.ahrq.gov/sites/default/files/publications/files/stakeholdr.pdf (accessed on Aug 17, 2018).

6. Healthcare policy makers and regulators: Government policy and regulatory mechanisms including the legal system influence cost and quality of healthcare, health insurance market development, and access to individuals and their families.
7. Healthcare research and research institutions: Local availability of research funds and quality of research institutions play a vital role in the development of the healthcare and medical insurance market in a region.
8. Insurance companies: The insurance companies provide coverage to the policyholder in return of an insurance premium. They underwrite the policy and collect the calculated premium based on the customer's risk profile. They are expected to honor the claims raised by a policyholder in case of unforeseen circumstances. The risks are shared by pooling a large number of customers with diverse risk profiles and with the help of reinsurers.
9. Reinsurance providers: The reinsurers provide coverage to the insurance companies in case of large unforeseen situations such as natural calamities, terrorist attacks, and bankruptcy. They assist in sharing the risk across geographies and a diverse pool of risky customers.
10. Third-party assistance (TPA) service providers: TPAs are the mediators between insurance providers and customers or consumers who assist in providing information to the underwriters and help in smooth processing of claims as and when raised. They assist in the information collection process of insurance providers and in submission of claims by customers.
11. Agents or banks: The insurance service provider either hires the agents/employees to sell various products or collaborates with financial institutions such as banks and third-party providers to cross-sell insurance products to their existing customer base.

The insurance firm acts as a major coordinator between the different players. The insurers help interpret laws and regulations. They are aware of drugs, side effects, treatment schedules, and procedures. They collect data on rates and costs of different services in different places. They also provide information and allied services to the customers. For example, they directly process bills and pay hospitals once the procedure has been approved. In many ways, a strong and competitive insurance industry is necessary to coordinate so many different interests.

## 6 Fraud in the Insurance Business

Insurance fraud is one of the largest sources of white-collar crime in the world, meaning that significant police effort is also devoted to tracking and eliminating it. However, given limited police resources and a universe of crime that encompasses far more than just the white-collar variety, hard insurance fraud perpetrated by organized criminals tends to be the focus of law enforcement. This leaves unorganized hard fraud and a plethora of soft fraud to remain within the purview of insurance companies.

The health insurance industry is no more immune to fraud than any other insurance subsectors. Experts estimate about 6% of global healthcare spending is lost to fraud annually.[6] In a world in which trillions of dollars are spent on healthcare by governments, nongovernmental organizations, and corporations alike, this amounts to tens of billions lost to criminal enterprises. In the USA alone, fraud is estimated to cause about US$80 billion in losses to the industry annually, with property casualty fraud accounting for US$32 billion.[7] These figures do not include fraud perpetrated on Medicare and Medicaid.

Health insurance fraud is an act of providing misleading or false information to a health insurance company in an attempt to have them pay to a policyholder, another party, or entity providing services (PAIFPA 2017). An individual subscriber can commit health insurance fraud by:

- Allowing someone else to use his or her identity and insurance information to obtain healthcare services
- Using benefits to pay for prescriptions that were not prescribed by his or her doctor

Healthcare providers can commit fraudulent acts (PAIFPA 2017) by:

- Billing for services, procedures, and/or supplies that were never rendered
- Charging for more expensive services than those actually provided
- Performing unnecessary services for the purpose of financial gain
- Misrepresenting non-covered treatments as a medical necessity
- Falsifying a patient's diagnosis to justify tests, surgeries, or other procedures
- Billing each step of a single procedure as if it were a separate procedure
- Charging a patient more than the co-pay agreed to under the insurer's terms
- Paying "kickbacks" for referral of motor vehicle accident victims for treatment
- Patients falsely claiming healthcare costs
- Individuals using false/stolen/borrowed documents to access healthcare

Tackling fraud is critical to the industry, especially with fraud becoming ever more complex. By its nature, insurance fraud is difficult to detect, as its aim is to be indistinguishable from genuine insurance claims. In each of the above cases, identifying the fraud that has been perpetrated can be a laborious process, consuming time and effort. Given that healthcare spending can be sudden, urgent, and unexpected, checking for fraud can be a complicated process. Companies must balance their financial constraints with the reality of healthcare situations.

