A Root Cause Analysis of Stock-outs in the Pharmaceutical Industry

By

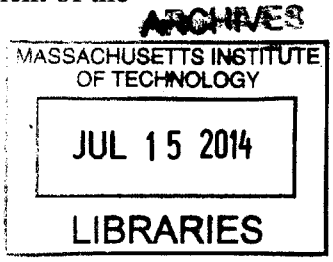
Xuewen Sun
LL.B., Dongbei University of Finance and Economics, 2011

And

Bangqi Yin
B.E. Electrical and Electronic Engineering, Nanyang Technological University, 2013

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Logistics
at the
Massachusetts Institute of Technology
June 2014



Signature redacted

Signature of Author .....
Master of Engineering in Logistics Program, Engineering Systems Division
May 7, 2014

Signature redacted

Signature of Author .....
Master of Engineering in Logistics Program, Engineering Systems Division
May 7, 2014

Signature redacted

Certified by .....
Dr. Roberto Perez-Franco
Research Director, MIT Center for Transportation and Logistics
Thesis Supervisor

Signature redacted

Accepted by .....
Prof. Yossi Sheffi
Director, MIT Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

A Root Cause Analysis of Stock-outs in the Pharmaceutical Industry

By

Xuewen Sun and Bangqi Yin

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Logistics

Abstract

PharCo (an assumed name) is a leading global healthcare company with well-recognized brands of both pharmaceutical and consumer healthcare products. As PharCo continues to expand its global presence, product stock-outs in their pharmaceutical business unit have been consistently increasing. PharCo suspected that manufacturing quality defects were a major cause of stock-outs, reducing the production yield and preventing the company from meeting customer demand. To help test this hypothesis and address the stock-out challenge, we reviewed existing research on the subject of product stock-outs within the pharmaceutical industry. To understand PharCo's manufacturing process, we conducted on-site visits and reviewed their quality control practices. Finally, we designed a mixed methods approach that combines qualitative and quantitative techniques to analyze the root causes of product stock-outs at PharCo. The analysis revealed that, instead of manufacturing quality defects, regulatory issues were the primary cause for stock-outs at PharCo. Regulatory challenges associated with developments such as new product launches, license renewals, and formulation modifications need to be addressed for PharCo to reduce their stock-out level.

Thesis Supervisor: Dr. Roberto Perez-Francos
Title: Research Director, MIT Center for Transportation and Logistics

Acknowledgments

We would like to express our sincere gratitude to our thesis advisor, Roberto Perez-Franco, for his constant encouragement and generous guidance throughout the research. The research experience has proven to be invaluable and memorable. We would also like to thank Dr. Bruce Arntzen for initiating and granting us the opportunity to participate in the project.

In addition, we would like to thank our thesis sponsoring company who has provided us the data, arranged onsite visits, and constantly helped us understand the pharmaceutical business. In particular, we want to thank Nicholas Izzi and Ben Owusu-Adjei for their support and guidance during this thesis research.

Last but not least, we thank our families for their love and support throughout our stay at MIT:

I would like to dedicate my portion of this paper to my wife, my two children, my mother-in-law, and my parents for their understanding, caring, and great support during my study at MIT. Without their support, I would never have achieved what I have today at MIT.

On behalf of Xuewen (Benny) Sun

I would like to thank my parents and my girlfriend for their invaluable support to me in the whole last year. They provided all kinds of help and enriched my experience in MIT Supply Chain Management program.

On behalf of Bangqi Yin

Table of Contents

List of Figures

List of Tables

1 Background and Introduction

The pharmaceutical industry develops drugs or pharmaceuticals licensed for use as medications. The business has a tremendous market size, with an estimated global market of 1.2 Trillion U.S. dollars in 2016 and a growth rate of average 3-6% from 2012-2016 (IMS Institute, 2012). With the rapid market growth and increased consumer needs, pharmaceutical companies are facing major challenges in keeping up with the demand and supplying sufficient products to their customers. Consequently, supply chain efficiency has become one of the key criteria for pharmaceutical companies to succeed (MCE Executive Issue No. 38, 2012, p. 12). By optimizing their supply chains, pharmaceutical companies can realize cost savings, increase their service level, and avoid loss of sales or potential negative publicity. Today, supply chain efficiency has even larger impact on businesses than in previous generations due to global outsourcing and consumption in the pharmaceutical industry. Therefore, many pharmaceutical companies have regarded supply chain development as a major portion of their business strategy.

Within supply chain strategies, stock-out management remains a very critical element. Stock-out exerts huge influence not only on revenue, customer satisfaction or public relations, but also on human health, due to the nature of the product. While most drug companies are motivated to improve the stock-out performance for their product lines, they usually lack a systematic tool to analyze their supply-demand process in order to determine the root causes of stock-outs. As a result, many drug companies are hindered from probing the problem and investing their resources to address stock-out challenges.

1.1 Problem Identification

As a leading pharmaceutical company, PharCo[1] faces the same challenge in managing their stock-out performance as many others. This has become more evident while they are expanding globally with the surge of product demand. Within five months between October 31, 2012 and April 01, 2013, more than 16,000 stock-out events were reported in PharCo. Stock-out has become a bottleneck for PharCo's further growth. In order to achieve its business objectives of delivering robust, high-quality products to customers in a timely manner, PharCo has decided to optimize their manufacturing processes and quality control practices to improve their supply chain performance, especially to reduce the stock-out level. As a prerequisite to start the changes, PharCo initiated a Root Cause Analysis with MIT graduate students to diagnose the causes of their stock-outs, mainly in their pharmaceutical business unit where most of the stock-outs have happened. In particular, PharCo suspected manufacturing quality defects as probably the main cause of stock-outs and wants to inspect the causality between manufacturing quality and stock-outs.

1.2 An Overview of PharCo's Supply Chain

PharCo's supply chain consists of five key players: Raw Material Supplier, Primary Manufactory, Secondary Manufactory, Distribution Center, and First-point Customer.

1.2.1 Raw Material Supplier

PharCo sources the raw material for all product lines from over 1,000 suppliers globally. Those raw materials are transported to their primary manufactories to produce the primary or semi-finished products. When the raw materials are received in the primary manufactory, they go

[1] PharCo is the company with whom we collaborated on this project. Due to confidentiality requirements, identifying information has been removed and the company is disguised as 'PharCo' in this paper.

through strict quality inspection processes and then are stored in climate controlled environment to prevent contaminations.

1.2.2 Primary Manufactory

The primary level manufactory includes the manufacture of active pharmaceutical ingredients and intermediates from basic chemical and biological substances (CAMH/MIN, 2007, p. 1). The active pharmaceutical ingredients and intermediates will normally make up a small portion of the final medicine; however, they are the essential parts that treat the disease. Primary manufacturing often involves chemical reactions to create new components and there can be many stages to these reactions. After the primary manufacturing, the semi-finished products (normally the ingredients) are delivered to the secondary manufactories for further processing, assembling, and packaging to the dosage forms.

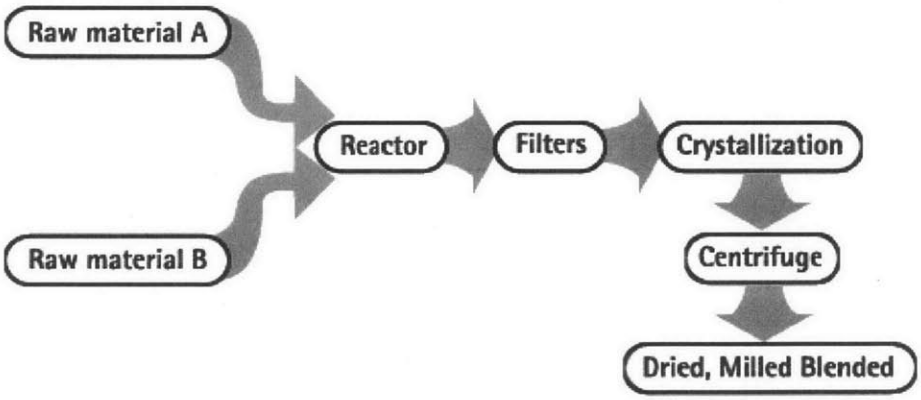


Figure 1 - Example of Ingredient Reaction (ABPI, 2014, p. 5)

1.2.3 Secondary Manufactory

Secondary manufacturing includes the production of finished dosage forms from the semi-finished products transported from the primary manufactory. After receiving the ingredients from the primary site, the secondary site checks the ingredients to make sure they meet the required

specification. Then, the active ingredient is turned into a medicine that can easily be taken by the patient, by mixing it with other substances. These are called excipients and they make up most of the volume of a medicine. Although they have no active role in curing a patient, they allow the active ingredient to be made into a medicine, such as a tablet.

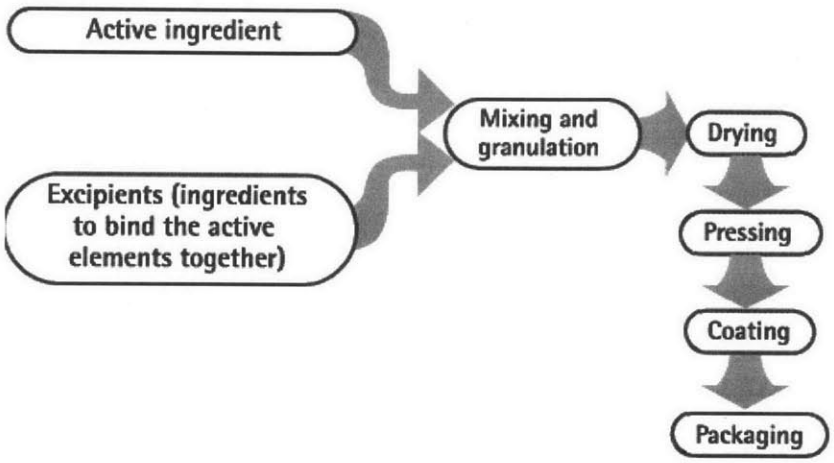


Figure 2 - Medicine Formulation in Secondary Manufactory (ABPI, 2014, p. 6)

1.2.4 Distribution Center

Distribution Centers consolidate product orders and arrange pharmacy-specific deliveries for PharCo's customers. The process in the distribution center includes inventory management, product packaging and distribution, and customer order fulfillment. This is the last stop during the process before the products reach to the customers.