According to an estimate of the US National Healthcare Anti-Fraud Association (NHCAA),[8] 3% of all healthcare spending is lost to healthcare fraud (LexisNexis

[6] "The Health Care Fraud Challenge," Global Health Care Anti-Fraud Network. http://www.ghcan.org/global-anti-fraud-resources/the-health-care-fraud-challenge/ (accessed on Jun 12, 2017).

[7] http://www.insurancefraud.org/statistics.htm (accessed on Jun 12, 2017).

[8] The Challenges of Healthcare Fraud. https://www.nhcaa.org/resources/health-care-anti-fraud-resources/the-challenge-of-health-care-fraud.aspx (accessed on Aug 17, 2018).

2011). Financial fraud including unlawful billing and false claim is the most common type of health insurance fraud and generally tied into aspects of organization and health information management (AHIMA Foundation 2010). The data mining tools and techniques and predictive analytics such as neural network, memory-based reasoning, and link analysis can be used to detect fraud in insurance claim data (Bagde and Chaudhary 2016).

Healthcare fraud leads to higher premium rates, increased expenses to consumers, and reduced coverage. It increases cost to employers for providing healthcare insurance to their employees affecting the cost of doing business. Besides financial losses, fraudulent activities lead to exploitations and exposure of people to unnecessary and unsafe medical procedures, which can have devastating health side effects.

Detecting healthcare insurance fraud is a long drawn-out, complicated process that costs companies time, effort, money, and the goodwill of their customers. Modern technology and statistical software have helped to reduce this cost, but it remains a significant burden on the resources of customer service departments the world over. Healthcare insurance fraud-proofing and management strategies and activities may include "improving data quality, building a data centric culture and applying advanced data analytics."[9] These provide opportunity for significant cost savings by the healthcare industry.

The innovations in insurance products and development in information communication technologies can help to design tailor-made insurance products with improved underwriting and pricing of healthcare insurance and coverage option. Technology and improved information systems can benefit stakeholders and market participants and lead to improved welfare and consumer satisfaction.

In the past, the primary manner in which insurers detected fraud was to employ claims agents who investigated suspicious claims. However, as data analytics software gains prominence and becomes more powerful, firms are becoming more able to identify patterns of abnormal behavior.

Fraud detection, however, must contend with the possibility of misidentifying fraud. Allowing false claims to go through the system hurts the company's profit and increases premiums. Forcing genuine claimants to go through the fraud detection process, however, increases costs and hurts customer satisfaction. As these constraints are diametrically opposed, any attempt to curb one will tend to increase the other.

[9]Deloitte (2012), Reducing Claim Fraud – A Cross Industry Issue. http://deloitte.wsj.com/cfo/files/2012/10/reducing_claims_fraud.pdf (accessed on Aug 17, 2018).

## 7 *Ideal*'s Business

*Ideal Insurance Inc.* is one of the largest healthcare insurance providers in the USA and other developed nations. It has expansion plans to enter into emerging markets where penetration is much lower. Most of the underwriting work is done in the US or the UK office. It has a large claim processing team located in all the countries of presence. It has back offices in other countries such as India, Singapore, and Thailand. With the increasing competition in the market, the company has had to focus on quick settlement, claims settlement ratio as well as profit margin. The company has been investing significant amount on automating the claim settlement process in order to increase customer satisfaction rate, reduce the length of the settlement cycle, and reduce the loss associated with claim leakage due to potential fraud claims. Table 25.1 shows some of the key performance measures that Sebastian was tracking. Though Ideal offers competitive premiums and maintains a high claim settlement ratio and low repudiation ratio, its net promoter score (NPS)—a metric of customer satisfaction—is significantly lower than the industry average.