1.2.5 First Point Customer

PharCo's first point customers are pharmaceutical wholesalers and distributors. Depending on the product line, customers order products through different sales systems or channels. Those orders are transferred to PharCo's distribution centers and then processed in a way that the first received and most critical orders are served in priority.

Figure 3 - Key Players of PharCo's Supply Chain

1.3 Thesis Research Area

The key research area for this thesis is to investigate various causes for pharmaceutical stock-outs in PharCo's supply chain, and further to determine whether manufacturing quality defects are the root cause for stock-outs. In PharCo, stock-outs are defined between the secondary manufactory and the distribution center (as shown in Figure 3). When the inventory controller requests a batch of products to be shipped to the distribution center from the secondary manufactory but the request is not fulfilled in a timely manner, a stock-out event occurs and is recorded in PharCo's stock-out management system. In order to identify the attributes causing stock-outs and further diagnose the root causes, our thesis analyzed all possible factors contributing to stock-outs in PharCo and evaluated the weight of each factor using a matrix we customized for this research.

2 Literature Review

The majority of available resources corresponding to our research area are theses and dissertations from MIT and other universities. Due to the scarcity of studies on the exact topic we are researching, i.e. the correlation between manufacturing quality and stock-outs in the pharmaceutical industry, we find it more relevant to evaluate current pharmaceutical research on three areas: stock-out, manufacturing quality, and root cause analysis.

2.1 Stock-out

Pharmaceutical stock-outs not only affect the financial performance of the drug companies, but more importantly, will cause the modification or discontinuation of medical treatments. Masters (2013) finds that "Stock-out of medicines has a profound effect on health in various ways". He states, "First, if a drug is not available then a sick patient who visits the health facility will not be able to receive the treatment they need; second, if a facility experiences stock-out a patient may be less willing to visit the health facility because they do not believe they will get the care and medicine they needed." (p. 1) Therefore, assessing the root causes of pharmaceutical stock-outs and further developing preventive solutions are essential in improving human health.

While Masters (2013) collects the data and analyzes the severity of pharmaceutical stock-outs in Uganda, Kenya and Ghana, and utilizes regression models to identify various affecting factors such as the rurality of the facility receiving drugs, he does not examine the manufacturing process in the factories and how the manufacturing process affects the stock-outs. Furthermore, the data collected by Masters (2013) focuses on the three countries mentioned above rather than the global regions required to be evaluated by PharCo. In addition, his definition of stock-out is based on the observation of drug unavailability by the surveyor, which is fundamentally different from the stock-out defined by PharCo. In our research, PharCo defines 'stock-out' as an event

that the finished goods inventory at their secondary manufactory fails to meet the demand of their distribution center.

2.2 Manufacturing Quality

It is very important to understand what 'quality' means before conducting any research on the subject. As Juran & De Feo (2010) point out, it is necessary to define 'quality' before one can know how to manage it (p. 5). According to Crosby (1984), 'quality' has to be defined as conformance to requirement, not as goodness (p. 64). The original definition for quality was 'fitness of use', which was defined by the consumers affected by the good or service. However, this concept falls short while more and more people use methods of managing for quality. Juran & De Feo (2010) further define 'quality' as 'fitness for purpose'. No matter what the product or service is, its quality must fit its purpose of satisfying the customer requirements (p. 5). This concept is useful for us to understand the specific definition of quality during the research.

The pharmaceutical industry is well known to be capital intensive; however, despite the heavy investment spent on drug development, and stringent oversight by regulatory bodies, quality assurance monitoring is surprisingly rudimentary when compared with other industries. There are two main reasons: (1) high profit margin made increased investment in quality control tools unattractive; (2) changes to manufacturing process require revalidation and extensive sampling which makes it uneconomical to continuously improve them (Srinivasan, 2011, p. 11). On the other hand, Shirazi (2001) states that "Quality assurance is highly significant in the success and sustainability of an organization and provides a basis for the future direction and motivation, ensuring improvement and ultimate success of an entity" (p. 1). Therefore, identifying the quality defects in pharmaceutical manufacturing process and developing corrective systems are critical to ensure the sustainability of the pharmaceutical industry.

Shirazi (2001) discusses the enhancement of leadership, training, and process management as a winning strategy to improve quality of beauty care and over-the-counter (OTC) products (p. 2). While this strategy emphasizes the initiatives to improve quality assurances, it does not assess the manufacturing process itself, which might be the source of poor quality. In addition, PharCo requires inspecting the stock-outs for prescribed medicines, which are different from the OTC products investigated by Shirazi (2001). Consequently, our methodology will take into consideration factors that Shirazi (2001) does not address to meet the specific requirement of our research.

Srinivasan (2011) evaluated the current software system of statistical process control in a pharmaceutical company and proposed a new Decision Support System to be integrated into the current process. This new system aims to automate the generation of inspection documents and improve process quality / productivity. The evaluation matrix presented in Srinivasan (2011) is useful in examining existing statistical quality control models used in pharmaceutical companies; however, the model examined and the data collected are based on the division of the pharmaceutical company located in Basel, Switzerland (p. 13). While there is common ground among pharmaceutical plants of different companies in different locations, the specific models and data collected from those plants will vary; hence, not all procedures are transferrable to other situations.

2.3 Root Cause Analysis

Root cause analysis (RCA) is a structured investigation that aims to identify the true cause of a problem and the actions necessary to eliminate it. It is a collective term used to describe a wide range of approaches, tools, and techniques used to uncover causes of problems (Andersen & Fagerhaug, 2006, p. 12-13).

Finding the Root Cause or Causes of an organizational problem is the single most important determinant of success or failure of any problem-solving method (Andersen & Fagerhaug, 2006, p. 388). This holds true when we try to diagnose and remedy pharmaceutical stock-outs. As Balasubramanian (2009) states, "In order to improve the reliability of a sub-system, thereby a system, it is necessary to identify the root causes of these problems precisely so that the corrective action taken can be effective."

Shirazi (2001) adopts the case study method to develop a strategy to improve the quality control for pharmaceutical contract manufactories. For example, in a case study regarding understaff challenge, Shirazi (2001) designed a training program to train new employees to supply the labor. This method fails to diagnose the root cause of the problem in the context and only addresses one possible solution. It does not identify and address other alternatives such as ineffective shift planning that may have caused the understaffing. Therefore, the case study method is not adequate to determine the root causes as it fails to identify alternative causes and further define an array of solutions.

One specific type of RCA method, Failure Mode Effect and Criticality Analysis (FMECA), is popular for inductive system analysis (an analysis that derives general principles from specific observations). FMECA is defined as "an engineering technique used to define, identify, and eliminate known and/or potential failures, problems, errors, and so on from the system, design, process, and/or service before they reach the customer" (as cited in Balasubramanian 2009, p. 66). This method quantifies the criticality of risk factors, allowing the remedial effort to be prioritized. Wang (2010) applies the FMECA model to classify the root causes of discarded inventories for pharmaceutical company SPM (p. 24). Balasubramanian (2009), on the other hand, combines a set of RCA models to minimize the errors of an automated prescription-filling

system. The model includes three RCA methods: first, it constructs a fault tree to identify the components affected and the errors; second, it uses FMECA to determine the initial criticality of the factors, which was then confirmed by means of an experiment; third, a dynamic Diagnostic Decision Tree (DDT) was constructed to generate preventive/corrective maintenance plans. This combined model reveals the possibility for us to customize RCA methods to meet the specific requirements in our research.

2.4 Summary of Literature Review

There are ample studies relevant to our thesis research, i.e. the stock-out, manufacturing quality, or root cause analysis in pharmaceutical industry. Those researches exemplify the severity of drug stock-outs, explain the necessity and principle of defining 'quality', provide examples of evaluating quality assurances, and supply plausible RCA methods. None of these, however, is tailored for the problem identified in this thesis research. In order to assess the correlation between manufacturing quality and pharmaceutical stock-outs on a global level, we need to further develop customized RCA methods and models to meet the needs.

3 Methodology

This chapter provides details on the methodologies we applied during this thesis research, both qualitatively and quantitatively. The chapter starts with an introduction of a specifically chosen Root Cause Analysis framework, "Do It Root Cause Analysis", which we used to guide our research process. Then, it provides details on both the qualitative and quantitative data analysis methods under the "Do It Root Cause Analysis" framework. Qualitatively, it introduces the "Manual Coding" method which we used to categorize each stock-out incident. Quantitatively, it describes the statistic regression model and quantitative Root Cause Assessment matrix we used to quantify the influence of each cause on stock-out incidents in PharCo.

3.1 Root Cause Analysis Framework

In order to select the most applicable Root Cause Analysis (RCA) framework to guide our research and analyze the causes of stock-out events, we need a standard against which different RCAs can be evaluated. According to Gano (2007), an effective root cause analysis process should provide a clear understanding of how the proposed solutions meet the goal. Gano (p. 1) suggests a list of six criteria to evaluate the RCAs.

1) Chosen RCA could clearly define the problem and its significance to the problem owners.

2) Chosen RCA could clearly delineate the known causal relationships that combined cause the problem.

3) Chosen RCA could clearly establish causal relationships between the root causes and the defined problem.

4) Chosen RCA could clearly present the evidence used to support the existence of identified causes.

5) Chosen RCA could clearly explain how the solutions will prevent recurrence of the defined causes.

6) Chosen RCA could clearly document criteria 1 through 5 in a final RCA report so others can easily follow the logic of the analysis.

Of all the popular RCA methods we researched, including "Events and Causal Factors Analysis", "Change Analysis", and "Barrier Analysis", there is no single one that would meet all six requirements so as to allow us to analyze stock-out incidents. Therefore, we further researched customized models which would combine the strengths of current RCA methods to meet all six criteria as indicated above. Among those customized models, DO IT Root Cause Analysis model meets all six criteria. Designed by Okes (2009), "the model consists of two major phases: steps 1-5 are the diagnostic phase, and steps 6-10 are the solution phase. And while the model looks linear, a unique feature is the iterative nature of the five diagnostic steps" (p. 7). Below exemplifies how the model works (figure 4 and 5).

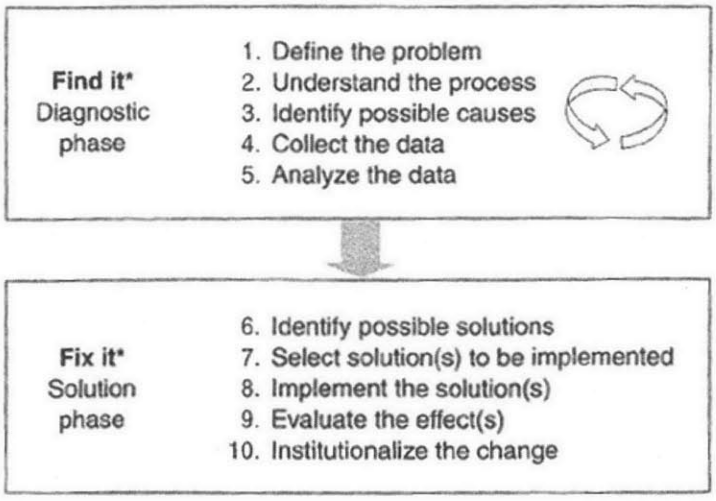


Figure 4 - The DO IT Problem-solving Model (Okes, 2009, p. 8)

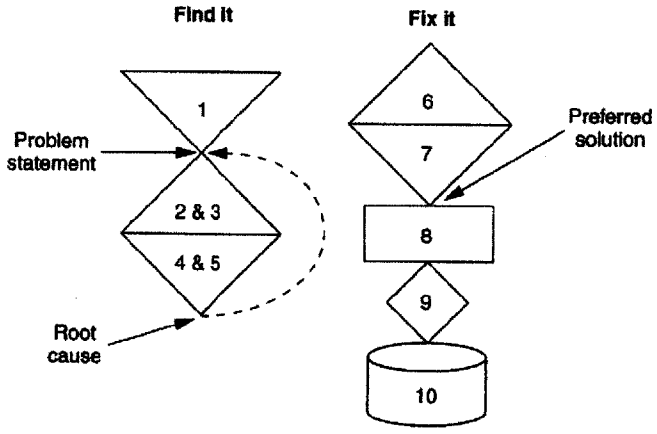


Figure 5 - Visual Depiction of The Model (Okes, 2009, p. 9)

This model provides a flow that readers can follow easily. Steps 1 through 7 break down the problem with creative thinking on both the convergent level (step 1, 4, 5, and 7) and the divergent level (step 2, 3, and 6). Steps 8 through 10 provide the project management phase to recommend the solution, implementing the solution, and maintaining the improvement and related knowledge.

The DO IT Root Cause Analysis model provides us with a logic framework to conduct our research and solve the identified problem. Section 3.2 and 3.3 below further introduces complementary tools within this framework.

3.2 Qualitative Data Analysis Method

The database provided by PharCo for this thesis project contains stock-out incidents and the comments from both customers and suppliers regarding the cause of each stock-out incident. Those comments were logged by end users into the stock-out management system and were categorized with default categories set up in the system. Those default categories are often vague and hard to understand. For example, a category named 'Trading Partner Issue' actually means 'Invoicing Issues with Trading Partners'.

To better describe the causes behind each stock-out incident and facilitate the statistical analysis, we applied a qualitative data analysis method, "Manual Coding", to manually code those comments regarding stock-out causes. Then, we translated those codes into new defined, meaningful categories that are more accurate and easier to understand. In this context, the term "code" is, according to Saldana (2013), "most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data" (p. 3). The process of transferring above-mentioned comments to categories and themes/concepts is described as below (figure 6).

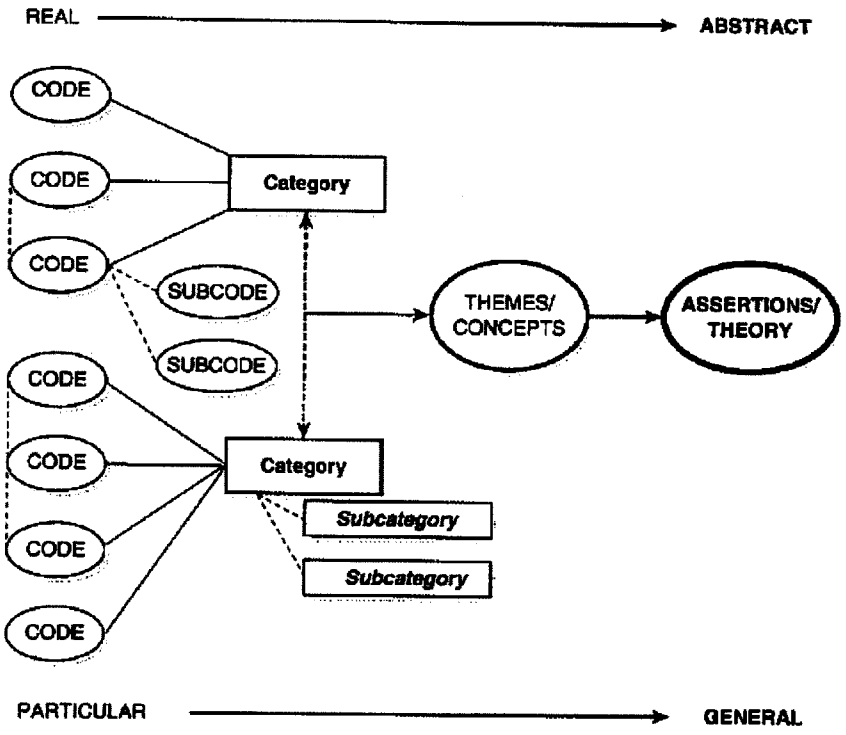


Figure 6 - A streamlined Codes-to-theory Model for Qualitative Inquiry (Saldana, 2013, p. 13)

Though diverse coding methods could be used, the most appropriate ones for this thesis are "initial coding" and "pattern coding". "Initial coding" breaks down qualitative data into discrete parts, closely examines them, and compares them for similarities and differences (Saldana, 2013,

p. 100). "Pattern codes" are explanatory or inferential codes that identify an emergent theme or explanation, and serve to pull together a lot of material into more meaningful categories that we need during this research (p. 210). Figures 7 and 8, shown below, provide examples of applying "initial coding" and "pattern coding" during our thesis research.