The company has an automated system in place that reviews the basic information of all the claims based on prespecified rules set by experts. The rules are used to classify the claims into three categories, namely, Genuine, Discuss, and Investigate. This information is passed to the claims settlement team to act further. The claims classified as "Genuine" are processed on high priority with a settlement turnaround time (TAT) of 7 working days. The claims classified as "Discuss" are forwarded to the data collection team in order to collect more and granular information about the claims. Such claims usually take on an average of up to 30 days. The claims classified as "Investigate" are forwarded to the claim investigation team for thorough enquiry of the stakeholders. These claims usually take between 1 and 2 months or sometimes even more than 3 months for settlement or closure based on the results of the investigations or if the claims are litigated. Some customers file litigation suits in case of rejected claims, and then it is the company's onus to prove that the claim is fraudulent. Anecdotal evidence suggested that Ideal's experienced claim settlement personnel did not completely trust the

**Table 25.1** Ideal performance vis-à-vis industry

| Parameter | Ideal Insurance[a] | Industry average[a] |
|---|---|---|
| Revenue per policy | 70 | 100 |
| Contribution from core operations | 20% | 35% |
| Contribution from investments | 80% | 65% |
| Claims settlement ratio | 92% | 87% |
| Claims repudiation ratio | 2.5% | 4% |
| Average settlement period | 72 days | 30–45 days |
| Net promoter score (out of 100) | 73 | 86 |

[a]Scaled to 100 if no metric given

current system. The feeling was that it was somewhat limited due to a "bureaucratic approach." They did their own checks, often uncovering fraud from claims identified as Genuine by the current system. Rachel discussed the same with the more senior personnel to understand the root cause. The argument given was that "data maturity is vital in uncovering fraud pattern and therefore we re-analyze the claims (even if it is identified as Genuine) when more information is populated in the database." Also, their experience suggested that it is difficult to uncover professional soft fraud which is usually identified only by thorough analysis or if there is any lead from an outside stakeholder or by doing network analysis of other fraud claims.

Rachel wanted a feedback on the current processes and hired an independent consultant to examine 100,000 claims from the historical claims data. The consultant with the help of an investigation team did a thorough examination of the provided claims and classified those claims as fraud or non-fraud. Out of 100,000 claims investigated by the consultant, they identified 21,692 as potentially fraudulent claims and 78,308 as genuine claims. Comparing these results with the previous settlement records of 100,000 claims showed that more than 90% of these fraudulent claims were not identified by the existing automated system. Claims of more than 6657 customers were delayed because of the investigation process suggested by the current system. The average claim cost is around US$5000 while the cost of investigation is approximately US$500. The cost of investigation was equally divided between the internal manpower cost and the external cost of obtaining information specific to the claim. Thus, conservatively, Rachel estimated that investigating a genuine claim leads to a loss of US$500 and increases customer dissatisfaction due to the delay in settlement. It also reduces the efficiency of the investigation team. Settling a potential fraudulent claim leads to a loss of US$5000 on average and negatively affects the premium pricing and the effectiveness of the underwriting team.

The management team discussed the report with its claims settlement team and sought advice from them on how to improve the processes. Several excellent suggestions were gathered, such as monitoring each stage of the process instead of the entire process, flagging issues overlooked by the current system, and using past similar-claims data to verify the claim. The claims settlement team also suggested hiring an analytics professional to build a fraud predictive model using advanced data science techniques. They explained that this would not only help in correctly identifying potential fraud claims but also in optimizing the efforts of the claims investigation and settlement teams. They also mentioned that their closest competitor had recently set up an analytics department, which was helping in various aspects of business such as conduct of fraud analytics, predicting claims, review of blacklisted stakeholders, effective underwriting, and developing customized products.