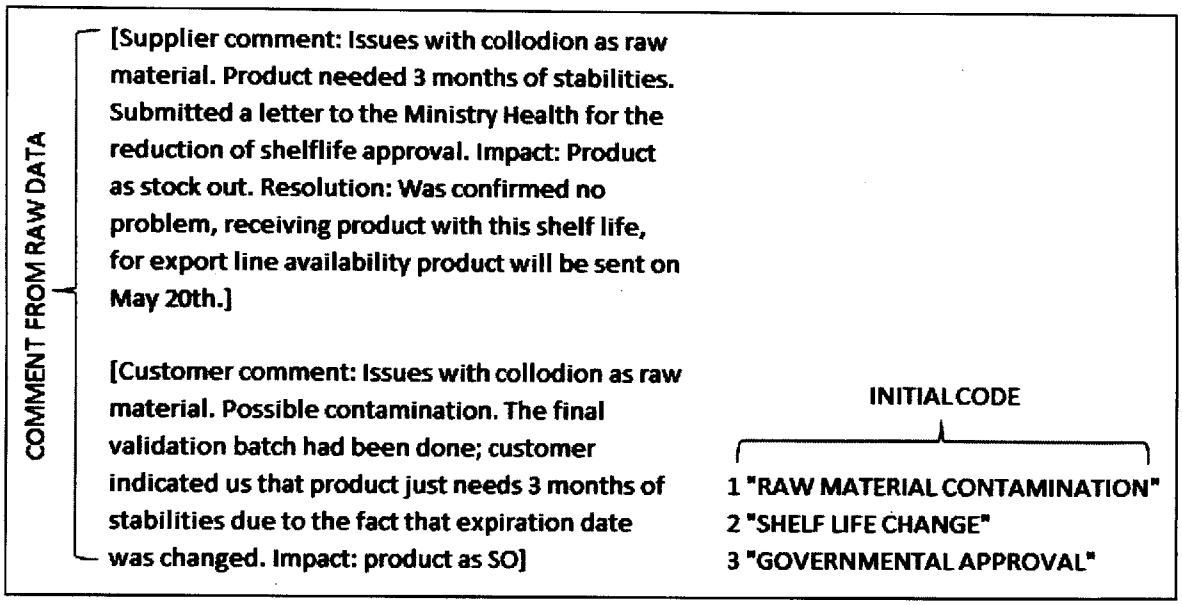


Figure 7 - Example of Initial Coding

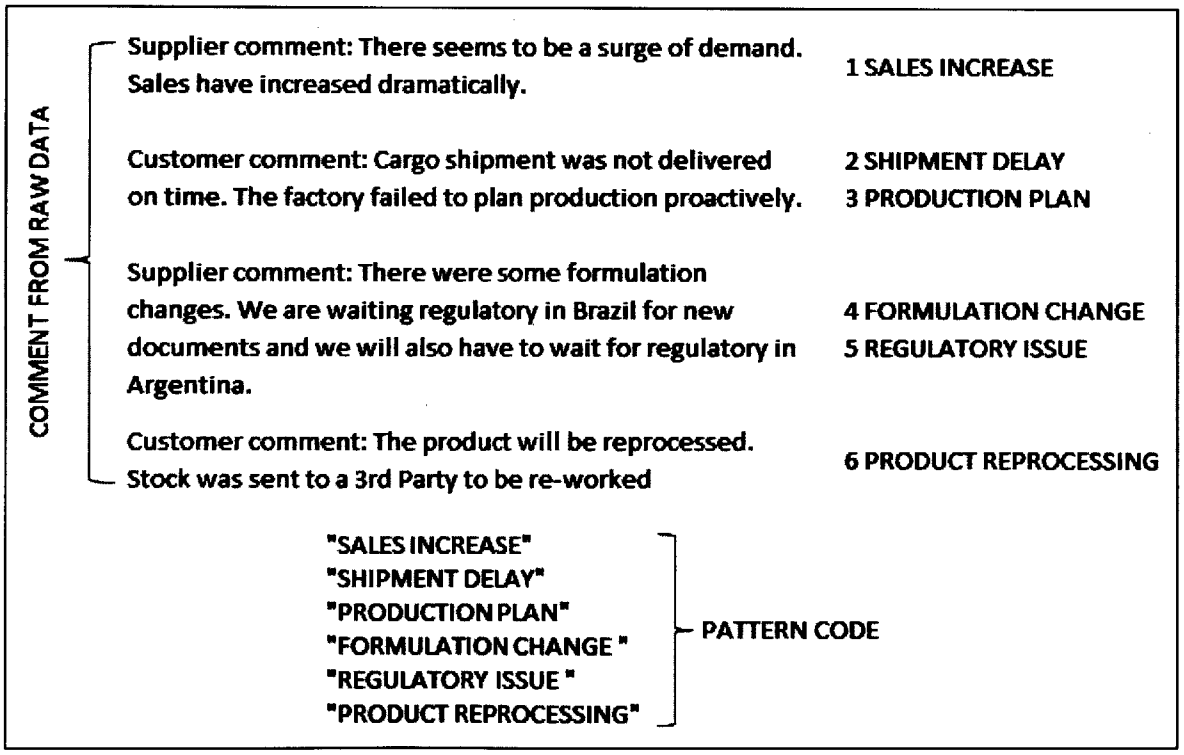


Figure 8 - Example of Pattern Coding

3.3 Quantitative Data Analysis Method

In step 5 of DO IT RCA Model, we are required to analyze the data pertaining to the research problem. This is also the essential part of the Root Cause Analysis, where we seek to find a statistical methodology to quantify and evaluate different variables contributing to stock-outs.

For any given incident, the influence of their causes can be measured by two criteria: the "predictability" and "frequency". For causes of stock-outs in PharCo, their "frequency" can be quantified by the sum of each stock-out cause reported in PharCo's stock-out management system. In order to quantify the other criterion "predictability", we developed logistic regression models to determine how each cause predicts stock-out incidents. Then, we designed a two-dimensional Root Cause Assessment matrix to evaluate the influence of each cause on stock-outs based on their individual "predictability" and "frequency".

3.3.1 Logistic Regression and Odds Ratio

One important indicator of the influence for a given variable to the incident is called "predictability". Predictability is the degree to which an outcome can be forecasted by the given variable. In order to determine the predictability of different causes, we need to build a model that quantitatively measures how a dependent variable (stock-out events) is predicted by a set of independent variables (different probable causes). We first examined the characteristic of the dataset provided by the PharCo. There are two variables within the dataset: stock-out causes (independent variable) and stock-out results (dependent variable). Both variables are binary. A binary variable, as explained by Powers & Xie (2008), is a variable that assumes one of two possible values, which are commonly described as true (1) and false (0). For the dataset we are provided, the independent variable, i.e. the cause of stock-outs, is either true or false and hence considered binary. The dependent variable, i.e. the stock-out event, also has only two possible

outcomes, i.e. a genuine stock-out (1) or a false stock-out (0), so it is a binary data as well. A false stock-out is an event initially reported as a stock-out, later proven not to be one.

The statistical approach to modeling binary data is based on the idea that there is a one-to-one correspondence between the sample data and the population quantities being modeled. Based on this idea, the variables can be modeled by linear probability models. However, as linear probability models (i.e. classical regression models) do not guarantee the conditional probabilities with the range from zero to one, we use Logistic Regression, which will avoid the shortcoming of linear probability models and constrain the conditional probabilities within the (0, 1) range (Powers & Xie, 2008, p. 32). In addition, Logistic Regression makes fewer assumptions than linear regression (e.g. homogeneity of variance and normality of errors are not assumed). It is also easily accessed because of the wide availability of statistical software packages nowadays (Meyers, Gamst, & Guarino, 2013).

A very important concept in Logistic Regression that we used to determine the predictability of each variable is called 'Odds Ratio'. The odds ratio is a ratio of odds for each cause to predict the incidents and a way of measuring the association between causes and incidents. The table below illustrates the concept 'Odds Ratio' with a simple example of the success rate for male/females within a given program.

Table 1 - Fictional Data Illustrating Gender Differences in Program Success (Meyers et al, 2013, p. 531)

	Program Successful	Program Not Successful	Total Count
Females (0)	200	100	300
Males (1)	50	150	200
Total	250	250	500

1. Odds is the probability of belonging to one group divided by probability of not belonging to that group

2. Odds Ratio (OR) is a measure of association between an exposure and an outcome
3. Odds of a male being successful: 50/150 = 0.33
4. Odds of a female being successful: 200/100 = 2
5. Odds Ratio (Whether there is a gender difference in program success): 2/0.33 = 6, which means that in this given example, women were six times more likely than men to be successful in the program

Another important concept is Statistical Significance. As Weinbach & Grinnell (2003, p. 94) indicated, Statistical Significance is the demonstration, through the use of mathematics and the laws of probability, that the relationship between variables in a sample is unlikely to have been produced by sampling error. In the logistic regression, statistical significance is evaluated by what the JMP software calls the Prob>ChiSq value. As per the internal standard of confidence level in PharCo, the rejection level is set up as 0.05, which means that the Prob>ChiSq value has to be less than 0.05 for a relationship in the logistic regression to be declared statistically significant.[2]

3.3.2 Quantitative Root Cause Assessment Matrix

As we discussed before, the influence of any cause on incidents can be measured by two criteria: the predictability and the frequency. The predictability, as we described in section 3.3.1, is the degree to which an outcome can be forecasted by the given variable. It is also sometimes called probability. For example, among all car accidents happening in US in 2009, around 0.011% of them caused fatality (US Census Bureau, 2012); therefore, the probability (predictability) for a car accident causing fatality in US was 0.011% in 2009. The frequency, on the other hand, explains the likelihood of the occurrence of the cause. For example, in 2009, there were

[2] This rejection level is adopted throughout this research.

5,505,000 car accidents happening in US; hence the frequency of car accident in US was 5,505,000 in 2009 (NHTSA, 2010, p. 2).

In order to evaluate the influence of each cause on stock-outs and further define which one is the root cause, we developed a two-dimensional Root Cause Assessment matrix based on the predictability and frequency. In our research, the predictability is indicated by the odds ratio of each cause as we discussed in 3.3.1, and the frequency is indicated by the total number of each cause reported in PharCo's stock-out management system. After we calculated the "predictability" and "frequency" of each stock-out cause, we plotted the predictability (as the x-coordinate) and frequency (as the y-coordinate) of each cause into the Root Cause Assessment matrix to map and compare their influences. As per the internal standard in PharCo, we define root cause categories as those with both the predictability (odds ratio) and frequency (sum of each cause reported) above the medians (Figure 9).

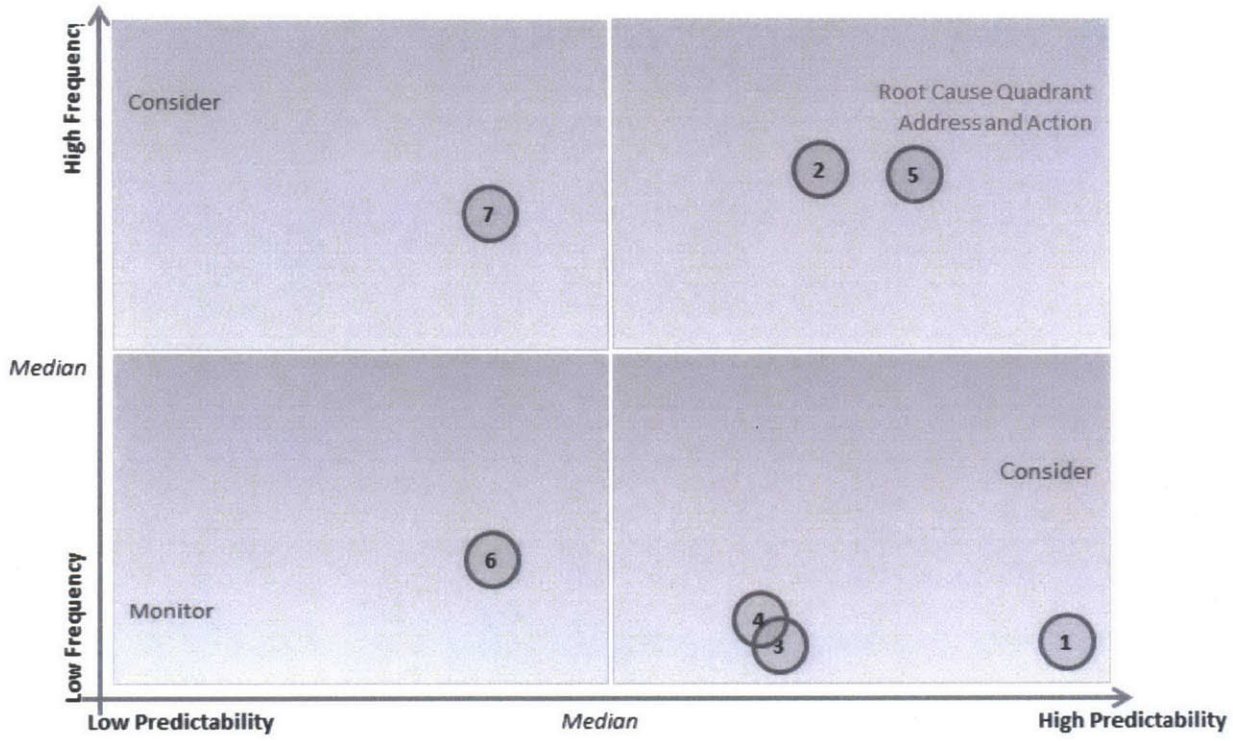


Figure 9 - Sample of Quantitative Root Cause Assessment Matrix

3.4 Summary of Methodology

The key to diagnosing the root cause for stock-outs in PharCo lies in finding and applying the most powerful methods to analyze the dataset. The DO IT Root Cause Analysis model provides us a logical framework to assess the problem, analyze the data, and identify possible solutions. The manual coding methodology offers us a qualitative tool to define meaningful root cause categories from the stock-out comments of customers and suppliers. The logistic regression and quantitative Root Cause Assessment matrix serve as quantitative methods to quantify the root cause influences. Through these methodologies, we were able to dive into the data and come up with useful results to help solve the identified problem. In the following chapters, further details of methodology application will be discussed.

4 Data Analysis and Results

This chapter explains the original data retrieved from PharCo's stock-out management system and the logistic regression model, which we named LR1, to analyze the predictability of PharCo's default stock-out root cause categories. It also introduces new stock-out root cause categories developed by us using the "Manual Coding" method and the logistic regression model, which we named LR2, to analyze the predictability of new defined categories. In addition, the Quantitative Root Cause Assessment matrix is applied to determine the influence of each stock-out cause and eventually the critical root causes. In the last section of this chapter, the meaning, the significance, and the limits of the data are discussed.

4.1 Original Data

PharCo uses a stock-out management system named 'PharSO'[3]. In PharSO, the stock-out is defined between the secondary manufactory (the supplier) and the distribution center (the customer). When an inventory controller from the distribution center orders a batch of products from the secondary manufactory and the secondary manufactory fails to fulfill the order in a timely manner, a stock-out occurs and is reported to PharSO. In this case, the distribution center (DC) will communicate with relevant parties involved such as the secondary manufactory and the transportation carrier to find out the reason and log the event into PharSO system. Meanwhile, the supplier (PharCo's secondary manufactory) will also investigate the causes and add their comments into the same event in PharSO. Sometimes, a false stock-out is reported for reasons such as product package changes or manual order discrepancies. In this case, the customer (inventory controller in the DC) will log back into PharSO and close the stock-out event with a comment explaining why the event has been excluded.

[3] Due to confidentiality requirements, identifying information has been removed and the system is disguised as 'PharSO' in this paper.

In order to determine the root causes of the stock-outs, a large number of incidents are needed for us to find out the patterns and trends. To facilitate the research, PharCo provided three historical reports extracted from PharSO consisting of 16,383 entries of stock-out events in the format of Excel spread sheet. Those reports include all the stock-out events in PharCo from October 31, 2012 to April 01, 2013 in their global territory, including Europe, North America, Japan, and other emerging markets. For each stock-out event, the dataset indicates the region of suppliers, supplier organization and location, customer organization and location, stock-out reason, event start date, event end date, type of medicine, product group, genuine stock-out indicator, and supplier/customer comments regarding the stock-out (Table 2).

Table 2 - Sample Data of PharSO's Stock-out Report

Supplier Region	Supplier Organisation	Supplier Location	Customer Organisation	Customer Location	Stock-out Reason
External Supply	XXX Company	France, XXX	Indonesia (PH)	Indonesia, XXX	Demand - Demand / Forecast Char
Internal Manufactory	XXX Company	France, XXX	Romania (Pharma)	ROMANIA, XXX	Supply - Pre Launch Production
Internal Manufactory	XXX Company	France, XXX	Colombia (PH)	Colombia, XXX	Supply - Transport / Distribution
Internal Manufactory	XXX Company	Belgium, XXX	Malaysia (PH)	Malaysia, XXX	Demand - Warehouse Issues
Internal Manufactory	XXX Company	France, XXX	Peru (PH)	Malaysia, XXX	Demand - Unspecified
Internal Manufactory	XXX Company	Belgium, XXX	Thailand (PH)	Thailand, XXX	Demand - Demand / Forecast Char
Internal Manufactory	XXX Company	India, XXX	Chile (PH)	Chile, XXX	Not Defined
External Supply	XXX Company	Switzerland, XXX	Cambodia (PH)	Cambodia, XXX	Demand - Unspecified
External Supply	XXX Company	USA, XXX	USA (PH)	USA, XXX	Not Defined
External Supply	XXX Company	Netherlands, XXX	Croatia (PH)	Croatia, XXX	Demand - Data Error
External Supply	XXX Company	Canada, XXX	Australia (PH)	Australia, XXX	Demand - Regulatory
External Supply	XXX Company	Canada, XXX	Australia (PH)	Australia, XXX	Demand - Regulatory
External Supply	XXX Company	France, XXX	Kazakhstan (PH)	Kazakhstan, XXX	Demand - Data Error
External Supply	XXX Company	France, XXX	Kazakhstan (PH)	Kazakhstan, XXX	Demand - Data Error
External Supply	XXX Company	Belgium, XXX	Venezuela (PH)	Venezuela, XXX	Demand - Demand / Forecast Char
Internal Manufactory	XXX Company	Canada, XXX	Canada (PH)	Canada, XXX	Supply - Data Error / Interface Err
Internal Manufactory	XXX Company	Germany, XXX	Colombia (PH)	Colombia, XXX	Supply - GSN Product Transfers or
Internal Manufactory	XXX Company	Canada, XXX	Colombia (PH)	Colombia, XXX	Demand - Expired Stock
Internal Manufactory	XXX Company	Spain, XXX	Spain (PH)	Spain, XXX	Supply - Raw Material / Componen
Internal Manufactory	XXX Company	Japan, XXX	Bosnia (PH)	Bosnia, XXX	Demand - Interface/Data Error

4.2 Data Cleaning and Evaluation

As our research aimed to diagnose the stock-out causes for PharCo's pharmaceutical products and the raw data contains both pharmaceutical and consumer healthcare stock-out events, we excluded the consumer healthcare record from the dataset. To analyze the data from a global perspective, we also combined all three reports from different regions into one Excel spread

sheet. We further deleted all the invalid data, including blank entries and errors, finally generating 11,501 entries of records with 35 stock-out cause categories defined by PharCo. Among the 35 stock-out cause categories, there are some related to the suppliers, such as "Supply – QA Release" and "Supply – Regulatory". Others are related to the customer side, such as "Demand – Import License Issues" and "Demand – Transportation/Distribution". The full list of all default stock-out categories can be found in Exhibit 1 of the Appendix.

In order to run the logistic regression, we also converted the stock-out categories (independent variable) and stock-out results (dependent variable) into binaries, i.e. true (1) and false (0) as shown in Table 3.

Table 3 - Binary Data after Conversion

Event ID	Demand - 3rd Party / Trading Partner	Demand - Allocation	Demand - Customs Clearance Issues	Demand - Data Error	Demand - Demand / Forecast Changes	Demand - Expired Stock	Demand - Import License Issues	Demand - Interface Error	Demand - Interface/ Data Error	Demand - Pack Change	Stock Out
1	0	0	0	0	0	0	0	0	1	0	0
2	0	0	0	1	0	0	0	0	0	0	0
3	0	0	0	0	1	0	0	0	0	0	1
4	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	1	0	0
6	0	0	0	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	1
8	0	0	0	0	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0	1
11	0	0	1	0	0	0	0	0	0	0	1
12	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	1	0	0	0	0	0	0	0
14	0	0	0	1	0	0	0	0	0	0	0
15	0	0	0	1	0	0	0	0	0	0	0
16	0	0	0	1	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	1	0	0	0

4.3 Analysis on PharCo's Default Categories

Based on PharCo's default categories, we first ran a logistic regression[4] that we named LR1. The result of LR1 is shown as Figure 10. We excluded category "Not Applicable", "Not Defined", "Demand Unspecified", and "Supply Unspecified" from the regression model because those three categories provide no insight into the nature of stock-out incidents. In addition, there are no comments from either suppliers or customers for those categories.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	5142.9825	31	10285.96	<.0001*
Full	1961.1660			
Reduced	7104.1484			

RSquare (U)	0.7239
AICc	3986.54
BIC	4217.89
Observations (or Sum Wgts)	10262

Measure	Training	Definition
Entropy RSquare	0.7239	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.8445	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.1911	$\sum$ -Log(ρ[j])/n
RMSE	0.2303	$\sqrt{\sum(y[j]-\rho[j])^2/n}$
Mean Abs Dev	0.1060	$\sum$ |y[j]-ρ[j]|/n
Misclassification Rate	0.0656	$\sum$ (ρ[j]≠ρMax)/n
N	10262	n

Figure 10 - Whole Model Test of LR1

Under the Whole Model Test (Figure 10), Prob>ChiSq[5] is below 0.0001, which indicates that, taken as a whole, the model represents a statistically significant improvement in terms of predictive power over having no model at all.

[4] Software used: SAS JMP 11 Pro for Mac.

[5] Statistical significance (p-value) is measured by the Prob>ChiSq value in JMP software. PharCo's rejection level is set as Prob>ChiSq = 0.05

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Demand - 3rd Party / Trading Partner	1	1	1.41711e-5	0.9970
Demand - Allocation	1	1	0	1.0000
Demand - Customs Clearance Issues	1	1	0	1.0000
Demand - Data Error	1	1	0	1.0000
Demand - Demand / Forecast Changes	1	1	0	1.0000
Demand - Expired Stock	1	1	0	1.0000
Demand - Import License Issues	1	1	0	1.0000
Demand - Interface Error	1	1	0	1.0000
Demand - Interface/Data Error	1	1	0	1.0000
Demand - Pack Change	1	1	0	1.0000
Demand - Product Changes / NPD Issues	1	1	0	1.0000
Demand - QA Release	1	1	0	1.0000
Demand - Quality Assurance / Documentation	1	1	0	1.0000
Demand - Regulatory	1	1	0	1.0000
Demand - Transport/Distribution	1	1	0	1.0000
Demand - Warehouse Issues	1	1	0	1.0000
Supply - 3rd Party / Trading Partner	1	1	3.01055e-5	0.9956
Supply - 3rd Party Logistic Provider	1	1	0	1.0000
Supply - Allocation	1	1	0.00043643	0.9833
Supply - Data Error	1	1	0	1.0000
Supply - Data Error / Interface Error	1	1	0	1.0000
Supply - GSN Product Transfers only	1	1	0	1.0000
Supply - Interface Error	1	1	0	1.0000
Supply - Manufacturing / Production / Capacity Issues	1	1	0	1.0000
Supply - Pack Change	1	1	0.00045237	0.9830
Supply - QA Release	1	1	0.00323295	0.9547
Supply - Quality Assurance / Documentation	1	1	0	1.0000
Supply - Raw Material / Component Supply Issues from 3rd Parties (Non GSK)	1	1	0	1.0000
Supply - Raw Material / Component Supply Issues from GSK Site	1	1	0	1.0000
Supply - Regulatory	1	1	0	1.0000
Supply - Transport / Distribution	1	0	0	.