Rachel turned to Raghavan, a recent hire who had graduated with a master's degree in business analytics from a famous school in South Central India. Raghavan had expertise in analytics specifically in the insurance domain. He was charged to hire professionals and supervise the project: to build a predictive model to identify potential fraudulent claims out of reported health insurance claims. This solution, Rachel and Sebastian felt, will help not only in reducing the losses due to fraud but also in improving efficiency and customer satisfaction and reducing the claim settlement cycle.

## *7.1 Project Details*

Raghavan's initial thoughts were to deliver a robust analytical solution that would improve fraudulent claim identification process at *Ideal*'s site without investing much time and effort in the field at the early stage. The potential fraud claims can be investigated more rigorously, while genuine claims can be settled quickly at the same time. He co-opted Caroline Gladys, who had also recently graduated in analytics from one of the premier business schools, who had been working with the business intelligence team and now wanted to switch to the advanced analytics team. Raghavan provided her the opportunity to work on this proof of concept and deliver a solution.

## *7.2 Data Description*

Caroline through her experience within *Ideal Insurance* quickly created a sample dataset at the transaction level for 100,000 health insurance claims. Each observation has up to 33 parameters related to the claim such as policy details, whether there was a third-party administrator, demographic details, and claim details. The complete details are shown in Appendix 1. Tables in Appendix 2 provide the coding of variables such as product type, policy type, and mode of payment.

The data in the tables are collected by the transaction processing system (1) when the policy is issued, (2) when a claim is recorded, and (3) while its progress is tracked. The ERP system did a fairly good job of collecting the necessary data.

**Table 25.2** Summary of identified fraud claims (by expert)

| Fraud | Number of claims | Proportion (%) |
|---|---|---|
| No | 78,308 | 78.3 |
| Yes | 21,692 | 21.7 |
| Total | 100,000 | |

**Table 25.3** Summary of current system's recommendation

| System's recommendation | Number of claims | Proportion (%) |
|---|---|---|
| Genuine (Green) | 91,986 | 92.0 |
| Discuss | 6560 | 6.6 |
| Investigate | 1454 | 1.5 |
| Total | 100,000 | |

Custom software helped put together the information into tables and created reports for further processing. *Ideal* had invested a great deal in automation of transactions in the past and was looking to reap dividends from the reporting system.

Caroline also obtained the classification of claims as Fraud/Not Fraud examined by the expert who had investigated 100,000 claims. The classification is shown in Table 25.2. Additionally, the data in Table 25.3 provide the classification of all 100,000 claims as Genuine, Discuss, and Investigate according to the current automated system.

Caroline put together all the data in a dataset (*idea_insurance.csv*; refer to the website) and also the detailed definitions of the variables available and data description required to decode the categories such as product type, policy type, and claim payment type.

Having collected all this information, Caroline was wondering how to begin the analysis. Was predictive analytics superior to the expert system used by *Ideal*? Would the experts who created the system as well as the senior settlement officers readily accept the changes? She was also worried about the ongoing creation of rules and maintenance of the system. That would cost significant investment in people and technology, not to mention training, obtaining data, etc. She would have to clearly convince the management that this was a worthwhile project to pursue!

## Electronic Supplementary Material

All the datasets, code, and other material referred in this section are available in www.allaboutanalytics.net.