Figure 11 - Effect Likelihood Ratio Tests of LR1

However, in the Effect Likelihood Ratio Tests for LR1 (Figure 11), all the stock-out causes have Prob>ChiSq around 1.00. This means that none of the default stock-out causes contributes significantly to the model fit although the model as a whole is significant. Therefore, it is worthwhile to reduce the number of stock-out categories into a handful of significant clusters. Using the qualitative data coding method, we manually coded the stock-out comments into 12 new defined stock-out cluster categories.

4.4 Analysis on New Defined Categories

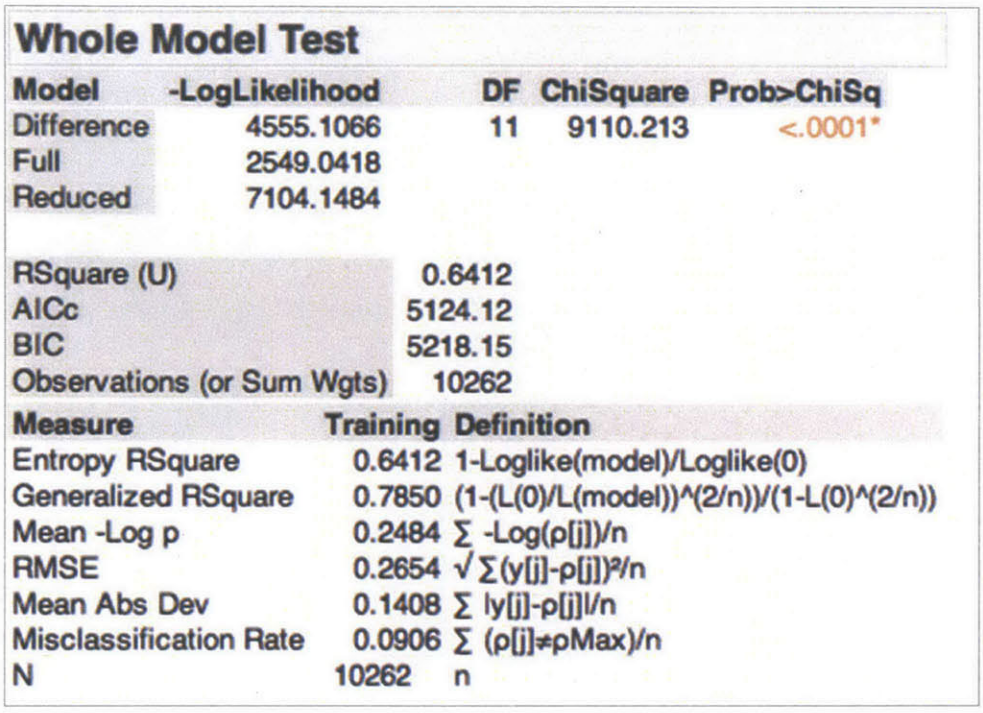
New defined stock-out categories were created through aggregating PharCo's default stock-out cause categories into the 12 new defined cluster categories through the "Manual Coding" method

(as discussed in chapter 3.2). New stock-out categories and their relationship with original ones are shown below (Table 4).

Table 4 - New Root Cause Category vs. Original Root Cause Category

New Root Cause Category	Original Root Cause Category
Invoicing Issues	Demand - 3rd Party / Trading Partner; Supply - 3rd Party / Trading Partner
Allocation Issues	Demand - Allocation; Supply - Allocation
Customs Clearance Issues	Demand - Customs Clearance Issues
Interface / Data Error	Demand - Data Error; Demand - Interface Error; Demand - Interface/Data Error; Supply - Data Error; Supply - Data Error / Interface Error; Supply - Interface Error
Forecast Error	Demand - Demand / Forecast Changes
Warehouse Quality Issues	Demand - Expired Stock; Demand - Warehouse Issues
Product Changes	Demand - Pack Change; Demand - Product Changes / NPD Issues; Supply - Pack Change
Manufactoring Quality	Demand - QA Release; Demand - Quality Assurance / Documentation; Supply - QA Release; Supply - Quality Assurance / Documentation
Production / Capacity Issues	Supply - Manufacturing / Production / Capacity Issues
Regulatory Issues	Demand - Regulatory; Supply - Regulatory
Transportation / Distribution	Demand - Transport/Distribution; Supply - 3rd Party Logistic Provider; Supply - Transport / Distribution; Demand - Import License Issues
Raw Material Issues	Supply - Raw Material / Component Supply Issues from PharCo Site, Supply Raw Material / Component Supply Issues from 3rd Parties (Non PharCo), Supply - GSN Product Transfers only

Using the 12 new defined categories, we created the logistic regression model which we named LR2 as shown below (Figure 12).



Whole Model Test

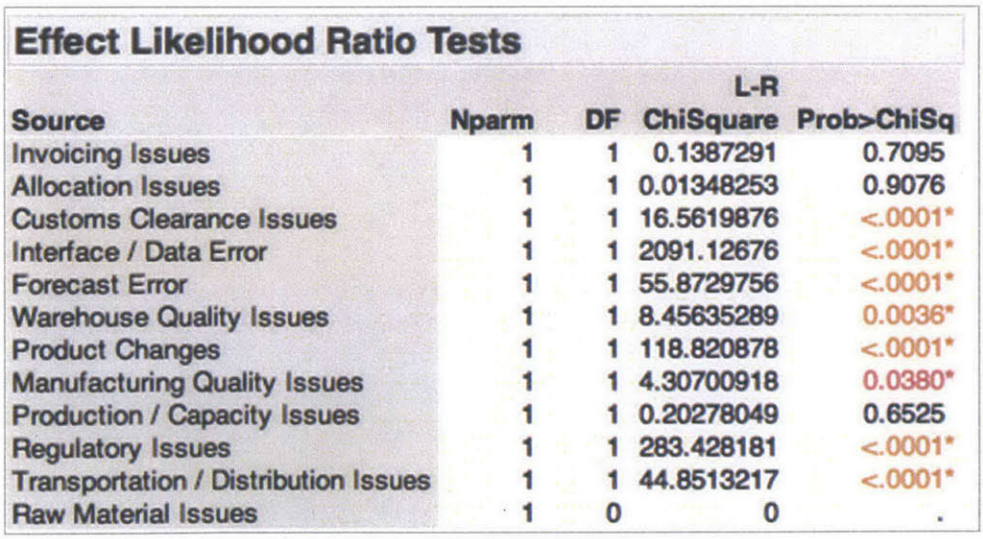
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4555.1066	11	9110.213	<.0001*
Full	2549.0418			
Reduced	7104.1484			

RSquare (U)	0.6412
AICc	5124.12
BIC	5218.15
Observations (or Sum Wgts)	10262

Measure	Training	Definition
Entropy RSquare	0.6412	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.7850	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.2484	$\sum$ -Log(ρ[j])/n
RMSE	0.2654	$\sqrt{\sum(y[j]-\rho[j])^2/n}$
Mean Abs Dev	0.1408	$\sum$ |y[j]-ρ[j]|/n
Misclassification Rate	0.0906	$\sum$ (ρ[j]≠ρMax)/n
N	10262	n

Figure 12 - Whole Model Test of LR2

Under the Whole Model Test (Figure 12), Prob>ChiSq is below 0.0001, indicating that the model as a whole is significant.



Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Invoicing Issues	1	1	0.1387291	0.7095
Allocation Issues	1	1	0.01348253	0.9076
Customs Clearance Issues	1	1	16.5619876	<.0001*
Interface / Data Error	1	1	2091.12676	<.0001*
Forecast Error	1	1	55.8729756	<.0001*
Warehouse Quality Issues	1	1	8.45635289	0.0036*
Product Changes	1	1	118.820878	<.0001*
Manufacturing Quality Issues	1	1	4.30700918	0.0380*
Production / Capacity Issues	1	1	0.20278049	0.6525
Regulatory Issues	1	1	283.428181	<.0001*
Transportation / Distribution Issues	1	1	44.8513217	<.0001*
Raw Material Issues	1	0	0	.

Figure 13 - First Effect Likelihood Ratio Tests of LR2

As is shown in the Effect Likelihood Ratio Tests (Figure 13), there is an obvious improvement compared to the last model LR1. Among 12 different categories, 9 of them have Prob>ChiSq under 0.05 (the rejection level set by PharCo), indicating that they contribute significantly to the model fit and should be included in the model. On the other hand, "Allocation Issues", with the highest Prob>ChiSq (0.9076), should be excluded to further improve the model LR2.

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Invoicing Issues	1	1	0.1529398	0.6957
Customs Clearance Issues	1	1	7.30333891	0.0069*
Interface / Data Error	1	1	433.791606	<.0001*
Forecast Error	1	1	10.7775769	0.0010*
Warehouse Quality Issues	1	1	3.50607097	0.0611
Product Changes	1	1	38.9715546	<.0001*
Manufacturing Quality Issues	1	1	0.90349294	0.3418
Production / Capacity Issues	1	1	0.11683641	0.7325
Regulatory Issues	1	1	58.4873899	<.0001*
Transportation / Distribution Issues	1	1	11.8435919	0.0006*
Raw Material Issues	1	1	0.01348253	0.9076

Figure 14 - Second Effect Likelihood Ratio Tests of LR2

After the exclusion of "Allocation Issues", as shown above (Figure 14) in the second simulation of LR2, "Raw Material Issues" now appears to be the category with greatest Prob>ChiSq (0.9076). Hence, it should be eliminated in the third simulation. To gain the best fit of the model, we iteratively excluded categories with Prob>ChiSq value greater than 0.05 (the rejection level set by PharCo) until all remaining variables have Prob>ChiSq value less than 0.05. We generated the final version of LR2 with the Effect Likelihood Ratio Tests as shown below (Figure 15).

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Customs Clearance Issues	1	1	22.9362715	<.0001*
Interface / Data Error	1	1	5999.61697	<.0001*
Forecast Error	1	1	164.420136	<.0001*
Warehouse Quality Issues	1	1	12.1456859	0.0005*
Product Changes	1	1	199.417332	<.0001*
Manufacturing Quality Issues	1	1	10.2953561	0.0013*
Regulatory Issues	1	1	765.266306	<.0001*
Transportation / Distribution Issues	1	1	87.5205686	<.0001*

Figure 15 - Final Effect Likelihood Ratio Tests of LR2

As shown in the Parameter Estimates (Figure 16), the p-values (Prob>ChiSq) of all independent causes are less than 0.05. As a variable with Prob>ChiSq less than 0.05 is said to be significant as per PharCo's internal standard of confidence level, all of the variables in the final version of model LR2 are significant.

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-8.6895577	0.8058311	116.28	<.0001*
Customs Clearance Issues[1]	-1.2903859	0.2240542	33.17	<.0001*
Interface / Data Error[1]	-3.8387269	0.1267411	917.36	<.0001*
Forecast Error[1]	-1.2575315	0.1291304	94.84	<.0001*
Warehouse Quality Issues[1]	-1.0109685	0.2442203	17.14	<.0001*
Product Changes[1]	-1.9406583	0.1482377	171.39	<.0001*
Manufacturing Quality Issues[1]	-0.5821014	0.1759941	10.94	0.0009*
Regulatory Issues[1]	-2.0445658	0.12568	264.65	<.0001*
Transportation / Distribution Issues[1]	-1.3368235	0.1489452	80.56	<.0001*

Figure 16 - Parameter Estimates for the Final Version of LR2

Since all the categories are significant, odds ratio of each category (as shown in Figure 17 below) will be discussed as the indication of "predictability" in the Root Cause Assessment matrix.

Odds Ratios for Customs Clearance Issues

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.0757155	<.0001*	0.0323192	0.1917067
0	1	13.207328	<.0001*	5.2163007	30.941388

Odds Ratios for Interface / Data Error

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	10.487719	<.0001*	4.7485205	29.697168
0	1	0.0953496	<.0001*	0.0336732	0.2105919

Odds Ratios for Forecast Error

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.0808578	<.0001*	0.0470082	0.1301954
0	1	12.367389	<.0001*	7.6807624	21.272894

Odds Ratios for Warehouse Quality Issues

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.1323988	0.0005*	0.0533879	0.3750228
0	1	7.5529409	0.0005*	2.6665046	18.73085

Odds Ratios for Product Changes

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.0206237	<.0001*	0.0112319	0.0361415
0	1	48.488016	<.0001*	27.668995	89.032469

Odds Ratios for Manufacturing Quality

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.3121714	0.0013*	0.1559499	0.6277879
0	1	3.2033679	0.0013*	1.5928947	6.4123146

Odds Ratios for Regulatory Issues

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.0167538	<.0001*	0.009852	0.0265526
0	1	59.688035	<.0001*	37.661121	101.50187

Odds Ratios for Transportation / Distribution

Level1	/Level2	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
1	0	0.0690001	<.0001*	0.0375593	0.1216674
0	1	14.492728	<.0001*	8.2191315	26.624561

Figure 17 - Odds Ratio of LR2 stock-out categories

As shown in Figure 17, odds ratio for "Regulatory Issues" is the highest among the all categories, at 59.7, followed by "Product Changes" (48.5), "Transportation/Distribution Issues" (14.5), and "Customs Clearance Issues" (13.2). Hence, above-mentioned four categories have the

most predictable impact on stock-outs. But this is only half the story: as shown in the Sample of Quantitative Root Cause Assessment matrix (Figure 9), both the predictability and the frequency ratio need to be considered for us to assess the influence of any cause factors on stock-outs. In this research, the frequency is defined as the total number of each category reported in the PharSO system. Table 5 shows the distribution of odds ratio and frequency for all significant stock-out categories in logistic regression model LR2. The interface/data error category has the highest frequency, followed by forecast error, regulatory issues, and manufacturing quality issues. Meanwhile, regulatory issues category has both the odds ratio and frequency above the medians.

Table 5 - Distribution of Odds Ratio and Frequency

Stock-Out Category	Odds Ratio	Frequency
Customs Clearance Issues	13.2173	69
Interface/Data Error	0.0953	4847
Forecast Error	12.3673	1408
Warehouse Quality Issues	7.5529	86
Product Changes	48.4880	160
Manufacturing Quality Issues	3.2033	519
Regulatory Issues	59.6880	1134
Transportation/Distribution Issues	14.4927	310
Median:	12.7873	414.5

4.5 Quantitative Root Cause Analysis and Results

As we discussed in the Methodology section 3.3.2, a quantitative Root Cause Assessment matrix is designed to evaluate the influence of each variable causing stock-out issues. Using the odds ratio from model LR2 and the frequency number extracted from the dataset (table 5), we plot them into the matrix to get a root cause map.

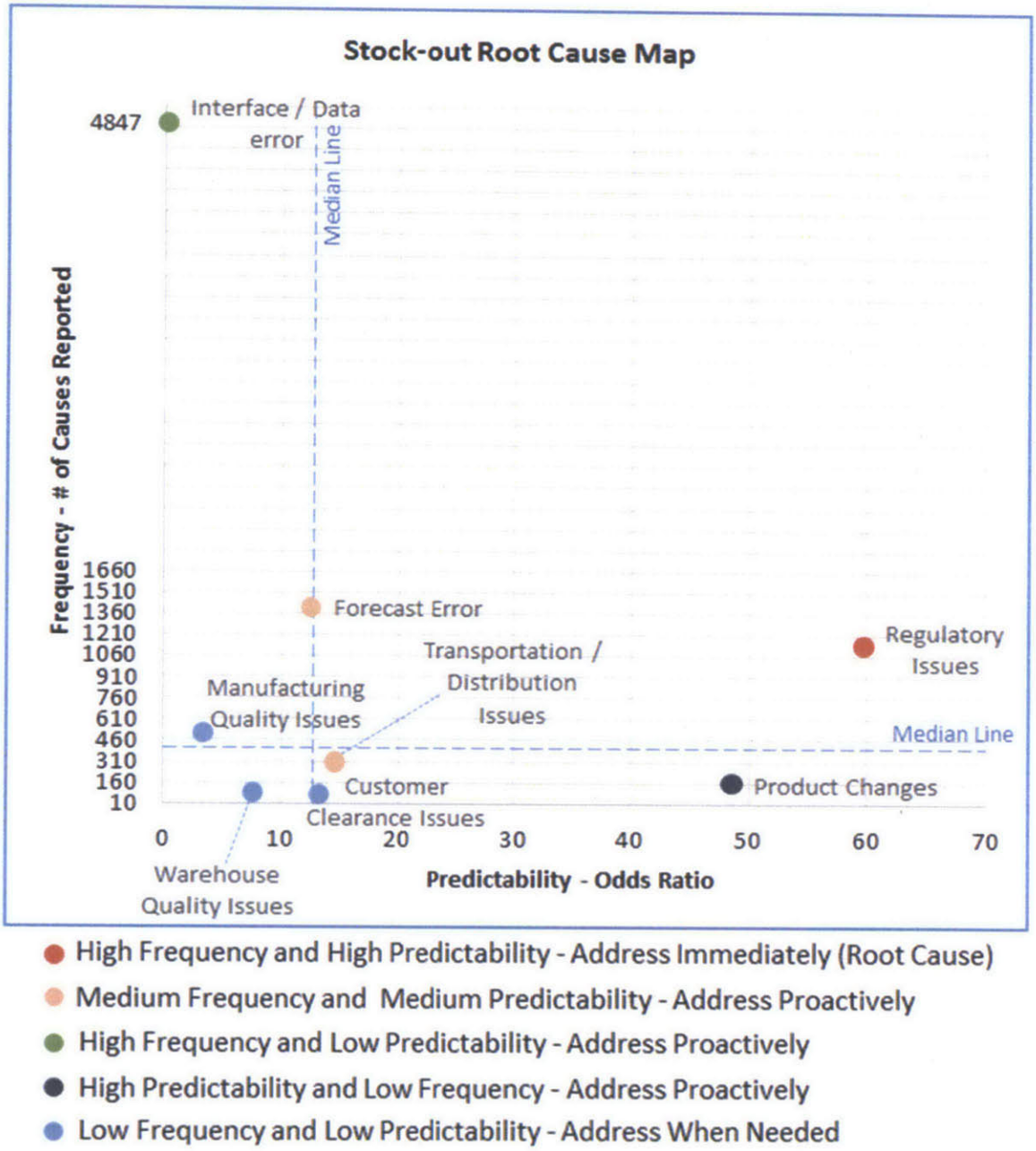


Figure 18 - Quantitative Root Cause Assessment Matrix for PharCo's Stock-outs

As indicated in figure 18, there are four stock-out cause quadrants divided by two median lines. The top right quadrant above both median values is considered as the root cause quadrant. There is only one stock-out category, regulatory issues (i.e. the lack of governmental regulatory approval to release products), with both predictability and frequency above the median values.

Since we define a root cause as one with both odds ratio and frequency above the median values, "Regulatory Issues" is considered a strong root cause for the stock-out problem in PharCo.

4.6 Limitations of This Study

There are three significant limitations to this study. First, out of the original 16,383 stock-out incidents in the raw data, 4,979 of them are categorized under "Not Applicable", "Not Defined", "Demand Unspecified", or "Supply Unspecified" categories. Since those categories do not reveal the nature of stock-out events and there is no comment from customers or suppliers for them, those data entries are eliminated from the analysis. Therefore, our study only explains the rest 70% (11,404) of stock-outs in the dataset.

Secondly, the dataset provided is dated within five months between October 31, 2012 and April 01, 2013. As we have no access to datasets with longer historical horizon, the analysis is limited within the selected samples. There are possibilities that these samples are not representative of the entire population and will hence cause sampling errors or biases.

Thirdly, since all root causes are entered by end users manually, the accuracy of our model is based on the assumption that users understood the definition of each category and categorized each stock-out incident correctly. This assumption might not be true in the real life and would require further study to validate.