- Data 25.1: ideal_insurance.csv

# Appendix 1

| Sr. no. | Variable name | Description |
|---|---|---|
| 1 | tpa | Third-party administrator ID |
| 2 | policy_ref | Policy reference number |
| 3 | member_id | Insured member ID |
| 4 | sex | Sex of the insured member |
| 5 | dob | Date of birth of the insured member |
| 6 | policy_start_dt | Date of commencement of policy |
| 7 | policy_end_dt | Date of expiry of policy |
| 8 | prod_code | Product type |
| 9 | policy_type | Policy type |
| 10 | sum_insured | Maximum sum insured available to policyholder |
| 11 | claim_ref | Claim reference number |
| 12 | claim_dt | Date of claim intimation to insurer |
| 13 | hospital_id | Unique ID given to hospital |
| 14 | hos_zipcode | Zip code of the hospital |
| 15 | admit_dt | Date of admission in hospital |
| 16 | discharge_dt | Date of discharge from hospital |
| 17 | payment_dt | Date of settlement of the claim |
| 18 | claim_amt | Amount claimed by the claimant |
| 19 | nursing_chg | Nursing charges incurred during treatment |
| 20 | surgery_chg | Surgery charges incurred during treatment |
| 21 | cons_fee | Doctor consultation charges incurred during treatment |
| 22 | test_chg | Investigation charges of medical tests prescribed by the doctor |
| 23 | pharmacy_cost | Medicines consumed during treatment |
| 24 | other_chg | Any other charges that cannot be categorized in above |
| 25 | pre_hosp_exp | Amount claimed for pre-hospitalization treatment |
| 26 | post_hosp_exp | Amount claimed for post-hospitalization treatment |
| 27 | other_chg_non_hosp | Other non-hospital charges (laundry, paid TV channels, guest F&B, etc.) |
| 28 | copayment | Co-payment or excess if applicable |
| 29 | settle_amt | Final amount paid to the insured or to the medical service provider on behalf of insured |
| 30 | payment_type | The mode of payment (refer description) |
| 31 | hosp_type | Whether hospital is networked, Y/N |
| 32 | recommendation | Classified as Green/Discuss/Investigate by the claim settlement team |
| 33 | fraud | Classified as Fraud/Non-Fraud by an expert consultant |

Note: All the amounts are in US$ and dates are in d-mmm-yyyy format. The identity data such as policy number, claim number, and hospital details are masked to maintain the data privacy.

# Appendix 2

| prod_code | Product type |
|---|---|
| A | Basic policy |
| B | Hospital cash plan |
| C | Outpatient coverage |
| D | Universal health policy |
| E | Microinsurance policy |
| F | Package policy (covering more than one type of health above) |
| G | Hybrid policy (covering other than health also) |
| O | Any other product type |

| policy_type | Policy type details |
|---|---|
| A | Individual |
| B | Individual floater |
| C | Group |
| D | Group floater |
| E | Declaration |
| F | Declaration floater |
| G | Declaration with group insurance |
| H | Declaration floater with group insurance |
| O | Any other cover type |

| payment_type | Claim payment type description |
|---|---|
| A | Cashless settlement |
| B | Reimbursement to claimant |
| C | Cash benefit |
| D | Per diem basis |
| E | Lump-sum basis |
| F | Cashless to the insured |

## References

AHIMA Foundation. (2010). *A study of health care fraud and abuse: Implications for professional managing health information*. Retrieved September 15, 2018, from https://www. ahimafoundation.org/downloads/pdfs/Fraud%20and%20Abuse%20-%20final%2011-4-10.pdf.

Bagde, P. R., & Chaudhary, M. S. (2016). Analysis of fraud detection mechanism in health insurance using statistical data mining techniques. *International Journal of Computer Science and Information Technologies, 7*(2), 925–927.

LexisNexis. (2011). *Bending the cost curve: Analytics driven enterprise fraud control*. Retrieved September 15, 2018, from http://lexisnexis.com/risk/downloads/idm/bending-the-cost-curve-analytic-driven-enterprise-fraud-control.pdf.

PA Insurance Fraud Prevention Authority (PAIFPA). (2017). *Health insurance fraud*. Retrieved September 15, 2018, from http://www.helpstopfraud.org/Types-of-Insurance-Fraud/Health.

Singhal, S., Finn, P., Schneider, T., Schaudel, F., Bruce, D., & Dash, P. (2016). *Global private payors: A trillion-euro growth Industry*. New York: McKinsey and Company Retrieved September 15, 2018, from http://healthcare.mckinsey.com/sites/default/files/Global%20private%20payors%20%28updated%29.pdf.