5 Conclusion

This thesis approached the question of finding the root cause of stock-outs through a mixed methods approach that combines qualitative and quantitative techniques. With the manual coding methodology, we first coded the supplier/customer comments into meaningful root cause categories that we could use for further analysis. Then, we utilized logistic regression models to quantify the predictability of each category. Finally, we applied the quantitative root cause assessment matrix to measure the influence of each stock-out category and further define the root cause. Although PharCo had a strong suspicion that manufacturing quality may have been a major cause of stock-outs within the organization, through the in-depth analysis, we concluded that this hypothesis is not supported by the data. A strong root cause, with ~60 odds ratio and a relatively high frequency, is the category of regulatory issues. The most frequent regulatory challenges, among the others, are those associated with new product launches, license renewals, packaging changes, registration renewals, and formulation modifications (a sample of regulatory issues can be found in Exhibit 2 of the Appendix).

5.1 Recommendations

Our research indicated that the category of regulatory issues poses a major challenge preventing PharCo from improving their stock-out performance. Though this challenge needs immediate attention, it is difficult to develop a simple strategy to handle it. In fact, regulatory issues have been a major problem for the whole pharmaceutical industry for a long period of time. During the 2014 MIT Research Expo where we had the opportunity to present our research, a supply chain executive told us that in his thirty year career in the pharmaceutical industry, regulatory issues had posed the greatest challenge to limiting the number of stock-outs. According to him,

and also OECD (2000), the industry is heavily regulated from patent application, to market approval, commercial exploitation, patent expiration, and competition with generics (p. 7). The process of obtaining regulatory approvals can be very slow and costly, taking anywhere from months to years. Moreover, as different countries impose different regulations, it is difficult for pharmaceutical companies to design one type of strategy to comply with all regulations. Due to the complexity involved, we suggest PharCo initiate another research to fully investigate the factors that underlie the regulatory challenges. Upon further research, PharCo may have to invent new procedures to address the regulatory issues and reduce the stock-out level.

There are several other stock-out causes that we identified, such as the forecast error and transportation/distribution issues, which are not as critical as regulatory issues; however, they still contribute to a large number of stock-outs. For example, forecast error caused 1,254 stock-out incidents out of 4,917 incidents reported. Unlike regulatory challenges, these causes can be addressed within PharCo through proper action plans. There are many specialized tools or models that we suggest PharCo look into to improve their forecast accuracy and optimize their transportation / distribution network.

Several other categories, such as customs clearance issues and product changes, come with relatively low frequency but high odds ratio. This means that they did not contribute to the majority of stock-outs within PharCo; however, if they happened at all, there was a high possibility that the stock-out would ensue because of the high odds ratio (predictability) associated with them. We suggest PharCo take proactive actions to address these categories because they could bear serious consequences.

We also noticed that due to the Interface / Data Error, many false stock-outs were reported and later excluded. This reveals another opportunity for PharCo to improve their system. For example, manual input orders could have a wrong quantity that triggers the false stock-out alarm. In this case, PharCo could benefit from designing and implementing a proper training program to increase the data accuracy. In addition, some default stock-out categories in PharSO are vague and difficult to understand. This may have caused the misinterpretation for end users and further mis-categorized stock-outs in the PharSO system. If PharCo surveys the end users and customizes the categories per their feedback, it could reduce the errors significantly.

5.2 Further Steps

In order to validate the analysis and learn more about PharCo's supply chain, we plan to visit PharCo's secondary manufactory and conduct several mini case studies of their stock-out incidents. Through the case studies, we hope to gain more insight into PharCo's supply chain activities and help identify further research areas deriving from our thesis study. We also would like to gain the perspective of people who work in PharCo's supply chain. After all, it is the people rather than the systems that make things happen.

Bibliography

Andersen, B., & Fagerhaug, T. (2006). *Root Cause Analysis Simplifed Tools and Techniques* (Second Edition.). ASQ Quality Press.

Balasubramanian, P. (2009). *Root Cause Analysis-based Approach for Improving Preventive/Corrective Maintenance of an Automated Prescription-filling System* (Master Thesis). Binghamton University.

Crosby, P. B. (1984). *Quality Without Tears*. McGraw-Hill, Inc.

Gano, D. L. (2007). Comparison of Common Root Cause Analysis Tools. In *Apollo Root Cause Analysis - a New Way of Thinking* (Third Edition.). Apollonian Publications. Retrieved from http://kscsma.ksc.nasa.gov/Reliability/Documents/Ganos_comparison_the_Root_Cause_Analysis_Methods.pdf

IMS Institute for Healthcare Informatics. (2012a). The Global Use of Medicines: Outlook through 2016. Retrieved from http://www.imshealth.com/deployedfiles/ims/Global/Content/Insights/IMS%20Institute%20for%20Healthcare%20Informatics/Global%20Use%20of%20Meds%202011/Medicines_Outlook_Through_2016_Report.pdf

Juran, J. M., & De Feo, J. A. (2010). *Juran's Quality Handbook* (Sixth Edition.). McGraw-Hill, Inc.

Management Centre Europe. (2012b). New Opportunities & Strategies in the Pharmaceutical Industry (the Executive Issue), (No. 38). Retrieved from http://www.mce-ama.com/executive-issue-38-pharma-industry-2012

Masters, S. H. (2013). *Pharmaceutical Stock-out in Uganda, Kenya and Ghana* (Master Thesis). University of Washington.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2013). *Applied Multivariate Research* (Second Edition.). SAGE Publications, Inc.

Okes, D. (2009). *Root Cause Analysis, The Core of Problem Solving and Corrective Action.* ASQ Quality Press.

Organisation for Economic Co-operation and Development (OECD). (2000). *Competition and Regulation Issues in the Pharmaceutical Industry* (Roundtable Meeting Report No. DAFFE/CLP(2000)29). Retrieved from http://www.oecd.org/regreform/sectors/1920540.pdf

Powers, D. A., & Xie, Y. (2008). *Statistical Methods for Categorical Data Analysis* (Second Edition.). Emerald Group Publishing Limited.

Saldana, J. (2013). *The Coding Manual for Qualitative Researchers* (Second Edition.). SAGE Publications Ltd.

Shirazi, H. (2001). *Quality Assurance in Beauty Care and Over the Counter Pharmaceutical Contract Manufacturing* (Master Thesis). California State University Dominguez Hills.

Srinivasan, A. (2011, June). *Application of Information Technology and Statistical Process Control in Pharmaceutical Quality Assurance & Compliance* (Master Thesis). Massachusetts Institute of Technology.

The African Union Conference of Ministers of Health. (2007). Pharmaceutical Manufacturing Plan for Africa, (Third Session). Retrieved from http://www.pambazuka.org/actionalerts/images/uploads/Pharmaceutical_Plan-CAMH_MIN._8(III).pdf

The Association of the British (ABPI) Pharmaceutical Industry (ABPI). (2014). Making Medicines. Retrieved April 29, 2014, from http://www.abpischools.org.uk/page/modules/makingmedicines/makingmedicines5.cfm?age=Age%20range%2016-19&subject=Business%20studies

U.S. Department of Transportation, N. H. T. S. A. (2010). *Traffic Safety Facts Research Note* (Summary of Statistical Findings No. DOT HS 811 363). Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/811363.PDF

Wang, X. (2010). *Inventory Management in a Pharmaceutical Company: Minimizing Discard Practices* (Master Thesis). Massachusetts Institute of Technology.

Weinbach, R. W., & Grinnell Jr., R. M. (2003). *Statistics for Social Workers* (Sixth Edition.). Allyn & Bacon, Inc.

Appendix

Exhibit 1 – Original Stock-out Root Cause Categories in PharSO

Original Stock-Out Cause Categories
Demand - 3rd Party / Trading Partner
Demand - Allocation
Demand - Customs Clearance Issues
Demand - Data Error
Demand - Demand / Forecast Changes
Demand - Expired Stock
Demand - Import License Issues
Demand - Interface Error
Demand - Interface/Data Error
Demand - Pack Change
Demand - Product Changes / NPD Issues
Demand - QA Release
Demand - Quality Assurance / Documentation
Demand - Regulatory
Demand - Transport/Distribution
Demand - Unspecified
Demand - Warehouse Issues
No Applicable
Not Defined
Supply - 3rd Party / Trading Partner
Supply - 3rd Party Logistic Provider
Supply - Allocation
Supply - Data Error
Supply - Data Error / Interface Error

Supply - GSN Product Transfers only
Supply - Interface Error
Supply - Manufacturing / Production / Capacity Issues
Supply - Pack Change
Supply - QA Release
Supply - Quality Assurance / Documentation
Supply - Raw Material / Component Supply Issues from 3rd Parties (Non PharCo)
Supply - Raw Material / Component Supply Issues from PharCo Site
Supply - Regulatory
Supply - Transport / Distribution
Supply - Unspecified

Exhibit 2 – Samples of Regulatory Issues from PharSO Report

Sample of Regulatory Issues
Awaiting for the Certificate of Pharmaceutical Product (CPP)
Package Change / Wrong Package / Unplanned Promotion Packs
Formulation Change
Technical Review
Remediation Issue
Active Pharmaceutical Ingredient (API) Failure / API Supplier Not Authorized
New Product Launch or Introduction
Regulatory Changes
License Not in Place / Waiting for License Resolution / Import License Quota Issue / Sales License Expired
Regulatory Approval Delay
Pending Discussions with Local Regulatory Authorities
COFEPRIS's Requirement Validation (COFERRIS is a regulatory body of the Mexican government)
Expiration Date Change
Certificate of Analysis (CofAs) Not Matching Requirements
Leaflet Issue
Delay in Manufacturing Application Approval
Waiting Ministry of Health (MOH) Approval
Discrepancy Between Manufactured and Registered Formula
Waiting for Registration of Raw Material / Pending Registration Renewal / Registration Name Change / Re-registration
Release Process Issue / Issue with Quality Release
Regulatory Issue at Manufacturing Site
Handling Medicines to Be Checked
Product Variation Approval
Awaiting Approval from Internal Regulatory Team
Missing Manufacturing Certification from Regulatory Authorities

Manufacturing Site Change / Site Transfer
Source Change
Pending Update on Reimbursement
Stability Data Issues
Pricing Approval
Non-conformance and Administrative Issue
Product Transfer
Re-submission Approval Delay
Awaiting Confirmation from Sales Team